Patient-reported outcomes (PROs) can be important measures of the impact and value of new drug treatments to patients. Recently, both multisector stakeholder groups and the U.S. Food and Drug Administration have carefully considered and issued guidance on best practices for the use of PROs in measuring treatment impact. When best practices are followed and PRO data are appropriately included in drug development strategy and clinical trials, these data can be part of the evidence submitted for drug approval and included in drug labeling. One study showed that PRO data were included in 30% of a sample of new drug labels and were more concentrated in certain therapeutic areas, such as anti-inflammatory agents, vaccines, gastrointestinal agents, and respiratory and urologic agents. PRO data included in labeling, or generated in a similar scientific manner, may often then be used in other communication vehicles, such as formulary submission dossiers, journal or direct-to-consumer advertisements, publications, or continuing medical education. Meaningful and reliable PRO results regarding the effects of new treatments on how patients feel and function provide useful information to those who must make decisions about the availability and utilization of such treatments.
be the focus of this paper and go well beyond getting the terminology correct. Key aspects will include: (1) identification of the appropriate role of PROs in drug development and early commercial strategy; (2) establishment of a clear conceptual basis for the PRO instruments to be used; (3) successful implementation of the PRO measures within the clinical trial program; and (4) effective and timely communication of the PRO study objectives and results, both with regulators and other stakeholders. An overview of how PRO results have been used in recent labeling for new drug products approved in the United States, as well as some specific examples, will also be given.

Identifying the Role of PROs in the Drug Development Program

As with any other aspect of a drug development program, PRO endpoints should be included only when there is clinical and/or practical medical value to those endpoints. Although clinical and practical medical value are closely related, the clinical value is derived from the information to be gained from PROs about treatment effects on symptoms and physiological measures per se, while the practical medical value is broader and includes the effects of treatment on patient well-being, as well as the relevance of this information to patients, payors, and others involved in decisions about treatment. Since practical medical value contributes to market access and uptake of a product, PRO endpoints are relevant to the early commercial strategy as well.

The value of PROs and their role in clinical trials in particular was outlined by the PRO Harmonization Group, an assemblage of stakeholders from academia, industry, and regulatory agencies, in a series of meetings in 2000-2002. The general value of the patient’s perspective was summarized along 4 lines, determined by whether the PRO was: (1) a unique indicator of the impact of disease, (2) essential for evaluating treatment efficacy, (3) useful for interpreting clinical outcomes, and/or (4) a key element in treatment decision-making. In other words, the PRO must be important for understanding the benefits of treatment for the patient. Furthermore, PROs can be “essential endpoints” for clinical trials when: (1) the patient’s self-report is the primary or sole indicator of disease activity; (2) the treatment may have a small impact on survival but may have a significant impact, positive or negative, on HRQOL; (3) the treatment may adversely affect patient functionality or well-being; (4) the treatment arms offer equal efficacy but differential PRO benefits; or (5) treatment-related decisions are based on a combination of objective and patient-reported subjective parameters. These describe the most common situations when PROs complement or replace other clinical endpoints in measuring the treatment benefits (or adverse impact) to the patient. In such situations, PROs can be useful, and perhaps instrumental, to regulators when deciding whether a new treatment should be approved for public use.

As with any other aspect of a drug development program, PRO endpoints should be included only when there is clinical and/or practical medical value to those endpoints.

The role of PROs in decisions about treatments in regular medical practice is the other major consideration in their selection for use in drug development programs. As is well-known, appropriate cancer treatment may be selected based on a tradeoff between likely improvements in survival, sometimes small, versus its impact, sometimes negative, on patient HRQOL. A less critical but nonetheless important example are the choices among the many arthritis treatments available, made in part based on pain relief and the effects on the patient’s ability to perform normal daily activities. An even more common example would be selection of allergy medication, which is often made based on degree of symptom relief versus drowsiness—both PROs. Thus, it is important for providers—physicians, pharmacists, nurses, and others—as well as patients to have meaningful and reliable information about these effects derived from clinical trial data.

PRO data can be useful to managed care decision makers in evaluating the mix of drugs needed on formulary to best suit the variety of their patients’ needs. In a given category it may be important to have both a drug with maximum efficacy and one that may be not quite as efficacious but more tolerable or “patient-friendly” to encourage adherence and to allow the physician some options in patient treatment (eg, chemotherapy, certain anti-infectives, and pain medications). There is some evidence that payors do use this type of information. In a survey of managed care medical and pharmacy directors, 54% thought
HRQOL information was “important” or “very important,” and 76% felt it would increase in importance. A similar survey found that, among factors considered in the drug benefit decision-making process, the HRQOL effects of the drug were more important than physician demand for the drug, rebate arrangements, and consumer demand.

