<table>
<thead>
<tr>
<th>QUERY NO.</th>
<th>Details Required</th>
<th>Author’s Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title of up to 50 characters that can be used instead.</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>Please confirm whether the authors and their affiliations given are correct.</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>Please confirm the statement “The author declares no conflict of interest” as conflict of interest statement. If there are any conflicts of interest to be declared, please provide.</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>Please indicate whether the personal communication stated in Table 2 was in an oral form or written form.</td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>Please confirm whether the deletion of ‘A.D.’ for references [3, 9, 35, 46, and 52] is OK.</td>
<td></td>
</tr>
<tr>
<td>Q6</td>
<td>Please provide the accessed dates for references [10, 32, 36, 54, 130, 133].</td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>Please confirm whether the insertion of ‘Cambridge University Press’ as the name of publisher for references [11, 12] is OK.</td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>Please confirm whether the deletion of ’72 refs’ in the article title for reference [16] is OK.</td>
<td></td>
</tr>
<tr>
<td>Q9</td>
<td>If there are fewer than 4 authors for reference [55], please supply all of their names. If there are 4 or more authors, please supply the first 3 authors’ names then et al</td>
<td></td>
</tr>
<tr>
<td>Q10</td>
<td>Please provide volume number for reference [56].</td>
<td></td>
</tr>
<tr>
<td>Q11</td>
<td>Please confirm whether the deletion of ‘[34 refs]’ from the article title of reference [74] is OK.</td>
<td></td>
</tr>
</tbody>
</table>
JOURNAL NAME: MLR
ARTICLE NO: 201472

<table>
<thead>
<tr>
<th>QUERY NO.</th>
<th>Details Required</th>
<th>Author’s Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q12</td>
<td>Please confirm whether the placement of the section ‘Definitions’ is OK as references have been cited.</td>
<td></td>
</tr>
<tr>
<td>Q13</td>
<td>Please note the given reference are not in order. We have renumbered the references both in text and reference list.</td>
<td></td>
</tr>
</tbody>
</table>
The Use of Patient-reported Outcomes (PRO) Within Comparative Effectiveness Research

Implications for Clinical Practice and Health Care Policy

Sara Ahmed, PhD,* Rick A. Berzon, DrPH,† Dennis A. Revicki, PhD,‡ William R. Lenderking, PhD,§ Carol M. Moinpour, PhD,|| Ethan Basch, MD,¶ Bryce B. Reeve, PhD,# Albert W. Wu, MD,** and on behalf of the International Society for Quality of Life Research

Background: The goal of comparative effectiveness research (CER) is to explain the differential benefits and harms of alternate methods to prevent, diagnose, and monitor a clinical condition or to improve the delivery of care. To inform decision making, information from the patient’s perspective that reflects outcomes that patients care about are needed and can be collected rigorously using appropriate patient-reported outcomes (PRO). It can be challenging to select the most appropriate PRO measure given the proliferation of such questionnaires over the past 20 years.

Objective: In this paper, we discuss the value of PROs within CER, types of measures that are likely to be useful in the CER context, PRO instrument selection, and key challenges associated with using PROs in CER.

Methods: We delineate important considerations for defining the CER context, selecting the appropriate measures, and for the analysis and interpretation of PRO data. Emerging changes that may facilitate CER using PROs as an outcome are also reviewed including implementation of electronic and personal health records, hospital and population-based registries, and the use of PROs in national monitoring initiatives. The potential benefits of linking the information derived from PRO endpoints in CER to decision making is also reviewed.

Conclusions: The recommendations presented for incorporating PROs in CER are intended to provide a guide to researchers, clinicians, and policy makers to ensure that information derived from PROs is applicable and interpretable for a given CER context. In turn, CER will provide information that is necessary for clinicians, patients, and families to make informed care decisions.

Key Words: patient-reported outcome, comparative effectiveness research, clinical care, health policy

(Comments; Medical Care 2012;00: 000–000)

A PATIENT-CENTERED CER

CER aims to explain the differential benefits and harms of alternate methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. Information generated from CER can help patients, clinicians, purchasers, and policy makers make informed decisions about which diagnostic, monitoring, and interventional approaches are superior under specific circumstances. Often, “conventional” endpoints such as disease-free survival do not provide all the information needed to understand the treatment effects. A comprehensive evaluation should incorporate patients’ perspectives of treatments, both in terms of patients’ actual experiences (eg, functional impact), and their judgments about the value of care (eg, access to services).

