



sanofi aventis

Because health matters

04 April 2006

Via fax and UPS

Division of Dockets Management (HFA-305)
Food and Drug Administration
5630 Fishers Lane, Room 1061
Rockville, MD 20852

Re: Docket No. 2006D-0044

Draft Guidance for Industry on Patient Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims

Dear Sir/Madam:

Sanofi-aventis U.S. Inc, a member of the sanofi-aventis group, appreciates the opportunity to comment on the above-referenced Draft Guidance entitled "*Patient Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*".

This draft guidance describes how FDA evaluates Patient-Reported Outcome (PRO) instruments used as effectiveness endpoints in clinical trials. It also describes the Agency's current thinking on how sponsors can develop and use PRO instruments to support claims in approved product labeling.

We have evaluated the content of the draft guidance and offer the following comments and/or clarifications for your consideration.

Section I: Introduction

Lines 21 – 29/Definition of treatment benefit

Draft Guidance: "*This guidance describes how the FDA evaluates patient-reported outcome (PRO) instruments used as effectiveness endpoints in clinical trials. It also describes our current thinking on how sponsors can develop and use study results measured by PRO instruments to support claims in approved product labeling. It does not address the use of PRO instruments for purposes beyond evaluation of claims made about a drug or medical product in its labeling. By explicitly addressing the review issues identified in this guidance, sponsors can increase the efficiency of their endpoint discussions with the FDA during the product development process, streamline the FDA's review of PRO endpoint adequacy, and provide optimal information about the patient's perspective of treatment benefit at the time of product approval.*"

2006D-0044

C14

In some clinical trials, such as those for Oncology, where a new agent is added to a current strategy, could we consider that no deterioration, for example in HRQOL, could be considered a treatment benefit?

Lines 31-43/ PRO assessment by caregiver/parent/bed-partner

Draft Guidance: *“A PRO is a measurement of any aspect of a patient's health status that comes directly from the patient (i.e., without the interpretation of the patient's responses by a physician or anyone else). In clinical trials, a PRO instrument can be used to measure the impact of an intervention on one or more aspects of patients' health status, hereafter referred to as PRO concepts, ranging from the purely symptomatic (response of a headache) to more complex concepts (e.g., ability to carry out activities of daily living), to extremely complex concepts such as quality of life, which is widely understood to be a multi-domain concept with physical, psychological, and social components. Data generated by a PRO instrument can provide evidence of a treatment benefit from the patient perspective. For this data to be meaningful, however, there should be evidence that the PRO instrument effectively measures the particular concept that is studied. Generally, findings measured by PRO instruments may be used to support claims in approved product labeling if the claims are derived from adequate and well-controlled investigations that use PRO instruments that reliably and validly measure the specific concepts at issue.”*

In some diseases, such as ALZHEIMER, SCHIZOPHRENIA and SLEEP APNEA, is it acceptable that PRO be assessed by a third-party (caregiver, parent, bed-partner for example)?

Lines 38-40 and 180 – 181 (Section IV)

Draft Guidance: *“Data generated by a PRO instrument can provide evidence of a treatment benefit from the patient perspective. For this data to be meaningful, however, there should be evidence that the PRO instrument effectively measures the particular concept that is studied AND Therefore, in such instances, we encourage sponsors to document the original development processes, all modifications made, and updated assessments of its measurement properties.”*

Documentation is repeatedly referred to in this guidance. It is clear that this implies a substantial obligation to the sponsor to provide such information prior to the start of discriminatory trials. It would otherwise be impossible to ensure that a claim is likely to be acceptable to the Agency.

Section II: Background

Lines 67 - 72

Draft Guidance: *“Although an assessment of symptom improvement or pertinent function depends on patient perception, historically these assessments were often made by physicians who observed and interacted with patients (depression scales, heart failure severity scales, activities of daily living scales). Increasingly, such assessments are based on PRO instruments. The purpose of this guidance is to explain how the FDA evaluates such instruments for their usefulness in measuring and characterizing the benefit of medical product treatment.”*

Clarify intention “vis à vis” assessment of symptoms made so far by physicians will equal validity (and conclusions) likely to be questioned in the future? Where would instruments, intended to capture the caregiver perspective, fit?

Lines 86 - 88

Draft Guidance: *“a sponsor should develop evidence to show not only a change in symptoms, but how that change translates into other specific endpoints such as ability to perform activities of daily living, or improved psychological state.”*

This implies that the sponsor must not only demonstrate that the instrument is qualified to measure a concept, but if a wider claim is requested that the instrument has convergent validity with a measure of wider impact.

