Patient-reported Outcome (PRO) Assessments in Clinical Trials: Navigating the EMA and FDA Regulatory Framework

Ingela Wiklund, PhD, Senior Research Leader, United BioSource Corporation
Kathy Wyrwich, PhD, Senior Research Leader, United BioSource Corporation
Olivier Chassany, MD, PhD, Professor, Department de la Recherche Clinique et du Development, Paris, France

Agenda

- Comparison of the EMA and FDA Guidance on PROs
  - Ingela Wiklund
- Statistical Considerations and PRO Interpretation Comparisons
  - Kathy Wyrwich
- A Comparison of Examples of EMA and FDA Disease Area Guidance Documents – Implications
  - Olivier Chassany

Evolution of the Term: “Patient-reported Outcomes”

- Mid-1990s: “Health-related quality of life” began to appear in regulatory submissions for labeling and promotion
- Evidence to support a HRQL claim was discussed and harmonized internationally with participation from FDA, ERIQA, ISPOR and ISOQOL
- Growing interest in individualized therapy and personalized medicine
- Growing need to assess specific concepts in measurement, particularly symptoms
- Treatment benefit = improvement in how patients feel and function
- Medical product claims = the concept measured and empirically substantiated
- Evidence to support measure validity and reliability generalizable across all PROs


Patient-reported Outcomes (PROs) in the Drug Evaluation Process

- In 2003, EMA begins drafting a guideline on HRQL evaluation in registration trials.
- In 2005, the final document is published
- AIM: define the place and give recommendations for use of HRQL measures in the evaluation of medicinal products
- In 2006, FDA publish the draft PRO Guidance to support label claims
- In 2009, FDA publish the Final PRO Guidance document
- AIM: define the place and give recommendations for use of PRO endpoints in drug evaluation process to support label claims

Quality of Life and Patient-reported Outcomes (PROs): What is the difference between the FDA and EMA?

- EMEA is more likely to grant multiple claims (HRQL, symptoms, treatment satisfaction...)
- EMEA tends to accept a patient completed questionnaire provided it is relevant to the patient population and has adequate evidence validity
- EMEA most likely to accept and instrument developed in agreement with the FDA Guidance
- FDA imposes the guidance strictly (even retrospectively)
- FDA not likely to grant label on HRQL
There are many similarities between FDA and EMA, but also some differences …

**Similarities**
- Endpoint model
- Content validity and relevance
- Previously validated in population
- Clinical endpoints vs. HRQL
- FDA use the Guidance for ObsRO and ClinRO
- Recall period
- Trial design: compare with standard of care, placebo not accepted
- Trial duration

**Differences**
- Terminology (HRQL vs. PROs)
- FDA recommends instruments for use in various therapeutic areas, whilst EMA recommends concepts of relevance to be driven by the desired claims, endpoint model, and anticipated therapeutic benefits of the molecule to be tested
- FDA in general have issues around recall and are concerned of memory bias if the recall period is longer than 7 days
- FDA recommends sponsors to use electronic data capture of diary symptoms in trials to secure a time stamp as to when the diary was completed
- FDA has detailed PRO Dossier template that sponsors are requested to follow, whereas EMA does not request any specific format. Likely, the FDA PRO Dossier is far too extensive for EMA

**Terminology and Endpoint Definition Differences**
- EMA provides guidance on HRQL, a specific type of PRO, distinguished by being multi-dimensional
- Symptoms such as pain, or migraine are well accepted primary and secondary endpoints (i.e., clinical endpoints) in registration trials and as such require no specific validation

**FDA – Patient Outcomes Classification Sources and Examples**

<table>
<thead>
<tr>
<th>Patient-reported Outcomes</th>
<th>Caregiver-reported Outcomes</th>
<th>Clinician-reported Outcomes</th>
<th>Biological &amp; Physiological Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global impression</td>
<td>Global impression</td>
<td>Global impression</td>
<td>BP</td>
</tr>
<tr>
<td>Global burden</td>
<td>Global burden</td>
<td>Global impression</td>
<td>FEV1</td>
</tr>
<tr>
<td>Global rating of change</td>
<td>Global rating of change</td>
<td>Global impression</td>
<td>HbA1c</td>
</tr>
<tr>
<td>Well-being</td>
<td>Well-being</td>
<td>Global impression</td>
<td>CPT4</td>
</tr>
<tr>
<td>Satisfaction with Rx</td>
<td>Satisfaction with Rx</td>
<td>Global impression</td>
<td>Tumor size</td>
</tr>
</tbody>
</table>

- Pain
- Shortness of breath
- Depression
- NYHA classification
- Functional status
- Global impression
- Sp
- Global rating of change
- Survival
- Observations
- FEV1
- HbA1c
- CPT4
- Tumor size

**FDA – PRO data in 30% or more of labels for NMEs**

**What about Non-PRO Measures?**

FDA uses PRO Guidance to evaluate all types of Clinical Outcome Assessments:
- Patient-reported Outcome (PRO)
- Observer-reported Outcome (ObsRO)
- Clinician-reported Outcome (ClinRO)
Instruments to be validated for the patient population/condition to be studied

- Protocol and analysis plan (how PROs will be measured and interpreted)

- Measurement properties (psychometric properties)
  - Reliability (including reproducibility)
  - Validity (including content validity and construct validity)
  - Responsiveness

- Translation/cultural adaptation (translation and back-translation)
  - Who translated and back-translated?
  - How many translators were involved?
  - Was cross-cultural validation performed?