Countries around the world have made investments in patient-centered CER including Canada with the Strategy on Patient Oriented Research,1 and the work of the Medical Advisory in 2001 and Ontario Health Technology Advisory in 2003.2 Likewise, as far back as 1999, Australia, France, Germany, and the United Kingdom established agencies to ensure that their investments in health care, including medications, treatments, and new medical technologies, are resulting in improved outcomes and to assist health care providers in improving their clinical practice.3,4

More recently in the United States, CER has been at the forefront of discussions surrounding health care reform. The American Recovery and Reinvestment Act appropriated an initial $1.1 billion to fund CER under the 2009 stimulus law. The 2009 Patient-Centered Outcomes Research Act (S. 1213, Baucus D-MT)5 emphasized the goal of person-centered medicine, and on March 23, 2010, established the PCORI.6,8
Annual funding for PCORI began at approximately $500 million, with increases in subsequent years. Patient-centered outcomes research (PCOR) "is research that is informed by the perspectives, interests and values of patients throughout the research process, from the selection of research questions to the dissemination of research results. PCOR is intended to be practically relevant. Its real-world impact on patients is known and included in decisions about prevention, diagnosis and treatment."9

THE VALUE OF PROs WITHIN CER

PROs provide a standardized method of capturing patient perspectives and experiences. A PRO is "any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else."10 Information collected using PRO questionnaires allows outcomes that patients care about to be assessed, beyond only survival, and biomedical findings that are often similar between treatments.11 In prostate cancer, for example, there is no conclusive evidence regarding the relative survival advantages of the leading therapeutic options12; information about symptoms and side effects related to urinary, bowel, and sexual function, HRQoL, and overall satisfaction can help providers and patients identify interventions with the greatest benefits and minimal harms given the circumstances.

Data on subjective experiences of treatment benefit and harm are already accepted in the preapproval regulatory setting in the United States. The US FDA released guidance for the use of PRO measures for the approval of pharmacological products and devices.10 In the postmarketing and clinical practice settings, such standards are not well established and may differ from the FDA in terms of the information, endpoints, study designs, and results needed to ensure a patient-centered evaluation. In addition, PROs are not consistently being used in AIDS and cancer where we would expect to see their use, given that therapies come with the risk of substantial side effects.13 Moreover, there are unique challenges associated with implementing PROs in real-world settings that differ from the clinical trial setting.

In this paper we discuss the value of PROs within CER, types of measures that are likely to be useful in the CER context and criteria for selecting appropriate PRO measures, and key challenges associated with using PROs in CER. We recommend how to select and apply PROs in research, clinical practice, and population surveillance. These recommendations can help inform a monitoring infrastructure to evaluate the effectiveness of interventions at the local and population level. The scope and content of the paper was selected based on a consensus process among members of the International Society of Quality of Life (ISOQOL) Strategic Partnership Working Group and consultation and request for feedback on an initial draft of the paper from ISOQOL members.

TYPES AND APPLICATIONS OF PROs

Types of outcomes that fall under the PRO umbrella include symptoms, functional status, health perceptions, and HRQoL; other health-related constructs include satisfaction with care, access to care, perceived treatment benefit or harm, health behaviors, comorbidities, treatment adherence, and caregiver burden.14 The domains of HRQoL are reflected in the definition by Cell a15 as “the extent to which one’s usual or expected physical, emotional, and social well-being is affected by a medical condition and/or treatment.” Table 1 summarizes the constructs that can be assessed using PRO measures and examples of commonly used instruments.

Table 2 presents studies that compare interventions where PRO outcomes play an essential role in evaluating treatment effectiveness. Studies were purposefully selected to represent a range of interventions and clinical applications where a PRO was used to support (or not) the intervention of interest. For example, a drug for non-small cell lung cancer (Navelbine), which was initially approved for clinical use without inclusion of PRO data, received an approval for mention of quality of life in the label.36-39 In another study, a drug for the management of asthma (salmeterol/fluticasone propionate, a β agonist combined with a long-acting inhaled corticosteroid) was shown to have a positive effect in more domains of patients’ asthma HRQoL when compared with increased doses of inhaled corticosteroids.15,41 In another study, 3 antidepressant drugs were compared and while 1 drug (fluoxetine) was found to be associated with fewer patient-reported adverse events, this did not produce clear differences in HRQoL, depression, and other clinical outcomes.42 Finally, a Southwest Oncology Group study incorporated a novel comprehensive CER measurement approach, including PROs.51 The objective of this study is to evaluate whether the 21-gene Recurrence Score (RS) (RS refers to a number between 0 and 100 that corresponds to a specific likelihood of breast cancer recurrence within 10 y of the initial diagnosis) assay provides sufficient information to indicate whether chemotherapy—in addition to standard endocrine therapy—will benefit a patient at low risk for disease recurrence (RS ≤25, 1–3 positive nodes, hormone receptor–positive and HER2-negative breast cancer) compared with standard endocrine therapy alone. This question will be investigated in the context of clinical, cost, and PRO outcomes within both an observational study and a randomized trial.52

PROs are also valuable for evaluating comparative safety.53 The US National Cancer Institute has developed a measure specifically for such evaluations called the PRO-CTCAE (Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events).54,55 The PRO-CTCAE is a library of over 125 questions to assess adverse symptom events experienced by patients under treatment, such as anorexia, sensory neuropathy, nausea, fatigue, etc. Individual items that are relevant to a particular CER context can be selected from the PRO-CTCAE item library, to capture information on safety and tolerability from the patient perspective. This approach is particularly useful if treatment alternatives are expected to yield different toxicity profiles (eg, more fatigue or rash is expected with one intervention compared with another).