Section III: Patient-Reported Outcomes – Regulatory Perspective

Lines 114 - 118

Draft Guidance: *“This is important because improvements in clinical measures of a condition may not necessarily correspond to improvements in how the patient functions or feels. For example, clinically meaningful improvements in lung function as measured by spirometry may not correlate well with improvements in asthma-related symptoms and their impact on a patient’s ability to perform daily activities.”*

Does this mean that we may face some potential issues with clinical validity during the psychometric validation of the PRO used?

Lines 234 - 237

Draft Guidance: *"The FDA intends to review all evidence based on multidomain PRO measurements with particular attention to the precise claim that is supported by the results in the measured concepts or domains."*

Also noted below that *a priori* claims are required. It is not stated in the guidance but has been noted in discussion with the Agency that negative findings in domains, which are not the basis for claims, would be considered in labeling.

Lines 249 - 251

Draft Guidance: *"For measures of general concepts, the FDA intends to review how individual items are associated with each other, how items are associated with each domain, and how domains are associated with each other and the general concept of interest."*

As noted above it is unclear whether this will indicate that other domains (other than the specified domain) will be considered to be important for labeling or whether overlapping domains must be more clearly delineated?

Lines 269 - 271

Draft Guidance: *"The PRO instrument can be developed for a variety of roles, including defining trial entry criteria, including excessive severity, evaluating treatment benefit, or monitoring adverse events."*

Is the same level of documentation required, no matter what the role?

Lines 275 - 278

Draft Guidance: *"The FDA plans to compare the patient population used in the PRO instrument development process to the study populations enrolled in clinical trials to determine whether the instrument is appropriate to that population with respect to patient age, sex, ethnic identity, and cognitive ability."*

Ethnicity for multinational trials could become an issue. If an instrument is developed in a cross-section of one nation's patient population will this be acceptable to the Agency?

Lines 298 - 300

Draft Guidance: "FDA plans to review instrument development (e.g., results from patient interviews or focus groups) to determine whether adequate numbers of patients have supported the opinion that the specific items in the instrument are adequate and appropriate to measure the concept."

In this context, "adequate" should be further defined.

Lines 304 - 308

Draft Guidance: "For example, in assessing the concept performance of daily activities, it is more appropriate to ask whether or not the respondent performs specific activities (and if so, with how much difficulty) than whether or not he or she can perform daily activities (because patients may report they are able to perform a task even when they never do so). Of course, it would be critical to know that each item refers to something that patients actually do."

Motivation to undertake tasks seems to be denied by this statement but may be important to the patient (e.g. for insomniacs the feeling that they are unable to participate in activities is more important than the fact that they may be able to do the activity if undertaken).

Lines 322 - 324

Draft Guidance: "The FDA intends to review the comparability of data obtained when using multiple modes of administration to determine whether pooling of results from the multiple modes is appropriate."

See also electronic PRO comments below. Multiple methods should be avoided due to the opportunity to be contradictory.

Line 331-334: Choice of recall period

Draft Guidance: "use of the instrument, the characteristics of the disease/condition, and the treatment to be tested. When evaluating PRO-based claims, the FDA intends to review the study protocol to determine what steps were taken to ensure that patients understand the appropriate recall period."

How does the Agency define an appropriate recall period?

Lines 334 - 337

Draft Guidance: *"If a patient diary or some other form of unsupervised data entry is used, the FDA plans to review the protocol to determine what measures are taken to ensure that patients make entries according to the study design and not, for example, just before a clinic visit when their reports will be collected."*

It is not clear what "measures" refers to in this context. The implication is that in-clinic administration is preferred but is not consistent with the remark that patients leaving trials should undertake an assessment in practice.

Type of PRO instrument

Will the Agency evaluate a PRO related to patient productivity? For example if a sponsor would like to support a claim on improving productivity by assessing the number of days missed from work?

Lines 341 - 343

Draft Guidance: *"It is usually better to construct items that ask patients to describe their current state than to ask them to compare their current state with an earlier period or to attempt to average their experiences over a period of time."*

This statement is contradictory to previous advice where responses to questions relating to a time period (e.g. over the past month) have been acceptable.

