DRUG DEVELOPMENT TOOL
QUALIFICATION PLAN DETERMINATION
DDT COA #000015

Stephen Joel Coons, PhD
Executive Director, PRO Consortium
Critical Path Institute
1730 E. River Road
Tucson, AZ 85718

Dear Dr Coons:

We have completed our review of the Qualification Plan (QP) for Drug Development Tool (DDT) COA #000015 received on March 19, 2019 by the CDER Clinical Outcome Assessments (COA) Qualification Program, submitted under section 507 of the Federal Food, Drug, and Cosmetic Act.

The QP is for the Patient-Reported Outcomes Measurement Information System® (PROMIS®) Fatigue Short Form 10a, a patient-reported outcome (PRO), proposed for the assessment of fatigue severity in adult patients, 18 years of age and older, with rheumatoid arthritis (RA).

FDA has completed its review and has agreed to accept your QP. The qualification review team's (QRT's) response to the questions included in the QP submission can be found below:

Question 1: Does CDER agree that this is an acceptable plan for the qualification of the PROMIS Fatigue Short Form 10a?

QRT Response:
In general, your qualification plan for the PROMIS Fatigue Short Form 10a appears reasonable. However, we have concerns with some aspects of the proposed quantitative analyses. See QRT Response to Question 2.

Question 2: What additional quantitative analyses and/or evidence are necessary to support the qualification of the PROMIS Fatigue Short Form 10a for use as a secondary endpoint measure in rheumatoid arthritis registration trials?

QRT Response:
In preparing to submit a Full Qualification Package (FQP), please ensure that the FQP submission addresses the quantitative analysis issues and recommendations outlined below:

1. The FQP should specify: model assumptions, evaluation of differential item functioning, and calculation of goodness-of-fit statistics.

2. Indicate how you will calculate the frequency and percentage of subjects missing each individual item, i.e., whether the denominator in this calculation will include all randomized subjects or be restricted to subjects who (a) completed the form, (b) remain in the study, or (c) meet some other criteria at this time point. We recommend using all randomized subjects for this analysis.

3. Provide rationale for the use of omega to assess dimensionality or use a more traditional approach to assess dimensionality using exploratory or confirmatory factor analysis. We recommend that the proposed method consider the ordinal nature of the data.

4. Clarify the use of the interaction term in the planned ANOVA model for test-retest reliability. Given that there is only one observation per subject at each time point, it is not clear what this interaction refers to and its purpose in the model.

5. We recommend the additional calculation of Spearman's correlation for convergent and divergent validity analyses, which, unlike Pearson's correlation, does not require linearity.

6. We have concerns with the use of CDAI for the known-groups validity analysis, given it may not be a relevant anchor appropriate for the concept of fatigue. We expect that known-groups validity will need to be evaluated again with data from another longitudinal study where more appropriate anchors were included (see comments in point 9 below). All within-patient meaningful change threshold analyses provided in this submission will be considered preliminary.

   In addition to these analyses, we ask that you include boxplots of PROMIS scores for each of the known groups.

7. We request additional elaboration on the missing data simulation plan to assist implementation and validation of the results by the statistical reviewer. For example, it is not specified how the missingness (dropping of items) will be applied across subjects (i.e., the level of missingness across individuals). Your simulation plan should be detailed enough for implementation by a reviewer.

8. Additional details are still needed on the thresholds and endpoints used to group subjects into improved/stable/worsened for the ability to detect change analyses. Given the targeted concepts, we see more relevance in the use of the SF-36 Vitality or MAF as anchors for this approach. Therefore, we recommend the requestor determine whether an established within-patient meaningful change scores for the SF-36 Vitality or MAF to define responders (not the MID) exists. However, meaningful within-patient change scores will likely need to be evaluated again with data from another longitudinal study where more appropriate anchors were included (see comments in point 9 below).

9. We have concerns regarding the adequacy of the ACR response and CDAI anchors proposed to evaluate PROMIS Fatigue Short Form's ability to detect change and to determine clinically meaningful within-patient change thresholds for the PROMIS Fatigue Short Form. It is not clear that these anchors are appropriate for these purposes as they do not represent change in the concept of interest. Global anchor scales, such as patient global impression of severity/change (PGIS/C) items, both specific to fatigue, would be more useful and readily interpretable. We expect that the ability to detect change analyses and determination of clinically meaningful within-patient meaningful change thresholds for the PROMIS Fatigue Short Form will need to be evaluated again using data from another longitudinal study where more appropriate anchors were included. The determination regarding a qualification statement for the PROMIS Fatigue Short Form will be based on our review of the data presented in the full qualification package (FQP) (e.g., whether or not we have data from another longitudinal study).

   We further note that we will consider the distribution-based methods as supportive to the anchor-based methods.

Appendix 1 of this letter contains the contents to include in your submission to reach the next milestone (full qualification package). Please contact the CDER COA Qualification Program at COADDTQualification@fda.hhs.gov should you have any questions (refer to DDT COA #000015).