In clinical care settings, PRO measures themselves can also serve as a behavioral intervention to improve the quality of care a patient receives by influencing patient-physician
interactions and informing clinical decision making. PRO information can be used to focus the clinical visit on health concerns most problematic to their patients. As a surveillance system, patients can report their health status periodically through their home computer, smart phone, or other device (eg, interactive voice response system) and be monitored for changes in health that can expedite more or less clinical visits depending on how their health changes. PRO data can eventually be stored in patients’ electronic health record (EHR) and integrated with other clinical data about the patient.13

**IMPLEMENTING PROs AS EFFECTIVENESS OUTCOMES IN CER**

Selecting the most appropriate PRO measure given the proliferation of such questionnaires over the past 20 years is challenging. Success of a CER study, however, to inform decision making is dependent on the selection of valid and reliable questionnaires for a given application. Practical issues, such as the availability of translations for different languages and cultures, respondent burden, and copyright issues, are also important criteria in the choice of instruments.13

<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
<th>Example Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td>Measures that evaluate the frequency, severity, and impact of symptoms16,17</td>
<td>Brief Pain Inventory (BPI)18 MD Anderson Symptom Inventory (MDASI)19 Memorial Symptom Assessment Scale Short Form (MSAS-SF)19 SF-36 Vitality Scale20 Center for Epidemiologic Studies Depression Scale (CESD)22 Distress Thermometer23 McGill pain questionnaire24,25 Kidscreen26</td>
</tr>
<tr>
<td>Functional Status</td>
<td>Functional status measures assess a person’s ability to carry out daily activities such as walking, working, or attending social events14</td>
<td>Arthritis impact measurement scales27 Saint George’s respiratory questionnaire28 Functional Assessment of Cancer Therapy (FACT) core plus symptom modules29 European Organization for Research of Cancer Quality of Life Questionnaire Core-30 (EORTC QLQ-C30) plus symptom modules29,30 SF-3621,29</td>
</tr>
<tr>
<td>HRQoL</td>
<td>The extent to which one’s usual or expected physical, emotional, and social well-being is affected by a medical condition and/or treatment</td>
<td>World Health Organization Quality of Life Assessment, Brief Form (WHO QOL-BREF)29,31 EuroQol (EQ-5D)</td>
</tr>
<tr>
<td>Non-preference</td>
<td>These generic measures provide data on functioning relative to both a minimal and maximal level of performance for each health concept and can be used with any group of individuals</td>
<td>Short Form 6D (SF-6D) Health Utility Index (HUI) Quality of Well-Being scale</td>
</tr>
<tr>
<td>Preference</td>
<td>The relative value or utility weight assigned to each of the levels of health is assessed based on patient preferences</td>
<td>Health education impact questionnaire (heiQ) Simplified medication adherence questionnaire (SMAQ)33 Parent Adherence Report Questionnaire34 Consumer Assessment of Health Plans Surveys (CAHPS)35</td>
</tr>
<tr>
<td>Health behaviors</td>
<td>Health directed behavior The healthful behaviors individuals engage in aimed at disease prevention and/or health promotion</td>
<td>Healthful Adherence Questionnaire (HAQ)36 Healthful behavior questionnaire (HBQ)37 Healthful behavior scale (HBS)38 Healthful behavior index (HBI)39 Healthful behavior measure (HBM)40 Healthful behavior questionnaire (HBQ)41 Healthful behavior scale (HBS)42</td>
</tr>
<tr>
<td>Adherence</td>
<td>The extent to which the patient continues the agreed-upon mode of treatment under limited supervision when faced with conflicting demands, as distinguished from compliance or maintenance32</td>
<td>Simplified medication adherence questionnaire (SMAQ)33 Parent Adherence Report Questionnaire34 Consumer Assessment of Health Plans Surveys (CAHPS)35</td>
</tr>
<tr>
<td>Satisfaction with care</td>
<td>Patient satisfaction with care received</td>
<td>Satisfaction with care questionnaire (SCQ)36 Patient satisfaction questionnaire (PSQ)37 Patient satisfaction scale (PSS)38 Patient satisfaction measure (PSM)39 Patient satisfaction index (PSI)40 Patient satisfaction questionnaire (PSQ)41 Patient satisfaction scale (PSS)42</td>
</tr>
</tbody>
</table>

*Selected measures are only select examples and may assess more than one construct.
### TABLE 2. Examples of CER that Incorporate PROs as an Endpoint