Draft Guidance: *"Response choices are generally considered appropriate when:*

- Wording used in responses is clear and appropriate (e.g., anchoring a scale using the term normal assumes that patients understand what is normal).*
- Responses are appropriate for the intended population. For example, patients with visual impairment may find the VAS difficult to complete.*
- Responses offer a clear distinction between choices (e.g., patients may not distinguish between intense and severe if both are offered as response choices to describe their pain).*
- Instructions to patients for completing the questionnaire and selecting response options are adequate.*
- The number of response options is justified."*

How can we justify the number of response options?

Line 390 - Development of format

Draft Guidance: *"Examples of changes that can alter the way that patients respond to the same set of questions include: • changing an instrument from paper to electronic format"*

With more clinical trials using electronic case report forms (e-CRF), does it mean that all PRO instruments initially developed in paper format should be re-validated if the sponsor plans to use an electronic format?

Lines 404 - 407 and Lines 740 - 751 (standardized quality control)

Draft Guidance: *"The FDA recommends that the PRO instrument development process includes the generation of a user manual that specifies how to incorporate the instrument into a clinical trial in a way that minimizes administrator burden, patient burden, missing data, and poor data quality." AND "quality can be optimized at the design stage by specifying procedures to minimize inconsistencies in trial conduct. Examples of standardized instructions and processes that may appear in the protocol include: • Standardized training and instructions to patients for self-administered PRO instruments • Standardized interviewer training and interview format for PRO instruments administered in an interview format • Standardized instructions for the clinical investigators regarding patient supervision, timing and order of questionnaire administration during or outside the office visit, processes and rules for questionnaire review for completeness, and documentation of how and when data are filed, stored, and transmitted to or from the study site"*

There are additional requirements for training and standards referred to here. Will the Agency require that training be documented (as with Investigator Meetings) for PRO instruments?

Lines 413 - 414

Draft Guidance: *"The FDA intends to consider whether a PRO measure conforms to assumptions that the response choices represent appropriate intervals by reviewing distributions of item responses."*

Response categories are reviewed in-depth in this chapter. It is unclear what "appropriate" may mean (apart from avoiding confusing terms or biased responses).

Lines 424 - 426

Draft Guidance: *"When empirically determined patient preference ratings are used to weight items or domains, the FDA also intends to review the composition of samples and the process used to determine the preference weights."*

If the preference weights are for allocation purposes it is not clear why this would be inappropriate?

Lines 429 - 430

Draft Guidance: *"However, this practice is discouraged unless the relationship of the preference weights to the intended study population is known and found adequate and appropriate."*

Is this feasible?

Lines 462 - 464

Draft Guidance: *"The FDA intends to examine the final version of an instrument in light of its development history, including documentation of the complete list of items generated and the reasons for deleting or modifying items, as illustrated in Table 3."*

Would it mean that the amount of documentation needed for a modified instrument should exceed that of a newly developed (original) tool?

Lines 474 - 480

Draft Guidance: *"The FDA generally intends to review a PRO instrument for: reliability, validity, ability to detect change, and interpretability (e.g., minimum important difference). The FDA plans to review the measurement properties that are specific to the documented conceptual framework, confirmed scoring algorithm, administration procedures, and questionnaire format in light of the study population, study design, and statistical analysis plan. The sociodemographic and medical characteristics of any sample used to develop or validate a PRO instrument determine its appropriateness for future clinical study settings."*

Ethnicity for multinational trials could become an issue. If an instrument is developed in a cross-section of one nation's patient population will this be acceptable to the Agency?

Lines 491- 493 Test-retest reliability

Draft Guidance: *"Test-retest reliability is the most important type of reliability for PRO instruments used in clinical trials. Test-retest is most informative when the time interval chosen between the test and retest is appropriate for identifying stability in reference to the clinical trial protocol."*

It is difficult to assess test-retest reliability in clinical trials where only one visit is planned before treatment allocation. Could it be acceptable to assess the test-retest reliability by focusing on stable patients?

Draft Guidance: *"The FDA recognizes that the validation of an instrument is an ongoing process and that validity relates to both the instrument itself and how it is used."*

Please clarify/define the term "ongoing process".

Lines 502 - 504

Draft Guidance: *"Sponsors should consider a PRO endpoint for evidence of content-related validity, the instrument's ability to measure the stated concepts, and the instrument's ability to predict future outcomes."*

Predictability of outcomes implies either that sensitivity to treatment effect or a link to clinician reported outcomes is required. This should be further clarified.