Developing PRO Claims—Not for Amateurs

To provide meaningful and reliable information for regulators, providers, patients, and payors, PRO data must be generated in a scientific way. The scientific process includes appropriate endpoint and instrument (aka, questionnaire) selection—and, when needed, instrument development—as well as data collection, analysis, and interpretation. In February 2006, the U.S. Food and Drug Administration (FDA) issued a 36-page draft guidance—not yet finalized at the time of this writing—that lays out much of the scientific process for generating PRO data, which the FDA expects sponsors to follow before such data can be included in product labeling.

The first step in this process is ensuring that the PRO endpoints used properly capture the aspects of treatment benefit that are most relevant to the patient, an exercise that begins with identifying the specific concepts of patient well-being that are both impacted by the disease and likely to be affected by treatment (Figure 1). For example, in the treatment of osteoarthritis, specific concepts considered important often include a “simple” one such as pain, as well as a more complex one such as limitations on activities of daily living. These PRO endpoints generally should be logical counterparts to other clinical or physiological endpoints; the interrelationships of all these endpoints may be described in what is called an “endpoint model.” In some cases, however, the PRO measures may be stand-alone primary endpoints if other types don’t apply; for example, migraine treatments rely primarily on PRO endpoints, since making objective measurements of pain and other symptom relief is difficult. When the PRO endpoint is of the more complex variety (ie, consisting of a summary score based on a number of individual questions), it needs to have an underlying “conceptual framework” that relates how the simpler concepts explored in the individual questions (eg, dressing oneself or bathing oneself) combine to form a more complex concept (eg, activities of daily living) represented by the summary score.

Proper delineation of the relationships among concepts and endpoints is important both to good measurement and to clear interpretation and communication of treatment benefit.

Asking patients about specific concepts relating to how they feel or function is not quite as simple as it may seem. PRO instruments must go through careful development based on input from patients with the relevant condition and characteristics, item (question) selection and testing, piloting, psychometric testing for several aspects of reliability and validity, as well as responsiveness to treatment effects, and then perhaps modification and re-testing. The draft FDA guidance depicts this process in a “wheel and spokes” diagram (Figure 1). PRO instrument development can be an expensive and time-consuming process, and sponsors will generally first seek to use existing, well-validated instruments in their programs. If no existing instrument applies precisely to the target population or nature of the treatment benefit, the second choice would be to modify an existing instrument, although some additional validation work must then be conducted. As a last resort, a new instrument may need to be developed from scratch—to be validated before phase 3 trials, this development must begin early during human clinical trials. Recently the FDA has been applying increasingly rigorous standards for PRO instruments, such that instruments previously thought to be validated may no longer pass muster.


Figure 1. The Process for Generating PRO Data

Identify concepts & domains that are important to patients.
Determine intended population and research application.
Hypothesize expected relationships among concepts.

i. Identify Concepts & Develop Conceptual Framework
Identify meaningful differences in scores. Finalize instrument formats, scoring, procedures & training materials.
Refine instrument & procedures.
Draft procedures for scoring & administration.
Pilot test draft instrument.

ii. Create Instrument

iii. Assess Measurement Properties
Assess score reliability, validity, and ability to detect change. Evaluate administrative & respondent burden. Add, delete, or revise items. Identify meaningful differences in scores. Finalize instrument formats, scoring, procedures & training materials.

iv. Modify Instrument
Change concepts measured, populations studied, research application, response options, recall period or method of administration.

PRO

Conceptual Framework


Developing PRO Claims—Not for Amateurs

To provide meaningful and reliable information for regulators, providers, patients, and payors, PRO data must be generated in a scientific way. The scientific process includes appropriate endpoint and instrument (aka, questionnaire) selection—and, when needed, instrument development—as well as data collection, analysis, and interpretation. In February 2006, the U.S. Food and Drug Administration (FDA) issued a 36-page draft guidance—not yet finalized at the time of this writing—that lays out much of the scientific process for generating PRO data, which the FDA expects sponsors to follow before such data can be included in product labeling.