<table>
<thead>
<tr>
<th>CER Question Posed</th>
<th>Intervention of Interest</th>
<th>Comparison Group</th>
<th>Timeline</th>
<th>PRO</th>
<th>Impact on decision Making</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study 1:</strong> What is the effect of Vinorelbine on health-related quality of life among individuals with stage IV non-small cell lung cancer?</td>
<td>Vinorelbine (Navelbine) compared with 5 fluorouracil plus leucovorin</td>
<td>5 fluorouracil plus leucovorin</td>
<td>Every 2 weeks for 2 months, and monthly for 2 additional months</td>
<td>Vinorelbine Quality of Life Questionnaire (variation of SWOG QOL questionnaire)</td>
<td>Initial FDA approval of Vinorelbine occurred without PRO outcomes. The label was later removed to indicate that quality of life was not adversely affected by Navelbine when compared with a control regimen.</td>
</tr>
<tr>
<td><strong>Study 2:</strong> What is the impact of inhaled salmeterol?</td>
<td>Advair</td>
<td>Pulmicort</td>
<td>12 weeks</td>
<td>Asthma Quality of Life</td>
<td>Intervention included in treatment guideline.</td>
</tr>
<tr>
<td><strong>Study 3:</strong> To compare the clinical, functional, and economic outcomes of budesonide on the health-related quality of life of patients with asthma</td>
<td>Fluoxetine</td>
<td>Imipramine or Desipramine</td>
<td>6 mo</td>
<td>SF-36 Health Survey, Hopkins Symptom Checklist-Depression Scale</td>
<td>Removed requirement for initial treatment with a TCA before treatment with fluoxetine</td>
</tr>
<tr>
<td><strong>Study 4:</strong> To evaluate outcomes of a chronic disease self-management program in a “real-world” setting</td>
<td>Chronic Disease Self-Management Program</td>
<td>Usual care at Kaiser Permanente</td>
<td>6 wk</td>
<td>Health Indicators (eg, health distress, self-rated health), Health behaviors, Self-Efficacy</td>
<td>Evidence of effectiveness has led to identifying mechanisms to make the program accessible Nationally and Internationally</td>
</tr>
</tbody>
</table>
| **Study 5:** Does the 21-gene Recurrence Score (RS) assay provide sufficient information to indicate whether is chemotherapy in addition to standard endocrine therapy beneficial to a patient at low risk for disease recurrence (recurrence score ≤25, 1–3 positive nodes, hormone receptor-positive and HER2-negative breast cancer) compared with standard endocrine therapy alone? | Main Group: Oncotype DX RS ≤25 (low-risk) randomized to Randomized trial: addition of chemo to hormonal treatment Women with RS <25  
Other groups:  
(a) Patients with RS scores >25 not randomized (receive treatment for higher risk disease)  
(b) Patients ≤25 who do not agree to randomization Both groups assessed pre/post-RS testing | First 2 PRO surveys: Women with ≤25 and >25  
Randomized trial: Hormonal treatment alone  
Pre/post-RS test surveys for those not on trial | 3 y for randomized study (PROs and Cost of medical care follow-up at 6, 12, and 36 mo)  
Randomized trial*: PROMIS Anxiety, Fatigue, Cognitive Function Concerns Short Forms (D. Cella, S. Garcia, J.S. Lai, personal communication, 2010),  
Decisional Conflict Scale,  
Assessment of Survivor Concerns,  
Oncotype DX testing questions,  
EQ-5D | Confirm whether or not women with low risk for recurrence need chemo in addition to hormonal therapy. Impact on a wide range of outcomes including cost of care & PROs |

---

*Secondary endpoint.

SWOG (formerly the Southwest Oncology Group) is a National Cancer Institute (NCI) supported organization that conducts clinical trials in adult cancers. SWOG is one of several cooperative groups funded by the National Cancer Institute to provide an infrastructure for the conduct of clinical trials of new cancer treatments as well as cancer prevention and control research (http://www.cancer.gov/cancertopics/factsheet/NCI/clinical-trials-cooperative-group).

Recurrence score is a number between 0 and 100 that corresponds to a specific likelihood of breast cancer recurrence within 10 years of the initial diagnosis.
the interpretation of results derived from PRO scores (Table 3). The selection of the PRO should be based on an a priori rationale driven by each of these parameters.

The CER goal or study question determines the range of outcomes that will help clinicians and patients select more effective treatments. The CER protocol should make explicit PROs included as primary or secondary outcomes. The objectives of the CER will determine whether the PRO will evaluate an important disease outcome (a common disabling symptom or limited function), or a treatment outcome (a comparison of side effects between a new drug and usual therapy). For example, in Table 1, study 1 aimed to evaluate the impact of Navelbine and a commonly used alternative (5 fluorouracil plus leucovorin) on the prevalence of symptoms, physical functioning, emotional functioning, and global quality of life. The stated objectives will also help determine whether to assess the impact of interventions on one or more domains such as emotional or social health, adherence, patient barriers to care, etc.; and whether a descriptive or preference-based (utility) measure (see the Selecting the appropriate PRO measure section) is appropriate, depending on whether an economic evaluation is included.

