DDT COA #000041 **COMMENTS ON SUBMISSION**

Dan Turner MD, PhD
Director, Pediatric Gastroenterology Unit
Shaare Zedek Medical Center,
P.O.B 3235 Jerusalem, Israel 91031
email: turnerd@szmc.org.il
Tel: +972-2-6666482
Cell phone +972-50-8685841
Fax: +972-2-6555756

Anne M. Griffiths MD
Head, Division of Pediatric Gastroenterology Hepatology and Nutrition,
SickKids Hospital
Northbridge Financial Corporation Chair in IBD
Professor of Paediatrics
University of Toronto
555 University Avenue. Toronto, Canada M5G1X8
Email: anne.griffiths@sickkids.ca
Tel: +1-416-8137826
Fax: +1-416-8136531

Dear Drs. Turner and Griffiths:

We have completed our review of the submission for the Clinical Outcome Assessment (COA) Qualification Program. The submission contained a point-by-point response to the FDA Qualification Review Team (QRT) letter dated July 15th, 2018, cognitive interview script for caregivers of children 2-7 years of age and for children 8-18 years of age for phase 2C, the TUMMY-UC scale for caregivers of children 2-7 years of age, and the TUMMY-UC scale for children 8-18 years of age.

We appreciate your progress towards developing an observer-reported and patient- reported outcome (ObsRO and PRO) to address the need for well-defined and reliable COAs in children with UC.

We continue to strongly recommend that this tool, which is being developed for use in clinical trials, include our recommendations to modify the instructions such that patients and caregivers are instructed to respond at bedtime every day rather than allowing the patient to choose the time of day. We also continue to strongly recommend the revision of the pain item such that if numbers are included beneath the faces, that the numbers do not read 0, 2, 4, 6, 8 as it is in the submitted

instrument, but rather something more simple and understandable from a patient's perspective such as 0, 1, 2, 3, 4. We also recommend removing the line as the respondent may mark between numbers in contrast to the instructions, which direct the respondent to select either a number or a face. Additionally, please provide a detailed scoring algorithm in your user's guide in your next submission.

In finalizing your instruments for your qualitative study (stage 2C), please review our comments embedded in our replies to your point-by-point responses below. Please note your responses are in regular font, and our QRT replies are in **bolded** font. Please also find our suggested tracked edits to your interview scripts in the appended documents.

1. Submitter's response: Indeed, the PRO and the ObsRO versions will be pooled for analysis and thus the three subjective items in the ObsRO (i.e. abdominal pain, fatigue and urgency) have been reconstructed to mirror the PRO format. This has been done also following the EMA comments and based on the input from caregiver interviews thus far (tables of behaviors that support the new structure were sent to you recently as part of the documents submitted to the EMA).

   As requested we plan to add an additional phase of cognitive interviews (phase 2C) to ensure accuracy and completeness of the revised items, establish conceptual equivalence of the ObsRO, and increase diversity of responders.

   We will thus now add ~34 more interviews in phase 2C:
   a. 7 interviews of caregivers to children 2-7 years (ObsRO)
   b. 10 interviews of children 8-12 years (PRO)
   c. 10 interviews of caregivers to children 8-12 years (including the 3 subjective items: stomach pain, weakness and urgency), in parallel to their children's scoring of the PRO
   d. 7 interviews of children 13-18 years (PRO)

   These interviews we will include African American, from two new US sites to be initiated now.

   The interviews of phase 2C will be used also to explore the new user-guide recommendation of the stool-frequency item (i.e. very frequent stools over a short period of time will be considered as one stool), the abdominal pain item in response to the FDA comments below, and all the other slight revisions of the TUMMY-UC done to address the comments by the EMA and FDA.

   QRT reply: We agree with your efforts to increase diversity in the interviewed sample by including more African American patients. In addition, we have the following recommendations for your consideration:
   - **To adequately establish comparability between the two instruments based on both qualitative and quantitative testing, we believe it is necessary for caregivers of 8-12-year-old children to complete the entire ObsRO (rather than**

only a subset of items), while the respective 8-12-year-old children should complete the entire PRO in parallel.

