DDT 000018 **COMMENTS ON COA DDT SUBMISSION**

August 17, 2017

Todd Saretsky, MPH
Outcomes Researcher
ICON Commercialisation and Outcomes
456 Montgomery St., Suite 2200
San Francisco, CA 94104
Phone: 415 856 0840
Todd.Saretsky@iconplc.com

Regarding: DDT #000018 Updated psychometric evaluation and SAP for the Community-Acquired Bacterial Pneumonia Daily Symptom Diary (CABP PRO) instrument for the measurement of respiratory and systemic symptoms of community-acquired bacterial pneumonia (CABP).

Dear Mr. Saretsky:

Please refer to your October 25, 2016 submission for COA DDT Qualification: CABP PRO, first submitted to the FDA on April 5, 2013 (DDT #000018). This letter also reflects the teleconference held on July 24, 2017. We appreciate your attention to our requests for revision of your psychometric evaluation study protocol and SAP. We have reviewed the revised documents and while you have addressed many of our previous concerns, the QRT believes that further revisions will be needed as development continues. We acknowledge that you will be proceeding with your psychometric study prior to item reduction. However, it is possible that administering the draft instrument items prior to item reduction phase can impact the sensitivity of the instrument and its ability to accurately assess core concepts of CABP. In the absence of a formal item reduction phase prior to psychometric testing, we recommend that you engage in multiple iterations of item reduction using qualitative (i.e., expert consensus panel review with subject matter experts, including FDA representatives) and quantitative methods to ensure that the most relevant items are included in the final instrument. Note that another clinical study may be needed to confirm the psychometric properties of the reduced instrument. Additionally, you will need to provide greater detail on your ePRO implementation plan, instrument administration schedule, and proposed analyses in your next submission.

Specific recommendations and comments related to your revised psychometric evaluation protocol and SAP submission are provided below:

General Comments:

1. We recommend that you focus on items 1-7, which constitute core symptoms of CABP. A thorough review by the project team of the Biomarkers Consortium of the FNIH found support for a symptom improvement efficacy endpoint based on these cardinal symptoms that can be used in a noninferiority trial as a part of the primary efficacy outcome. FDA concurred with this approach and found historical evidence for a treatment effect and noninferiority margin that supports the selection of an early symptom improvement efficacy endpoint for the noninferiority trial. Any deviation from these cardinal symptoms has potential to alter the assay sensitivity and therefore create difficulty for the use of a noninferiority trial to establish efficacy of a new antibacterial drug for treatment of CABP. The remaining sections may be considered for use as part of supportive endpoints.
2. Your protocol still lacks details regarding your study administration. In your next submission you should include additional information about the following:
    a. Data collection procedures for inpatients and outpatients: Procedures will differ for these subpopulations, especially in instances where a patient's condition worsens to the point of hospitalization or hospitalization with ventilation over the course of the study, following the initial diagnosis of CABP. In these cases, a patient may be enrolled in the study in the outpatient setting, but complete the study in the inpatient setting. Details regarding how these administrations when the setting changes will differ need to be added to the protocol.
    b. Exit interviews: Exit interview procedures lack detail and it is unclear whether accommodations will be made for non-English speaking Hispanic/Latino patient populations. Details regarding any accommodations need to be outlined in the study procedures.

3. Information regarding your ePRO system and implementation plan is not included in your protocol. We recommend the following:
    a. Submit screenshots and training materials (site and patient) for your ePRO implementation for Agency review and comment.
    b. Plan to perform usability testing of ePRO devices and implement a back-up plan (e.g., paper, web-based) in case of any malfunctions with the electronic devices, prior to using the devices in your psychometric evaluation study. Please include details regarding this stage of development and submit protocols and materials related to this usability testing for Agency review and comment.

4. Currently, your protocol indicates that your mode of administration will either be ePRO or a telephone interview. We recommend that you move forward with only the ePRO mode of administration (with paper backup in case of device malfunction only) as this will be the least complicated and in alignment with development efforts to date. If telephone interviews are also adopted, you will need to provide details on how patients will be selected for the telephone interviews. Likewise, you will need to develop and submit an interviewer

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov

administered version of the CABP PRO instrument (including prompts) for review and comment.