The patient population and intervention will also drive the PRO measurement strategy, questionnaires, and timing of evaluation. PRO selection needs to be based on the plausibility of the intervention having an effect on the outcome in the population of interest. Given, for example, the breadth and mechanism by which self-management programs are expected to improve outcomes43 (ie, improving long-term outcomes by improving individuals’ self-efficacy thereby improving skills to manage their disease and health behaviors such as adherence), studies that evaluate the impact of self-management programs often include measures of self-efficacy, health behaviors, and adherence.

Moreover, prior clinical observation or research in the given population should also be considered. These observations include evidence of appropriate psychometric characteristics, clinical perceptions of use, knowledge of feasibility of the PRO measure in a research or clinical setting, and understanding of public health consequences. Patients’ views concerning the importance and meaning of the domains addressed in the PRO measure should also be considered in measure selection.68 Most important is evidence that the PRO assesses the relevant level of health status and that those accessed for a given evaluation are representative of the target population. The Asthma Quality of Life Questionnaire (AQLQ), for example, has been extensively evaluated in a wide range of studies that assess pharmacological and behavioral interventions for asthma, across all levels of asthma severity.69–73

The decisions regarding selection of PROs with respect to analyses and interpretation may vary depending on whether the user is deriving CER information from research versus “real-world” clinical settings, as well as consideration of the timeline over which the evidence will be generated.74 Feasibility of integrating PRO questionnaires in the clinical or research setting including frequency of administration, respondent burden, and choices for mode of administration will influence the selection of the PRO measure. There may also be the need to create a measure of the relevant construct among the population of interest if not yet created and validated in that population. Caution is also needed in using measures, which have been validated for group-level comparisons, for assessing individual variations on the PRO. Individual level comparisons require a higher level of precision.75

### Strategies for Incorporating PROs in CER

Once the CER context is understood, the PRO domains to be assessed, including how they relate to other relevant outcomes, can be defined and a measurement strategy developed. Standardized and defined strategies for integrating PROs across clinical settings and research studies will facilitate CER in “real-world” situations. In turn, this will likely increase the uptake of findings in patient, clinical, payer, and health policy decision making.

### Identifying the Relevant Domains to Measure and Placing Them Within a Theoretical Framework

A conceptual model will help guide how the PRO and biomedical domains will interact with respect to patient outcomes. To ensure a patient-centered approach to evaluation, the framework would ideally highlight outcomes that are important to members of the target population. This may include, for example, the expected relationship between outcomes and access and adherence to care. A well-defined conceptual model will help to refine the measurement strategy and set priorities for selecting PROs that will be used to support treatment effectiveness versus those that will add to comprehensive evaluation.

© 2012 Lippincott Williams & Wilkins
Existing frameworks and taxonomies can help with the conceptualization phase. Examples include the International Classification of Functioning (ICF) (Fig. 1), which is a classification of the health components of functioning and disability.76 The goal of ICF is to provide consistent terms for describing and classifying health domains and health-related states and provides a universal model for health outcome measurement.77 The Wilson and Cleary model62 (Fig. 2) explains the relationships of clinical variables to measures of HRQoL. The model provides a theoretical approach to conceptualizing HRQoL as a multidimensional construct. There is overlap between the 2 models, although only the Wilson and Cleary model specifically includes general health perceptions and overall QOL.

Valderas and Alonso14 developed a model-based classification system that is a combination of the Wilson and Cleary Model and the ICF. The classification system proposes 3 main axes: the construct, population, and measurement. The construct can then be linked to specific ICF codes. For example, symptoms related to the respiratory system would receive a code of J00-J99. Similarly, mobility, which falls under the construct functional status, would receive a code of d410-d499. Once a clinician or researcher decides on the constructs of interest these can be linked to the classification model and PROs that have been mapped to these components of the model can be selected. Placing CER within known models, especially when mapped to common coding systems such as the ICF, may help consolidate information about the relative effectiveness of treatments, diagnostic, and monitoring approaches across research studies and/or clinical practice settings.

Selecting the Appropriate PRO Measure(s)

Once the context of the CER is characterized, and the domains and conceptual model linking them defined, questionnaires must be selected to measure the domains. There are many questionnaires that measure HRQoL and other PROs with distinctions in the way they are conceptualized, developed, and applied. There are also several families of PRO measures that can be characterized as either health profile or preference-based measures, and if they are generic or disease-specific measures. Each type will be briefly summarized below.