- Focus group composition and number of patients included
  - Who participated in focus groups?
  - How many patients were included?

- Endpoint model (provide justification of endpoints and how they will support claims, missing data, multiply)

- User manual and training material (standardized administration and quality control)

- Statistical analysis plan (how PROs will be analyzed and quality controlled, including burden of illness and clinical trials endpoints)

- PRO Evidence Dossier to Incorporate the Following:
  - Copy of PRO instrument(s), supporting references
  - All targeted claims (incl. specific claims and endpoints) and how they will support claims
  - Endpoints model (provide justification of selected instruments, relationship among all endpoints)
  - PRO conceptual framework (incl. diagram)
  - Item generation - Content validity documentation (PRO development history, saturation grid, cognitive debriefing, item banking, response options, recall period)
  - Measurement properties (psychometric performance)
  - Interpretation of score (incl. responder definition)
  - User manual and training plan (how PROs will be standardized administration and quality controlled)
  - User manual and training plan (how PROs will be analyzed and quality controlled, including burden of illness and clinical trials endpoints)

- PRO submission documents
  - Generic « Generic » instruments such as SF-36
  - HRQL « Generic » instruments such as SF-36
  - HR-QOL index
  - HR-QOL summaries
  - HR-QOL profile

- Review of EWP NfG – Examples of PRO Questionnaires

- AIMS, HAQ-S (Spondylitis)
- HAQ-Sk (Psoriasis)
- Inflammatory Bowel Disease Questionnaire
- Liebowitz Social Anxiety Scale
- Minnesota Living with Heart Failure
- Patient self-rated symptoms scales (VAS, NRS)
- Psoriasis Disability Index
- St Georges Respiratory Questionnaire
- Sheenan Disability Scale
- WOMAC
- HRQL « Generic » instruments such as SF-36

- Example – Rheumatoid Arthritis

- Swollen joints
  - Frequency of Pain

- Anemia

- Specific Symptom
  - Specific Impact

- Health-Related QOL (HRQOL)

- Generic Impact
  - SF-36
  - Health Survey, Nottingham Health Profile

- Similarities between FDA and EMA

- Endpoint Model
  - Hierarchy of endpoints with specification of primary, co-primary and secondary endpoints
  - Content validity
    - Content and domains of relevance to the disease population and its treatments

- Previously validated in population/condition

- Validation completed prior to use in phase 3

- Differences between FDA and EMA

- Recommendations of HRQL instruments to use
  - FDA provides a “shopping list” of validated instruments for use in various therapeutic areas
  - FDA recommends identification of concepts of relevance to be driven by the desired claims, endpoint model, and anticipated therapeutic benefits of the molecule to be tested

- Recall period
  - FDA advice sponsors to use a daily diary for symptom assessments, while EMA has no such requirements
  - FDA in general has issues around recall and are concerned of memory bias if the recall period is longer than 7 days

- ePRO use
  - FDA advice sponsors to use electronic data capture of diary symptoms in trials to secure a time stamp as to when the diary was completed

- PRO submission documents
  - FDA has a detailed PRO Evidence Dossier template that sponsors are requested to follow, whereas EMA does not require a specific format. Likely, the FDA PRO Dossier is far too extensive for EMA?
Content Validity Evidence – Beyond PROs:

What is being measured and does the outcome assessment measure that thing?
- Patient reported outcome assessments
  - Identify the concepts (symptoms + other) and context of use
  - Qualitative research to minimize variability between patients
  - Cognitive debriefing to confirm content
- Clinician reported outcome assessments
  - Identify the concept (observable + interpretation) and context of use
  - Qualitative research to minimize variability between clinicians
  - Cognitive debriefing to confirm content
- Observer reported outcome assessments
  - Identify the concept (observable) and context of use
  - Qualitative research to minimize variability between observers
  - Cognitive debriefing to confirm content


EMA and FDA: How to support a HRQL/PRO claim?

- The claim will always be considered depending on the strength of the evidence and the relevance (pertinence and importance) of the findings.
- The strength of the evidence should be based on:
  - the rationale for HRQL/PRO assessment in the context of the disease/medicinal product;
  - the justification of the choice of the HRQL/PRO questionnaire(s);
  - the objectives of HRQL/PRO assessment and the hypotheses of HRQL/PRO changes;
  - the evidence of validation (and of cultural adaptation/translation if applicable) of the HRQL/PRO questionnaire(s);
  - the adequacy of the statistical analysis plan; and
  - the relevance of observed changes.