5. Please provide further details regarding your quality assurance procedures, including: 1) requirements and methods for site and study staff qualifications and training; 2) data monitoring; and 3) data entry quality assurance (for paper backup entry into the electronic system).

6. Please provide details regarding plans for translation and cultural adaptation of the CABP PRO. This instrument will need to be culturally adapted and adequately translated for all intended study populations for use in multinational trials. We refer you to the ISPOR principles for the translation and cultural validation process.[1]

We offer the following additional comments and suggestions related to your submission:

CABP PRO Instrument

1. For items 24-29, we recommend removing the "Not Applicable" response option. We don't believe that "Not Applicable" is a meaningful response option for these items (e.g., Item 24 – "Did you have difficulty sleeping?") and it is unclear how these options would be scored. Additionally, we are concerned that Item 24 (difficulty sleeping), Item 25 (difficulty doing your usual activities), and Item 27 (social activities) will not be applicable to the inpatient population as level of independence (doing usual activities, social interaction) and sleep schedules would likely be influenced by hospital protocol.

Psychometric Evaluation Protocol

1. We recommend that you add further details and procedures (e.g., detailed data monitoring at regular intervals; program daily reminders and/or implement daily reminder phone calls or texts for outpatient participants) in order to minimize missing data.
2. P. 11 – Please specify whether respondents will be allowed to skip answers or whether each response will be forced choice. We would prefer if respondents are allowed to skip to avoid erroneous answers. We recommend that you add a skip option to each question and program a logic check that will ask respondents to indicate whether they intentionally skipped items. This way, there is a systematic way to account for missing data.
3. P. 16 – An error was found. Please correct "0.8" to "0.08."
4. P. 17-18 – There is discrepant information regarding your test-retest reliability analyses. You initially indicate that scores from days 7 and 10 will be used to assess test-retest reliability. However, on p. 18, you state that scores from days 7 and 14 will be used. Please

[1] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P; ISPOR Task Force for Translation and Cultural Adaptation. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov

Health. 2005 Mar-Apr;8(2):94-104.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov

clarify your analysis plan and correct the protocol and SAP accordingly. Please also consider using consecutive days' CABP PRO scores from participants whose supplemental question 1 (p. 55) response is "About the same" (this question asks: *Overall, how are your pneumonia symptoms today compared to yesterday*?).

Statistical Analysis Plan

1. Section 4.2 Handling of Missing Data
    a. Three studies are referred to in your description of test-retest reliability ("Test-retest reliability for all three psychometric evaluation studies"). Please specify what three studies you are referring to.
    b. Item and assessment level missingness needs to be assessed. Consider using multiple imputation to handle the missing responses, or consider conducting weighted data analyses with inverse probability of missingness weights. Single imputation with the mean of the observed item responses does not adequately account for variability due to missingness and should be avoided. Additionally, depending on the missingness MCAR may not be a valid assumption. If the MCAR assumption does not hold, then factor analyses and other psychometric data analyses may yield biased results.
    c. If items are not reduced, participant burden will be high given the frequency of administration and you might have increased levels of missing data due to respondent fatigue. In order to increase your power, we ask that you consider consider the use of multiple imputation beyond handling missingness at baseline.
    d. If, after symptom resolution prior to Day 14, participants do not complete daily diaries, then the post-resolution responses are missing, contrary to the SAP. Instead of handling these responses with LOCF imputation, we recommend that you use multiple imputation to handle post symptom resolution missingness. In general, LOCF has poor statistical properties, and it is unwarranted to assume that symptoms will remain resolved after the initial rating of symptom resolution.
    e. The *Guidance for Industry: Patient-Reported Outcome Measures* (p. 30) recommends at least two sensitivity analyses if multiple imputation is used to handle missing data.
    f. For all case report forms (CRFs), we recommend that you add a field at the top of each page to where patient personal identification numbers can be inputted by site staff or pre-populated.

2. Section 4.3 Distributional Considerations
    a. We recommend that you also consider generating Q-Q plots to assess normality.