The first step in this process is ensuring that the PRO endpoints used properly capture the aspects of treatment benefit that are most relevant to the patient, an exercise that begins with identifying the specific concepts of patient well-being that are both impacted by the disease and likely to be affected by treatment (Figure 1). For example, in the treatment of osteoarthritis, specific concepts considered important often include a “simple” one such as pain, as well as a more complex one such as limitations on activities of daily living. These PRO endpoints generally should be logical counterparts to other clinical or physiological endpoints; the interrelationships of all these endpoints may be described in what is called an “endpoint model.” In some cases, however, the PRO measures may be stand-alone primary endpoints if other types don’t apply; for example, migraine treatments rely primarily on PRO endpoints, since making objective measurements of pain and other symptom relief is difficult. When the PRO endpoint is of the more complex variety (ie, consisting of a summary score based on a number of individual questions), it needs to have an underlying “conceptual framework” that relates how the simpler concepts explored in the individual questions (eg, dressing oneself or bathing oneself) combine to form a more complex concept (eg, activities of daily living) represented by the summary score.

Proper delineation of the relationships among concepts and endpoints is important both to good measurement and to clear interpretation and communication of treatment benefit.

Asking patients about specific concepts relating to how they feel or function is not quite as simple as it may seem. PRO instruments must go through careful development based on input from patients with the relevant condition and characteristics, item (question) selection and testing, piloting, psychometric testing for several aspects of reliability and validity, as well as responsiveness to treatment effects, and then perhaps modification and re-testing. The draft FDA guidance depicts this process in a “wheel and spokes” diagram (Figure 1). PRO instrument development can be an expensive and time-consuming process, and sponsors will generally first seek to use existing, well-validated instruments in their programs. If no existing instrument applies precisely to the target population or nature of the treatment benefit, the second choice would be to modify an existing instrument, although some additional validation work must then be conducted. As a last resort, a new instrument may need to be developed from scratch—to be validated before phase 3 trials, this development must begin early during human clinical trials. Recently the FDA has been applying increasingly rigorous standards for PRO instruments, such that instruments previously thought to be validated may no longer pass muster. In particu-
lar, recall periods commonly used in questions (e.g., “Over the past week...”) have been a subject of scrutiny. The FDA has shown a strong preference for short recall periods as a guard against recall bias; however, only a limited amount of research has been done on the extent to which recall bias affects the comparative estimates of treatment effect. This regulatory pressure, together with the increasing availability and feasibility of electronic patient diaries, has triggered a new round of development of instruments better suited for use in such circumstances.

Because of the importance of regulatory judgments on the suitability of PRO instruments and results for use in product labeling, regular interaction with the FDA about the PRO strategy and studies during the course of the development program is advisable. This interaction should include discussion of the potential language to be used in the product label if the studies are successful. Whereas the FDA therapeutic area review divisions retain the ultimate authority for the labeling of individual products, the FDA Study Endpoints and Labeling Development Team is the focal point of the agency’s PRO expertise and will often be involved in consultations concerning PRO endpoints. These interactions may start as early as phase 1, when human trials first begin, through phase 3 and the new drug application (NDA), and extend into post-marketing periods. The end-of-phase-2 meeting is generally an important point for the sponsor to review the plans for PRO endpoints in detail with the FDA. At each stage of interaction, including the final NDA submission, the sponsor typically submits a briefing document or dossier with the background on the PRO instruments used and any data available to date; a proposed outline for this dossier is expected to be included with the final FDA PRO guidance.

Diligent implementation of PRO studies in clinical trials is also crucial. Without reasonably complete, properly collected data, even the best PRO instrument cannot yield reliable and meaningful results. Planning the logistics of the data collection during or between patient study visits, training the investigational site staff on administration procedures, early and ongoing communication with sites and monitoring of the success of the data collection, and quick remediation of problems are the hallmarks of successful studies. Even though the inclusion of PRO data in clinical studies has become relatively common, study sites may still consider such data to be of secondary importance and not place sufficient priority on good data collection procedures unless proactively managed by the sponsor. Inadequate attention to these concerns often results in too much missing or poor quality data and has resulted in the failure of many PRO studies.