Health profile measures generally provide a broad range of health outcomes and include measures like the SF-36 and the WHO-QOL.79 The SF-36 includes multiple health domains, such as physical functioning, role-physical, general health, bodily pain, mental health, vitality, role-mental, and social functioning, and overall summary scores of physical and mental component. These health profile measures provide data on functioning relative to both a minimal and maximal level of performance for each health concept. The SF-36 provides norm scores relative to the US general population and age-adjusted and sex-adjusted reference values.

Preference-based or utility measures were designed to address the issue of importance of health status in a systematic way. Although these instruments may include a health classification system similar to those found in health profile measures, the relative value or utility weight assigned to each of the levels of health is assessed based on patient preferences. Combining the classification system and the preference weights creates an overall preference-based
HRQoL score. Common preference-based measures include the EuroQol (EQ-5D),\(^{80-86}\) Short Form 6D (SF-6D),\(^{87,88}\) Health Utility Index (HUI),\(^{89}\) and the Quality of Well-Being scale.\(^{90,91}\) Preference-based measures can also provide information about the impact of an intervention relative to the quality of life gained; and, quality-adjusted life-years (recommended by the Panel on Cost-Effectiveness in Health and Medicine\(^{22}\)), is one example of this type of outcome. The question of cost relative to benefit of treatment to the individual and population is an inevitable concern when expanding interventions on a wide-spread or population level and when making health policy decisions about where to invest scarce resources.\(^{93}\) Quality-adjusted life years provide a common currency to assess the benefits or burdens that patients experience in terms of quality and quantity of life.

The second categorization of PRO measures is whether they are generic or disease specific. Generic measures of HRQoL include broad domains and can be used across a wide range of healthy and chronic disease populations, and allow for comparisons across groups. They have often been used in population-based and health services delivery studies. Generic single measure and global assessments can play a particularly important role for comparisons that involve participants with multiple chronic conditions to provide an overall assessment of stability, improvement, or deterioration. Commonly used measures include the SF-36,\(^{21,78}\) and the WHO-QOL.\(^{94}\) and encompass preference-based measures as well such as Health Utilities Index Mark 3,\(^{95-97}\) EQ-5D,\(^{89,98}\) SF-6D,\(^{99,100}\) and Quality of Well-Being scale.\(^{91,101}\)

Clinical researchers observed that generic measures were not sensitive enough to capture changes in specific clinical populations. Consequently, numerous disease-specific measures have emerged that capture domains of relevance to a specific patient population. Examples for HIV include the MOS-HIV,\(^{102,103}\) HAT-QOL,\(^{104,105}\) WHO-QOL-HIV,\(^{106}\) and for cancer include the Functional Assessment of Cancer Therapy-General\(^{107-109}\) and the European Organization for the Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30).\(^{110}\)

A modular approach can be used to capture disease-specific outcomes while being able to compare outcomes across populations by adding a disease-specific component to a core generic measure. This type of model permits a comprehensive assessment of a patient’s health.\(^{111}\)

More recently, the National Institutes of Health (NIH) funded the development of the Patient-Reported Outcomes Measurement Information System (PROMIS; \(\text{http://www.nihpromis.org/}\)) that provides researchers access to standardized measures of health domains (eg, physical function, fatigue, anxiety, social satisfaction) that have undergone extensive psychometric testing.\(^{112,113}\) Central to PROMIS are its item banks, which include a comprehensive set of questions to assess each domain. Questions can be selected to generate a targeted short form for specific patient populations, such as an 8-item depression-specific measure. Alternatively, one can use PROMIS to assess a patient’s health status using computerized adaptive testing. Computerized adaptive testing automatically tailors the items administered to a specific patient based on his or her answers to previously administered questions, and provides an efficient, precise, and valid measure of a patient’s health status.

The selection of a tool should be based on the ability, appropriateness, and integrity of the instrument for assessing the domain of interest in the target population of the CER study. For descriptive measures, guidelines with standardized criteria for the evaluation of the quality of PRO questionnaires can be used to guide the selection of appropriate PROs. These include those developed by the Scientific Advisory Committee of the nonprofit Medical Outcomes Trust,\(^{63}\) and the COncensus-Based Standards for the selection of health status Measurement INstruments (COSMIN).\(^{64-67}\)

Further, the proportion of participants at the highest (ceiling effect) and the lowest (floor effect) possible score for a tool is also an important consideration. For example, there are well-documented ceiling effect problems with EQ-5D\(^{114}\) and floor effect problems with SF-36/SF-12.\(^{115,116}\) Ceiling effects may underestimate the impact of mild disease, and similarly, floor effects may underestimate the burden of severe disease.