3. Proposed analysis order of operations (Figure 1): We recommend that you modify the figure as follows:
    a. Have only one arrow coming out of "EFA" going to "Rasch"
    b. Remove other arrows currently coming out of "EFA" (e.g., going to "Ability to Detect Change") and have these instead come out of "Rasch." This is because the proposed Rasch analysis could result in item deletion and these deleted items would not be included in your analyses to assess the ability to detect change.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov

4. Section 5.2.3 Rasch Analysis
    a. Please indicate which of the two Rasch models for polytomous items (partial credit model or rating scale model) you intend to use.
    b. Many items on Day 1 may have no or few responses of "Not at all," which would compromise the accuracy of parameter estimates. Consider whether there is value added to use not only Day 1 but also other days' data in your analysis. Please specify whether separate Rasch/IRT analysis will be conducted for each subscale if results from your factor analysis (Section 5.2.2) reveal multidimensionality. Also, if your factor analysis reveals multidimensionality, indicate whether a multidimensional Rasch/IRT analysis will be performed

5. Section 5.2.4 Scoring Algorithm
    a. If your factor analysis reveals multidimensionality, you should consider whether it is appropriate to compute an overall CABP PRO score.
        a. Please clarify how you envision a CABP PRO total score will be used to define efficacy endpoints for CABP clinical trials.
    b. The SAP presupposes that classical test theory should be used for scoring; Rasch/IRT analysis is intended to play a subsidiary role in determining adequate items. Please clarify why you are not using Rasch/IRT analysis to generate scores.

6. Section 5.2.5.1 Internal Consistency
    a. You propose that Cronbach's alpha be estimated based on the Day 1 sample. We recommend that Cronbach's alpha be estimated at several different time points in order to capture possible changes in alpha over time.[2]
    b. It might be useful to recreate Reeve and Fayers' Figure 1.5.4 using your IRT analyses to examine the reliability of the CABP PRO at different levels of health status.[3]
    c. If multiple subscales are defined we recommend that you evaluate internal consistencies separately for each subscale, in addition to the total score.

7. Section 5.2.8 Exploring Responder Definitions
    a. Your proposed use of the Patient Global Impression of Severity measure (PGI-S; referred to as PGI in your submission) in determining response thresholds for the CABP PRO instrument is acceptable. However, we recommend that you use both the PGI-S and PGI-C to provide an accumulation of evidence to help interpret a clinically meaningful score change in the CABP PRO instrument using anchor-based methods and cumulative distribution function (CDF) analysis. We prefer this approach over using ROC curves.

[2] Biemer, PP, Christ, SL, Wiesen, CA. A General Approach for Estimating Scale Score Reliability for Panel Survey Data. *Psychological Methods* 2009 December; 14(4): 400-412.
[3] Reeve, BB, Fayers, P. Applying ItemResponse Theory Modelling for Evaluating Questionnaire Itemand Scale Properties. In P. Fayers & RD Hays (Eds.) Assessing Quality of Life in Clinical Trials: Methods of Practice (2nd ed.). New York; Oxford University Press: 2005.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov

i. We recommend that you consider using the Clinician Global Impression of Severity (CGI-S) and Change (CGI-C) measures as supportive evidence to help further bolster the patient observations. Both the patient-rated and clinician-rated anchor scales should be assessed at the same time points as, but completed after, the CABP PRO instrument.

ii. You should generate CDF plots depicting changes in the CABP PRO score(s) by corresponding patient and clinician global impression of change and severity item response options (i.e., separate curves for each response option).

If you have any questions or would like to set up a teleconference to answer questions, please contact the Clinical Outcome Assessments Staff at COADDTQualification@fda.hhs.gov.

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: 2017.08.22 12:24:50 -04'00'

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

Sumathi Nambiar -S

Digitally signed by Sumathi Nambiar -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300145731, cn=Sumathi Nambiar -S
Date: 2017.08.19 20:32:03 -04'00'

Sumathi Nambiar, MD, MPH
Director
Division of Anti-Infective Products (DAIP)
Office of Antimicrobial Products
Center for Drug Evaluation and Research

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20903
www.fda.gov