- **Regarding the scope of disease severity, it is important to have an adequate representation of moderate to severe patients in the cognitive interview stage of the study. If the severity representation that was included in stage 2B (55% in remission, 28% moderate to severe and 17% mild disease severity) can be approximated in stage 2C, that should be acceptable.**

TUMMY-UC PRO Instrument

Instrument instructions:

2. Submitter's response:

i. The instruction to compete the TUMMY-UC at a consistent time daily has been added also to the instructions to patients and caregivers.
ii. The time of the day has been added to the PRO and ObsRO versions.
iii. The electronic diary will be developed at a later stage but will include a reminder function and time and date stamps.

QRT reply: We acknowledge your changes; however, we continue to strongly recommend that this tool, which is being developed for use in clinical trials, include our recommendations to modify the instructions such that patients and caregivers are instructed to respond at bedtime every day rather than allowing the patient/caregiver to choose the time of day.

3. Submitter's response: Wording has been improved as suggested.

QRT reply: We acknowledge your changes.

Abdominal pain item:

4. Submitter's response:

i. The numbers under the pain rating scale have been reverted to the original scoring
ii. The weighted scores were removed from the TUMMY-UC draft and added in the User's guide for calculation after completion (in the electronic form this will be done automatically).
iii. The additional cognitive interviews planned at phase 2C, will explore also whether children can tell the differences among the faces, whether they can put the faces in the intended order during a card sorting exercise and how they relate to the crying face (see attached- interview script).

QRT reply (item I-updated user guide): We acknowledge your changes, and we have the following recommendations and comments for your consideration:

- **We agree with using a card sorting exercise to assess if patients can tell the difference between each of the faces in the pictorial scale. Please note that it is important to have the child first complete the card sorting exercise and only then he/she should see the faces in order and answer item 1, which asks about his/her pain severity. Please keep track of the specific faces that children have difficulty differentiating.**
- **Given that the fifth face in the scale shows tears, we believe it is important to properly assess how patients relate to the crying face and whether this face would be endorsed by them as the most severe option. We suggest that patients should be asked the following three questions during the cognitive interviews:**

    1- Have you ever cried from very bad tummy/stomach pain?
    2- Can you ever imagine crying if your tummy/stomach hurt a lot?
    3- Do you believe the tearing/crying face could be you if you ever have very bad tummy/stomach pain?

- **Regarding the two combined scales included in item I of the TUMMY-UC's user guide (pictorial scale and numeric scale), can you please explain your reasoning as to why the scale includes the numbers 0, 2, 4, 6, and 8 as opposed to another type of numeric scale (e.g., 0 to 10; 0, 1, 2, 3, 4, 5, etc.).**
    - **Please note that including a horizontal line across the scale may create confusion for some patients and may encourage them to put a vertical line between the pre-specified hatch marks (e.g., placing a vertical line between 0 and 2).**
    - **If you plan to include both numbers and faces in item I, we suggest including the numbers underneath each face, without a horizontal line running between the numbers, as this may yield more interpretable data.**
- **Additionally, in both the instructions in the cognitive interview guide and the user guide (i.e., TUMMY-UC items), we recommend that the children should be presented with the tummy/stomach pain item during the cognitive interviews in three separate ways as follows:**

    1. **First, with instructions asking them to circle a face;**
    2. **Second, with instructions asking them to circle a number; and**
    3. **Third, with instructions asking them to circle <u>both</u> a face and a number.**

    Patients should then be prompted to justify their choices using probing questions, such as: "Do you think it makes the most sense to circle a face <u>and</u> a number?"; "Do you think it makes the most sense to circle only face?"; "Do you think it makes the most sense to circle only number?" And ask patients to justify their responses by explaining in their own words why they prefer the option they chose.

- **With regard to scoring item I, if you plan to include both numbers and faces in your final version of item I, we recommend that you consider selecting for**

scoring purposes the "most severe" response option regardless of whether a face or number was chosen.