### Analysis and Interpretation of PRO Data

In addition to considerations for analyses of any outcome, there are 2 particular considerations when analyzing PRO data. The first is that PROs, especially in the context of CER, are often administered at multiple time points to be able to characterize longitudinal changes. In such instances, missing data are often a problem, as over a prolonged period of time patients may experience morbidity or mortality due to disease or treatment. If missing data are related to the outcome being measured, then the resulting estimate of the effect of treatment may be biased. The same is also true for respondents who have a positive response to treatment and discontinue participation in the study, making the treatment seem less effective. Depending on the mechanism of the missing data, the appropriate analytic approach must be used for estimating change in the PRO.\(^{117}\) The prevention or minimization of missing outcome data is critical for CER studies,\(^{118}\) and the use of appropriate statistical techniques in advanced stage disease studies is very important due to the occurrence of substantial missing data.\(^{119}\)

Second, the interpretation of change scores must be considered carefully in longitudinal studies. Consensus about what constitutes a minimal important difference (MID) or a threshold that identifies a responder is needed to judge whether a statistically significant change reflects an important clinical improvement in the PRO.\(^{120}\) Ideally, an anchor-based approach would be used to assess MID and can be further supported using distribution-based approaches.\(^{120,121}\) Several studies that have used multiple approaches for determining an MID across different patient populations have estimated MIDs to be as low as 0.2 to 0.33 of an SD; however, without such evidence, many researchers feel comfortable using half an SD as an MID.\(^{122}\) An estimate of the MID has to be confirmed for a given instrument and patient population to identify a relevant threshold.\(^{121}\)
EMERGING CHANGES THAT FACILITATE CER USING PROS AS AN OUTCOME

Electronic Medical and Personal Health Records (PHRs)

The current implementation of internet-based PHRs that provide patients with access to their health information will facilitate and provide a means to monitor PROs to inform clinical care and CER initiatives. PHRs are available in the United States and Europe, and slowly gaining ground in Canada. The NIH’s Office of Behavioral and Social Sciences Research in collaboration with the National Cancer Institute and the Society of Behavioral Medicine are currently working to identify a core set of patient-reported constructs to include in the EHR.

The US Department of Veteran Affairs (VA) has more than 6 million active users of their EHR system. Multiple data sources linked through the EHR provide information for conducting medication effectiveness studies and have included evaluating the effect of statins on both Parkinson and Alzheimer diseases. Likewise, the Kaiser Permanente health network provides clients with access to parts of their individual health records with secure messaging between them and their care provider. Studies have shown more efficient and improved patient management using this system. The availability of PHRs that allow patients to enter symptom and HRQoL information in the EHR will maximize the potential for utilizing EHRs to inform CER analyses. These data can be entered from within the clinic or from the patient’s home and combined with other clinical and laboratory data to provide a comprehensive picture of a patient’s health status. The added benefit of electronic information systems that incorporate PROs as compared with the traditional paper-based interventions is in itself a CER question.

Registries

Hospital and population-based registries provide a mechanism for monitoring disease progression and patient responses to long-term disease management strategies. Some registries link clinical data to PROs. One example is the ORBIT-atrial fibrillation (AF) registry, a multicenter prospective outpatient registry of patients with incidental or prevalent AF to analyze treatment patterns and outcomes in the United States. PRO questionnaires will be administered to a subsample of approximately 1500 patients to assess outcomes including AF quality of life, anticoagulation treatment satisfaction, caregiver assistance, comorbidities, and adherence.

Use of PROs in National Monitoring Initiatives

Local health monitoring agencies have started incorporating PROs to inform health service decisions. In the United Kingdom, a national PRO measures program has been implemented to guide NHS decision making. As part of the NHS guidance, which started in April 2009, all licensed providers of NHS-funded Unilateral Hip replacements, Unilateral Knee replacements, Groin Hernia Surgery or Varicose Vein Surgery are mandated to ask patients undergoing one of these procedures to complete preoperative PRO questionnaires (called PROMs in the United Kingdom). Similarly, PROs are being used to compare health service providers in the United States, and increasingly in national population health surveys including the use of the HUI3 in the Canadian census survey.

Linking the Information Derived from PRO Endpoints to Decision Making

Many governments and patient advocacy groups are prioritizing their activities around interventions and services that are most likely to improve the HRQoL of individuals. Data to support CER may come from randomized clinical research trials, observational studies, clinical care settings, and the population; and each has strengths and limitations with respect to providing PRO data. Innovative approaches—such as linking administrative claims data and other forms of EHRs with PRO data—will take time to evaluate, as will other methods to collect, combine, or link information. Nevertheless, for value-based purchasing of health care services to succeed, the goal must be to make available better and more comprehensive evidence to inform decision making. Linking PROs to administrative databases and EHRs that feed into systems, such as the Adverse Event Reporting System designed to support the FDA’s post-marketing safety surveillance program for all approved drug and therapeutic biological products, can have a tremendous impact on safety and quality of care. PROs will also be critical for real world adoption of guidelines by tailoring recommendations to the patients’ self-reported level of symptoms or impact of disease.

