

Task Force Report of the Patient-Reported Outcomes Harmonization Group: Too Much Harmony, Not Enough Melody?

William Lenderking, PhD

Pfizer, Groton, CT, USA

The Ad Hoc Task Force of the Patient Reported Outcomes Harmonization Group had an important opportunity to interact with the FDA and move toward consensus on unresolved issues within the field. The Task Force is to be commended for building bridges to the FDA and working for several years to increase the level of regulatory acceptance of outcomes research. No doubt these insights will be incorporated in the forthcoming FDA guidance on patient-reported outcomes. However, the initial report of the Task Force, published in this issue of *Value in Health* [1], has failed to communicate the value of these meetings. The report does not mention the areas of controversy and debate that led to the need for meeting with the FDA, and deliberately avoids providing prescriptions or recommendations, fearing they would inhibit the growth of the field.

Whereas it is entirely appropriate for the Task Force to limit the relevance of their comments to the drug approval and regulatory process, I disagree that a set of prescriptions would limit our field. In fact, standards for the assessment of subjective variables already exist, based on literally decades of research by clinical and educational psychologists [2]. The difficulty of applying these guidelines to the drug development process should not obviate the need for doing so. There are many places where the Task Force could have further developed a given issue, for example, the issue of patient versus proxy reports. Patient and proxy can provide a different perspective, and in some instances, there is no choice but to use proxies, i.e., when cognitive development or impairment prevents the reporting of valid self-referential opinions. As the Task Force rightly points out, the proxy is more biased the less observable the behavior or characteristic of the patient—so at what point do we stop collecting the proxy reports, or how do we interpret them? Can we use proxies to identify drug benefits? The report raises these issues but does not provide any recom-

mendations. The issue of multidimensionality is another issue the Task Force could have provided more guidance on. While it is true that the requisite number of domains to cover the concept of health-related quality of life will vary from one disease to another, would it not have been possible to make a simple recommendation about minimal coverage required?

Although there are many active debates going on within the field of outcomes research, to which the Task Force alludes, they seem to have deliberately avoided these debates and shied away from making any statements that might help resolve the questions. For example, what is the consensus about what constitutes sufficient evidence for validation of a new scale? Validation is an ongoing process, but at what point can we agree that sufficient evidence has accumulated that would justify using a particular scale for a regulatory submission? In diseases where there are multiple scales available to measure outcomes, how does a researcher select the most appropriate measure and avoid bias? How much needs to be known about a scale before it can be used to guide treatment? What would the authors recommend around clinical trial designs which incorporate patient-reported outcomes? Should power calculations be universally provided for such endpoints, so that results can be interpreted in the context of the power of the instrument to detect changes or treatment effects? What is the group's opinion about the use of item banks and computerized adaptive testing during clinical trials for product registration? Has the technology progressed to the point where it can be considered comparable or better than standard instruments? Should the FDA and other regulators accept results based on such technology? While it may not be possible yet to endorse a specific model of health-related quality of life above all others, which models were considered? Certainly, there are excellent models available that illustrate the integration of subjective and clin-

ical outcomes [3,4]. A discussion of clinical significance, and commentary on the enormous problem of missing data, is totally lacking. Should data on responders only analyses be used to evaluate the effects of drugs? The authors provide examples of these sorts of results without discussing the bias potentially inherent in such analyses. Guidance from this Task Force on these issues would have been very helpful.

The distinction between “clinical” and “patient reported” outcomes is unfortunate and contributes to the conceptual lack of clarity that has plagued outcomes research. Certainly this cannot be blamed on the Task Force. The danger lies in confusing a specialized methodology with the concepts it is used to measure. However, the Task Force has not made any recommendations about improving this situation, and instead has made statements such as to collect patient-reported outcomes when “the treatment arms offer equal clinical efficacy but differential PRO benefits.” What this statement means is that patients might perceive differences between treatments that would otherwise be undetectable. This is a much more precise statement. The recommendation that outcomes endpoints should be treated by the same standards as clinical endpoints is essential, though hardly original.

Whereas operational definitions and measurement are cornerstones of the field, the main purpose of our research should be to demonstrate relationships among constructs, evaluate treatment effects, improve the health of samples or populations of patients, and allow us to make predictions. We seem to be lost in the forest of definition and measurement, and need to do more to show the clinical benefits of incorporating the patients’ perspective. The moniker “patient-reported outcomes” is unfortunate even though it has helped improve the acceptance of the field by FDA, primarily because it is inaccurate (it also includes outcomes provided by clinicians and proxy respondents), and also, one might argue, because it contains no information about the content of the field. If no theoretical content should be alluded to in our label (i.e., we moved away from describing our field as “quality-of-life measurement” because of the endless discus-

sions about definitions), perhaps a term such as “perceived clinical outcomes” would be preferable as more accurate, because it implies that the variable being measured is perceived by a person, whether a clinician, patient, or proxy. The Task Force laments the lack of a “clearly developed conceptual framework for understanding the relationship between HRQL and PROs,” but this makes no sense, as the term “patient-reported outcomes” has no theoretical content.

In summary, the Task Force is to be commended for their efforts to improve communication with regulatory authorities and to harmonize recommendations for measuring these outcomes across a variety of organizations. Unfortunately, little evidence of this important work is apparent in the current report, which holds way too much to safe territory and provides little guidance on many important issues facing the field. While the issues raised in this editorial are beyond the scope of any one paper to answer, let’s hope the forthcoming papers from the Task Force tackle at least some of them.

These statements reflect the personal opinions of the author and should in no way be construed as representing the opinion of Pfizer Inc.

References

- 1 Acquadro C, Berzon R, Dubois D, et al. Incorporating the patient’s perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health* 2003; 6:522–31.
- 2 American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC, 1999.
- 3 Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995;273:59–65.
- 4 Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med* 1996;334:835–40.