Sincerely,

Elektra J. Papadopoulos -S

Digitally signed by Elektra J. Papadopoulos -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300170743, cn=Elektra J. Papadopoulos -S
Date: 2019.09.09 11:33:33 -04'00'

Elektra Papadopoulos, MD, MPH
Associate Director
Clinical Outcome Assessments Staff
Office of New Drugs
Center for Drug Evaluation and Research

Sally M. Seymour -S

Digitally signed by Sally M. Seymour -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300222097, cn=Sally M. Seymour -S
Date: 2019.09.13 07:52:45 -04'00'

Sally Seymour, MD
Director
Division of Pulmonary, Allergy, and Rheumatology Products (DPARP)
Office of New Drugs
Center for Drug Evaluation and Research

Appendix 1: COA FULL QUALIFICATION PACKAGE

The COA Full Qualification Package (FQP) should be accompanied by a cover letter, the following completed sections, a copy of the instrument, the scoring algorithm, and the user manual. This package should contain the results of both the completed qualitative research and the quantitative research (measurement properties). Some sections may be less relevant for certain COAs (i.e., performance outcome measures) than others. If literature is cited, please cite using the number assigned to the source in a numbered reference list.

Note: Sections 1 and 2 will be posted publicly under Section 507 as well as any appendices or attachments referred to in those sections. In sections 1-2, please also refrain from referring to other sections within the FQP. Section 507 refers to section 507 of the Federal Food, Drug, and Cosmetic Act [FD&C Act] which was created by Section 3011 of the 21st Century Cures Act.

Section 1: Plan for COA Qualification

1.1 Introduction and overview

- This should include a concise description of the disease and the clinical trial setting in which the COA would be used, the limitations of existing assessments, a brief description of the existing or planned COA, and the rationale for use in drug development.

1.2 Concept of Interest for meaningful treatment benefit

- Describe the meaningful aspect of patient experience that will represent the intended benefit of treatment (e.g., the specific symptom and/or sign presence or severity or limitations in performance or daily activities relevant in the targeted context of use)

1.3 Context of Use

- Identify the targeted study population, including a definition of the disease and selection criteria for clinical trials (e.g., baseline symptom severity, patient demographics, language/culture groups)
- Identify the targeted study design. Most commonly the COA will be used to assess the change (compared to a control) induced by a medical treatment.
- Identify the targeted study objectives and endpoint positioning (i.e., planned set of primary and secondary endpoints with hierarchy). Usually, the COA will serve as a primary or secondary study endpoint measure.

1.4 Critical details of the measure to the degree known

- Reporter, if applicable
- Item content or description of the measure
- Mode of administration (i.e., self-administered, interview-administered)
- Data collection method

1.5 Description of the involvement of external expertise, including scientific communities or other international regulatory agencies, if applicable (i.e., working group, consortia)

Section 2: Executive Summary

- High-level summary of what is included in the Full Qualification Package and results to be described in the sections below

Section 3: Qualitative Evidence and Conceptual Framework

- Evidence of content validity (i.e., documentation that the COA measures the concept of interest in the context of use)

3.1 Literature review
3.2 Expert input
3.3 Reporter input (e.g., for PRO measures, concept elicitation, focus groups, or in-depth qualitative interviews to generate items, select response options, recall period, and finalize item content; for PerfO measures, evidence to support that the tasks being performed are representative of the meaningful health aspect of the concept of interest and are relevant to ability to function in day-to-day life)
3.4 Concept elicitation
3.5 Item generation
3.6 Cognitive interviews
3.7 Conceptual Framework

Sections 4, 5, and 6: Proposed Quantitative Analysis Results

Section 4: Cross-sectional evaluation of measurement properties (Results)

4.1 Item Level Description
  4.1.1 Item descriptive statistics including frequency distribution of both item response and overall scores, floor and ceiling effect, and percentage of missing response
  4.1.2 Inter-item relationships and dimensionality analysis (e.g., factor analysis or principal component analysis and evaluation of conceptual framework)
  4.1.3 Item inclusion and reduction decision, identification of subscales (if any), and modification to conceptual framework
4.2 Scoring algorithm (e.g., include information about evaluation of measurement model assumptions, applicable goodness-of-fit statistics). The scoring algorithm should also include how missing data will be handled.
4.3 Reliability
  4.3.1 Test-retest (e.g., intraclass correlation coefficient)
  4.3.2 Internal consistency (e.g., Cronbach's alpha)
  4.3.3 Inter-rater (e.g., kappa coefficient)

4.4 Construct validity
  4.4.1 Convergent and discriminant validity (e.g., association with other instruments assessing similar concepts)
  4.4.2 Known groups validity (e.g., difference in scores between subgroups of subjects with known status)

4.5 Score reliability in the presence of missing item-level and if applicable scale-level data
4.6 Copy of instrument
4.7 User manual and plans for further revision and refinement
    4.7.1 Administration procedures
    4.7.2 Training administration
    4.7.3 Scoring and interpretation procedures

Section 5: Longitudinal evaluation of measurement properties (Results If Known)

5.1 Ability to detect change

Section 6: Interpretation of Score (Results If Known)

6.1 Evaluation and definition of meaningful within person change (improvement and worsening)

Section 7: Language translation and cultural adaptation (If Applicable)

7.1 Process for simultaneous development of versions in multiple languages or cultures
7.2 Process of translation/adaptation of original version
7.3 Evidence that content validity is similar for versions in multiple languages

Section 8: References

- References and copies of the most important references that the submitter feels CDER reviewers may want to review

Section 9: Appendices and Attachments

- Study documents (e.g., protocols, analysis plan, interview guide, data collection form(s))