How Is Your Patient Really Doing?
Patient Reported Scales, Tools and Shared Decision Making

Lisa M. Shulman, MD
University of Maryland School of Medicine
Baltimore, MD

Outcomes Measurement

The quality of clinical care and clinical research depend upon reliable assessment of patient history and patient-reported outcomes (PRO's). Inaccurate patient history is likely to interfere with medical decision-making. And the quality of outcome measures is a "common denominator" resulting in limitations across clinical research. Data on health outcomes can be collected in 3 general ways: 1) PROs, 2) clinician-reported outcomes (CROs) and 3) physical and cognitive performance measures (e.g. timed gait speed, cognitive assessment, activity monitors). Health outcomes can be broadly divided into 3 categories: 1) Impairments (symptoms and signs; subjective and objective data), 2) Disability (performance of daily activities: subjective and objective data) and 3) Quality of Life (subjective perception of health and well-being; subjective data, by definition).

Are measures of impairments, disability and quality of life highly correlated with one another?
Studies show there is considerable variability in the correlations between these measures. In general, impairment correlates better with disability than with quality of life. Nonetheless, people with similar disease severity report different levels of daily function and quality of life. There are many reasons why these 3 outcomes diverge including medical co-morbidity, treatment side effects, age, mental health and socioeconomic factors. In clinical practice, patient-reported quality of life and daily function may trump level of impairment (symptom severity) in guiding management since preservation of quality of life is our primary objective in the setting of chronic illness. However, in clinical trials, discrepancies between outcomes often pose a problem. For example, *should a treatment be provided when symptoms are reduced but there is no improvement in disability or quality of life? Should we prescribe therapies when outcomes reveal statistical differences that aren't clinically meaningful?*

Common Impediments to Accuracy of PROs

All outcome measures- whether patient-reported, clinician-reported and even quantitative biometric assessments, have flaws and limitations. Over the years, my work with PROs has demonstrated many common barriers in the collection of sensitive and precise data.
Here are some examples:

- Discordance of subjective vs. objective reporting
    - In a study comparing subjective and objective ratings of performance of activities of daily living (ADLs), nearly half of study participants under-rated their disability compared to the observer rating
- Discordance of patient vs spouse reporting
    - Spouses regularly report the patient is more disabled than the patients report
- Effects of changes of symptoms
    - In Parkinson disease, patients *with and without* motor fluctuations report different levels of disability at best and worst function.

- In Parkinson disease, patient-reported disability and quality of life were responsive to symptomatic decline but not improvement. Therefore, it's more difficult to translate symptomatic improvement to improved disability and quality of life.

Tools: PROMIS®, Neuro-QoL, NIH Toolbox

Advances in measurement science and technology are resulting in new opportunities in outcomes assessment including measures with improved sensitivity, precision and practicality.

NIH initiatives including the PROMIS, Neuro-QoL and NIH Toolbox measurement systems resulted from the 2003 NIH Road Map strategic plan: *"The clinical outcomes research enterprise would be enhanced greatly by the availability of a psychometrically validated, dynamic system to measure PROs efficiently in study participants with a wide range of chronic diseases and demographic characteristics."* This is important since different investigators and chronic conditions tend to have different preferred measures for disability, depression, anxiety etc. resulting in the inability to compare results across clinical trials. These new tools also capitalize on technologic advances including computerized adaptive testing and digital data entry.

Instruments developed by PROMIS, Neuro-QoL and NIH Toolbox are all in the public domain and can be found at the HealthMeasures website (www.healthmeasures.net). These are comprehensive measurement systems that assess physical, mental and social health, symptoms, well-being and sensory, motor and cognitive function. PROMIS (The Patient-Reported Outcomes Measurement Information System) developed and validated numerous measures in populations with a broad range of chronic illnesses including chronic neurologic conditions. Neuro-QoL measures recapitulate a subset of PROMIS measures with additional validation in neurologic populations. In contrast, NIH Toolbox assesses healthy ageing with brief, comprehensive assessments in motor, cognitive, emotional and sensory domains. Notably, NIH Toolbox technical guides can be found at the HealthMeasures website with detailed descriptions of the instruments and tables with benchmarked performance by age and gender.

PROMIS responds to the 2003 NIH charge to measure across a wide range of chronic diseases with standardized formats and scoring systems. Scoring is based on t-scores with a mean of 50 and standard deviation of 10 based on extensive validation testing in US populations. Numerous domains of mental, physical and social function are all available in several formats: short forms with 4 to 10 items, computerized adaptive tests, and full item banks. A simple choice is the very practical PROMIS Profile with as few as 29 items (or also 43 and 57 item versions are available) that assess 8 key PRO domains (depression, anxiety, fatigue, sleep disturbance, physical function, satisfaction with roles, pain interference and pain intensity). All PROMIS measures are developed in the same rigorous standardized manner and calibrated with item response theory resulting in measures that are more precise and responsive, with reduced floor and ceiling effects.

Future Directions in Outcomes Measurement

Data collection by computer or tablet enables use of computerized adaptive tests (CATs) where precision is optimized by an algorithm that selects questions based on what is known about the examinee from previous questions. Whether patients complete PROMIS profiles, short forms or CATs, digital data collection enables a new level of integration of patient-reported data into clinical practice. PRO data completed before the office visit can be visualized by the clinician during the visit as either

cross-sectional or longitudinal data reflecting disease progression or response to treatment. Integration of PRO data in the EMR also has potential to respond to regulatory mandates. Clinical research and trials can be enhanced with PROMIS, Neuro-QoL and Toolbox measures.

Shared Decision-Making

Shared decision-making is a process where a patient and clinician, with more than one acceptable treatment option, jointly decide which option is best, based on current evidence and the patient's needs, preferences and values. Benefits of shared decision-making include respecting patients' wishes, improving care, and decreasing costs. Successful shared decision-making relies upon patient engagement and patient insight into their impairments, disability and quality of life. Patient-reported outcomes offer unique opportunities to raise our patients' awareness of their health, function and well-being, thereby increasing patient engagement and preparedness for shared decision-making.

Managing Big Data

Technologic advances in data collection and storage have enabled development of massive datasets with clinical and biologic data. The expanding scope of clinical and biologic markers of disease in large patient cohorts results in increasing complexity and quantity of data. These new opportunities have led to new challenges including: 1) visualizing informative patterns in massive datasets with multiple, complex domains, and 2) investigating disease progression based on diverse motor/cognitive/psychiatric domains and diverse biologic markers (genetics, imaging, serology, CSF). Traditional analytic approaches fall short in their ability to interrogate large multidimensional datasets. Novel tools and innovative methodologies are needed to efficiently sift through arrays of longitudinal data, identify patterns, generate hypotheses, and subject these hypotheses to rigorous statistical analysis.