Lines 510 - 512

Draft Guidance: *"If developers expected the instrument to discriminate between patient groups (e.g., between patients with different levels of severity), the FDA is interested in evidence that shows the instrument meaningfully discriminates."*

Meaningful discrimination is not illustrated in the guidance and requires further clarification

Lines 539 - 548 and 800 - 811 (Discussion with Division re interpretation)

Draft Guidance: *"Many PRO instruments are able to detect mean changes that are very small; accordingly it is important to consider whether such changes are meaningful. Therefore, it is appropriate for a critical distinction to be made between the mean effect seen (and what effect might be considered important) and a change in an individual that would be considered important, perhaps leading to a definition of a responder. For many widely used measures (pain, treadmill distance, HamD), the ability to show any difference between treatment groups has been considered evidence of a relevant treatment effect. If PRO instruments are to be considered more sensitive than past measures, it can be useful to specify a minimum important difference (MID) as a benchmark for interpreting mean differences. An MID is usually specific to the population under study." And subnote 4 (The FDA is specifically asking for comment on the need for, and appropriate standards for, MID definitions applied to PRO instruments used in clinical studies.)"*

It is not clear whether the FDA prefers responder analysis or MID. In neither case is it clear from the guidance how group-to-group analysis may be employed particularly for MID where this is normally an inpatient comparison.

Lines 694- 697

Draft Guidance: *"Over the course of some clinical trials, it can be anticipated that patients may become too ill to complete a questionnaire or to respond to an interviewer. In such cases, proxy reporting may help to prevent missing data. When this situation is anticipated, the FDA encourages the inclusion of proxy reports in parallel with patient self-report from the beginning of the study"*

Proxy respondents need to be clarified. Should these be patient caregivers, nurses or doctors? How will contradictions be handled when dual reports are gathered?

Lines 715 – 738: Study design – Blinding

There are many cases where blinding is not possible. The statement that PRO data from open label clinical trials is "rarely credible" should be revised.

Lines 709 - 713*

Draft Guidance: *"If the goal of PRO measurement is to support claims, we recommend that measurement of the PRO concept be clearly stated as a specific study objective. It is important that the protocol include the exact format and version of the specific PRO instrument to be administered. In the process of considering the NDA/BLA/PMA or NDA/BLA/PMA supplement, the FDA intends to compare both the planned and actual use of the PRO instrument and its analysis"*

This implies that the final form of the PRO instrument must be available before discriminatory studies are started (to be included in the protocol). Will further sophistication of the questionnaire (separate to the group-to-group analysis) thus invalidate the questionnaire?

Lines 763 – 768

Draft Guidance: *"The protocol can increase the likelihood that a trial will still be informative by establishing plans for gathering all treatment-related reasons for patients withdrawing from a trial and by trying to minimize patient dropouts prior to trial completion. We recommend the study protocol describe how missing data will be handled in the analysis. It could also establish a process by which PRO measurement is ascertained before or shortly after patient withdrawal from treatment exposure due to lack of efficacy or toxicity."*

Additional burden on studies may be envisioned if a follow-up visit is required to complete the PRO questionnaire. In some diseases such as ALZHEIMER, SCHIZOPHRENIA, SLEEP APNEA, is it acceptable that PRO will be assessed by a third-party (caregiver, parent, bed-partner for example)?

Docket No. 2006D-0044
04-April-2006

Lines 785 – 787 Duration of study

Draft Guidance: *“In a trial for a progressive disease where the PRO concept of interest does not change until after the follow-up required for other clinical efficacy parameters, longer study duration can be indicated.”*

In oncology trials, would it be possible to assess PRO only during chemotherapy and not during the follow-up period?

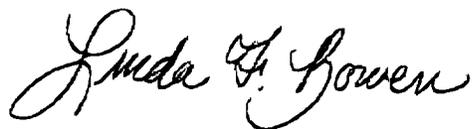
Lines 793 – 796: Design Considerations for Multiple Endpoints

Draft Guidance: *“A PRO instrument could be the primary endpoint measure of the study, a co-primary endpoint measure in conjunction with other objective or physician-rated measurements, or a secondary endpoint measure whose analysis would be considered according to a hierarchical sequence.”*

Does this mean that a label claim can be gained even if the endpoint is not a main secondary endpoint?

Sanofi-aventis appreciates the opportunity to comment on the Draft Guidance for Industry *Patient Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims* and are much obliged for your consideration.

Sincerely,



Linda F. Bowen
Director Regulatory Intelligence – Region US