Given valid, concept-based instrumentation and complete, high-quality data, the third major step toward a PRO claim is defensible analysis and interpretation of the data. Major endpoint hypotheses and analysis methods are declared in the original protocol; a full statistical analysis plan for the PRO data must also be established, preferably at the beginning of the study and absolutely before the data blind is broken. Key issues to address in this plan typically include, but aren’t limited to, approaches to multiple endpoint testing, missing data (hopefully minimal but rarely zero), and clinically meaningful differences. Since different approaches to these issues exist and can yield different results, it is important to prospectively declare the approaches to be used to avoid the appearance, or reality, of data-mining. Assuming successful studies and an approvable drug, the final step is negotiating final PRO label language with the FDA. Although draft language should have been discussed earlier in the development program, the FDA will want to ensure that the final labeling to be used with the general public appropriately communicates the concepts tested and the strength of the data, based on current standards for PROs. Shifting standards can be, and have been, an issue here since PRO study decisions made early in a development program may not yield results for 5 years or more. The issuance of the draft guidance was a big step forward in that it solidified a number of standards that had been evolving for some time. Even though not all current issues are resolved there—the science is always evolving—it gives sponsors a much clearer basis for planning PRO studies during development. Standards for PRO data have essentially become at least equivalent to those applied to more traditional clinical data. These developments should give users of that information—the
providers, patients, and payors—a greater degree of confidence that PRO claims made for new drugs are meaningful and reliable and represent treatment benefits of value to patients.

**Actual PRO Claims**

PRO data are not new to drug development and labeling. In studies for pain medications, for example, patients have always had to be asked about how much pain or pain relief they experienced, and those results were included in labeling as efficacy measures. (It should be noted that listings of side effects in labeling as evidence regarding safety are often patient-reported but have not been the focus of this discussion.) Broader health status measures also have a long history in both clinical trials and health policy, but it was not until about 20 years ago that sponsors began to regularly consider collection of such broader measures in clinical trials and as having potential for labeling claims; the first apparent HRQOL claim appeared in a U.S. label in 1989 (a number of HRQOL endpoints, for erythropoietin alfa). Since that time, a variety of simple and more complex PRO claims have been cited in labeling.

In a study of 215 new drugs approved in the United States from 1997 to 2002, we found that 64 of these products (30%) included PRO data in the clinical trials section of the label as a measure of treatment benefit (Figure 2). As can be seen in Figure 2, PROs are not as commonly used as more traditional clinician-reported outcomes (eg, mortality, fractures, tumor response, the Unified Parkinson’s Disease Rating Scale) or laboratory tests and device measurements (eg, FEV₁, HbA1C, blood pressure); however, 23 of these drugs are used only in PRO endpoints as measures of treatment benefit, including antimigraine products and some drugs for pain relief, antiepileptics, antiflu, and 3 drugs for allergic conjunctivitis. Other types of drugs that relied heavily, but not exclusively, on PRO endpoints included anti-inflammatory agents, gastrointestinal agents, vaccines, and respiratory and urologic agents.

In some cases, these PRO endpoints were relatively simple patient-experienced events (eg, patient recall of angina attacks, partial epileptic seizures, and days with bowel urgency). Also common were symptom ratings—simple questions about influenza symptoms, ocular itching, bloating, headache severity, nausea, and so on—generally in a Likert scale or visual analog scale format. Responses for related symptoms were sometimes combined to get a symptom score, such as in asthma, pruritus, or allergic rhinitis. Such endpoints are not considered HRQOL measures but, given their subjective nature, may be affected by the same validity issues as more complex scales and thus have an important place in the PRO spectrum. More formal, complex PRO scales were represented in 16 of these labels, including both disease-specific scales (eg, Walking Impairment Questionnaire, WOMAC Osteoarthritis Index, and International Index of Erectile Function) and more general health status measures (Short Form-36 and Sickness Impact Profile). The most common scale used in 6 labels of arthritis drugs was the Modified Health Assessment Questionnaire, a physical function and health perception instrument particularly relevant to that disease.

No comprehensive assessment of PROs in labels since 2002 is available, but an examination of labels of recently approved drugs shows that their use continues. The label for cilengitide, a treatment for allergic rhinitis, says that patients treated with it “exhibited statistically significantly greater decreases in total nasal symptom scores than placebo-treated patients.” The label for varenicline, a smoking cessation product, indicates that it “reduced urge to smoke compared to placebo in all studies.” The label for ecotizumab, a product for primary nocturnal hemoglobinuria, includes the statement “after 3 weeks … patients reported less fatigue and improved health-related quality of life.” These 3 examples range from a very disease-specific score important to efficacy measurement to a general...
HRQOL claim; the varenicline example is unusual in that “reduced urge to smoke” is related to how the drug helps achieve the ultimate outcome of treatment rather than being the ultimate outcome per se.