It remains to be seen whether CER is likely to be used by insurance plans to prevent payment for specific health care services or to be used to set reimbursement rates, or both. An innovative model for Medicare use to CER to pay for services proposes that evidence be assigned to 1 of 3 categories, based on findings of superior, comparable, or insufficient comparative clinical effectiveness. Medicare would determine that a service should be covered and simultaneously assess its comparative effectiveness. PROs would feature prominently in this type of paradigm because of the need for information to determine from the patient’s perspective whether the new service is more effective and/or has fewer side effects compared with the relevant clinical standard for the treatment under consideration.

SUMMARY

In this paper, we have outlined recommendations for selecting and incorporating PROs that can be applied to research studies, clinical practice, and population surveillance. The recommendations presented for incorporating PROs in CER provide a guide for the selection of PROs that are applicable and interpretable for a given CER context. In turn, CER provide information that is necessary for clinicians, patients, and families to make informed care decisions.

ACKNOWLEDGMENT

The authors thank the members of ISOQOL for their review and feedback of initial drafts of the paper.

© 2012 Lippincott Williams & Wilkins
APPENDIX

Definitions

Health-related quality of life (HRQoL): “the extent to which one’s usual or expected physical, emotional, and social well-being is affected by a medical condition and/or treatment.”15

Patient-reported outcomes (PROs): PROs provide a standardized method of measuring the patient perspective and the term PRO is used “to include any outcome based on data provided by patients or patient proxies as opposed to data provided by other sources.”132

Patient-centered Outcomes Research Institute (PCORI): A nongovernmental agency created by the US Health Care legislation to establish research priorities and methodological standards for federally supported CER in the United States.133

Comparative effectiveness research (CER): CER is defined as the generation and synthesis of evidence that compares the benefits and harms of alternate methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care.40

REFERENCES


# Author Reprints

For **Rapid Ordering** go to: [www.lww.com/periodicals/author-reprints](http://www.lww.com/periodicals/author-reprints)

## Medical Care

### Order

<table>
<thead>
<tr>
<th>Author(s) Name</th>
<th>Title of Article</th>
<th>*Article #</th>
<th>*Publication Mo/Yr</th>
<th>______</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>______</td>
</tr>
</tbody>
</table>

*Fields may be left blank if order is placed before article number and publication month are assigned.

### Quantity of Reprints

- **Quantity of Reprints ____ $____**
- **Covers (Optional) ____ $____**
- **Shipping Cost ____ $____**
- **Reprint Color Cost ____ $____**
- **Tax ____ $____**
- **Total ____ $____**

### Reprint Pricing

<table>
<thead>
<tr>
<th>Reprint Pricing</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>50 copies</td>
<td>$336.00</td>
<td>100 copies</td>
<td>$420.00</td>
<td>200 copies</td>
<td>$494.00</td>
</tr>
<tr>
<td>300 copies</td>
<td>$571.00</td>
<td>400 copies</td>
<td>$655.00</td>
<td>500 copies</td>
<td>$732.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Shipping

- **Within the U.S. -**
  - $15.00 up to the first 100 copies
  - $15.00 for each additional 100 copies
- **Outside the U.S. -**
  - $30.00 up to the first 100 copies
  - $30.00 for each additional 100 copies

### Plain Covers

- $108.00 for first 100 copies
- $18.00 each add’l 100 copies

### Reprint Color

($70.00/100 reprints)

### Payment

- **Payment**
  - MC
  - VISA
  - Discover
  - American Express

<table>
<thead>
<tr>
<th>Account #</th>
<th>/</th>
<th>/</th>
<th>Exp. Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th></th>
<th>Dept/Rm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>State</td>
<td>Zip</td>
</tr>
<tr>
<td>Telephone</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Ship to

<table>
<thead>
<tr>
<th>Name</th>
<th></th>
<th>Dept/Rm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>State</td>
<td>Zip</td>
</tr>
<tr>
<td>Telephone</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use this form to order reprints. Publication fees, including color separation charges and page charges will be billed separately, if applicable.

Payment must be received before reprints can be shipped. Payment is accepted in the form of a check or credit card; purchase orders are accepted for orders billed to a U.S. address.

Prices are subject to change without notice.

For quantities over 500 copies contact our Healthcare Dept. For orders shipping in the US and Canada: call 410-528-4396, fax your order to 410-528-4264 or email it to Meredith.Doviak@wolterskluwer.com. Outside the US: dial 441 829 772756, fax your order to 441 829 770330 or email it to Christopher.Bassett@wolterskluwer.com.

**MAIL** your order to:
Lippincott Williams & Wilkins
Author Reprints Dept.
351 W. Camden St.
Baltimore, MD 21201

**FAX:**
410.528.4434

For questions regarding reprints or publication fees, **E-MAIL:** reprints@lww.com

**OR PHONE:**
1.866.903.6951

For **Rapid Ordering** go to: [www.lww.com/periodicals/author-reprints](http://www.lww.com/periodicals/author-reprints)