Communicating the Value of Treatment-Reported Outcomes

The presence of the PRO claim in the approved label is important for 3 reasons: (1) it indicates that the FDA judged the information to be valid, reliable, and worthy of mention; (2) it makes the results widely available in a public document; and (3) it enables further use of the data in a variety of postlaunch communications. Such postlaunch communications are often intended for a variety of stakeholders and may include formulary submission dossiers, journal or direct-to-consumer advertisements, publications, or continuing medical education. They may contain more detail about the PRO results than present in the label per se (eg, results of individual items in a scale), as long as they are consistent with the labeling and are scientifically supportable. By reaching a wider audience, with more relevant details, the practical medical value of the PRO results can be more completely realized via these communication vehicles than through the label alone.

Under some conditions, PRO data can still be utilized even if it doesn’t appear in the original approved label. Postlaunch communications about marketed drugs, outside of scientific publications per se, are regulated by the FDA Division of Drug Marketing, Advertising, and Communications (DDMAC). DDMAC may deem that off-label PRO data are appropriate for use in communications if supportive data are brought forward that were not present in the original NDA, such as subsequent validation work, or data from phase 3B or phase 4 trials. In cases where development and validation of a new instrument cannot be completed before the phase 3 trials, or where the phase 3 trials reveal a potential PRO effect that must be confirmed with further trials, DDMAC-sanctioned use of nonlabel PRO data may be the most feasible route.

For those wishing to obtain more information about PRO results from a drug development program, several options are available. The most direct method is to contact the Medical Information Department of the drug sponsor, which is responsible for responding to external information requests. Many of the results will also be published in the scientific literature, either as conference abstracts or full journal publications. PRO results are also sometimes included in documents posted on ClinicalTrials.gov or ClinicalStudyResults.org.

In summary, PRO endpoints included in drug trials may capture treatment benefits that can only be measured by obtaining patient input. They can translate the effects of a new drug from the very disease-specific, objective clinical measures traditionally used in trials to everyday concepts that are meaningful to patients and those who care for them. They can help make decisions about whether a treatment should be started, or about which treatment should be chosen. Recent regulatory guidance for, and enforcement of, standards for PRO evidence, together with the increased attention by sponsors to well-considered use of PROs in drug development, allows patients, providers, and payors to be confident that PRO data included in product labeling or seen in promotion are meaningful and reliable. When properly communicated, PRO results can help the full value of medical treatments to patients be realized.

Acknowledgments: The thoughts expressed in this paper have benefited from many discussions in recent years with Penny Erickson, Laurie Burke, and several Pfizer employees and other colleagues; however, none of them bear any responsibility for what is written here, nor does it reflect an official position of Pfizer, Inc.

References

AHDB Stakeholder Perspective

When the American public rebelled against managed care in the late 1990s, it was demanding a healthcare delivery system that addressed their specific, or personal, healthcare preferences. Today there is terminology to describe this goal: personalized medicine. The public outcry against managed care is a useful backdrop for examining the phenomenon of patient-reported outcomes (PROs), which can be viewed as a part of personalized medicine.

PROs offer both challenges and opportunities to various stakeholders. Physicians incorporating PROs into their treatment strategies may see improved patient satisfaction, but the benefit comes with some costs: first education on the process, then obtaining PRO data, and finally interpreting it. However, if PROs put the clinical picture in full relief, it will be attractive to providers and their patients alike. Obviously, physicians must be given incentives to learn PROs, including increased compensation from plans.

If physician interest in PROs is the pursuit of quality, payors have a higher hurdle to climb because they do not practice medicine but manage overall resource allocation. Their concern is the pursuit of value, which is a composite of quality, cost, and access. Hence, payors need objective evidence that their professional interests are served in implementing PROs in order to face their more complex learning curve involved in incorporating them into their corporate paradigm: education, implementation, and then interpretation of PRO data.

Large employers and CMS have similar incentives to pursue personalized medicine, as both want healthy patients. But again, they must be able to finance it and establish value. Funding must be secured for CMS to pursue PROs.

Manufacturers will pay more to incorporate PRO data into their labeling, but the resulting product will have a more compelling value proposition than traditional safety/efficacy labeling. Demonstrating value of new drugs is essential in today's climate, and in appropriate disease states PROs can form an important part of the value proposition.

The basic tenet of PROs is that healthcare is a multilateral process, not a unilateral one. It involves really involving the patient into the process. The public has been criticized for not appreciating the value proposition of healthcare intervention. Perhaps PROs are a means of integrating patients into the decision-making process, and finally achieving that elusive goal of being a responsible healthcare copartner.