Blood in stool domain:

5. Submitter's response: We suggest leaving the two items in the TUMMY-UC and with the current response options. Indeed, we appreciated the FDA's earlier request at the end of the concept elicitation interview phase to separate the items because they measure two different concepts and prefer to retain the items and their response options as approved by the FDA prior to embarking on Phase 2B. We do believe that both items capture different concepts and thus one response is independent of the other. Binary response will limit the discriminant validity of the TUMMY-UC since it will not differentiate the severity of the disease. Finally, physicians all over the world are accustomed to scoring the pediatric UC bleeding item on an ordinal response option as has been previously validated in the PUCAI.

   QRT reply (item III-updated user guide): We understand your desire to capture concepts represented by items 2 & 3 (old user guide); however, we have the following recommendations for you:

   - **Patients are asked to quantify the amount of blood in their "worst poop." We believe that it may be difficult for children to identify their "worst poop" over the past 24 hours. We suggest that you make the following modification to your item stem: "When thinking about your poop, please choose the best answer that describes the most blood you have seen since yesterday at this time (last 24 hours)?"**

   - **Please note that in previous validation studies where the hospital setting was used, patients may not have pooped directly into a toilet filled with water, but rather into a dry stool collection container (e.g., a hat-shaped plastic container) placed under the toilet seat. We acknowledge that it is difficult to clearly record patients' experiences with examining blood in their stool without any potential variability while reporting. However, given the importance of including patients' experiences with bloody stools for clinical and regulatory decision-making, and the lack of an existing and effective method of data collection in a clinical trial setting, we believe it would be useful for you to test two sets of response options – those currently included in item III in addition to the following response options to further explore patients' understanding and preference regarding the provided choices.**

     a) **I had no poop & no blood**
     b) **I had only poop & no blood**
     c) **There was a mix of poop and blood**
     d) **I had no poop & only blood**

6. Submitter's response: This option has been added, as suggested.

QRT reply (item III-updated user guide): We acknowledge your changes.

7. Submitter's response: We acknowledge the reservation but would like to clarify and emphasize that each item captures a different aspect of the concept. If a child has one stool he will be scored much lower in the stool frequency item. This item has been approved by the FDA before embarking on the cognitive interviews in stage 2B and we respectfully ask to maintain the original approval.

   QRT reply (item IV-updated user guide): We appreciate your response; however, we note that the item stem is assessing the frequency of bleeding in stools while the response options query patients about the proportion of stools containing blood. Therefore, we are concerned that responses to this item will not be interpretable. It is important to include frequency-based response options in order to respond to a frequency-based item stem. We suggest using category-based response options for item IV to be used along with an open-ended response for item II (i.e., patients would enter the exact number of times they pooped). By doing so, patients can answer item IV within the context of what they responded to in item II and can better estimate the number of times they saw blood in their stools in the past 24 hours per the total stool count they had reported having in the past 24 hours.

Stool consistency domain:

8. Submitter's response: The option of "no-poop" has been added to the first category, as suggested.

   QRT reply (item V-updated user guide): We acknowledge your changes. In order to be consistent with your instructions to patients within the item (i.e., "Circle the picture that describes what most of your poops looked like."), we suggest that you change your item stem to "What did most of your poop look like since yesterday at this time (last 24 hours)?" and include the same aforementioned instructions to patients.

Stool frequency domain:

9. Submitter's response: We believe that adding a log to the TUMMY-UC will unnecessarily increase responders' burden since the TUMMY-UC is completed for the last 24 hours only! As found in the previous phases, children have no difficultly in reporting the number of stools in 24 hours, especially when the responses are categorized and not explicit. Indeed, almost all children replied that they have no difficulty in remembering the number of stools in the last 24 hours (>95%). Indeed, pediatric gastroenterologists are accustomed to interviewing children regarding daily stool frequency. Children reply intuitively and without hesitation. Indeed, the reliability of this item, when used as part of the PUCAI, has a very high reliability (ICC>0.9).

QRT reply (item II-updated user guide): We understand your concerns. Please see our response to question #7 above. We recommend that you test an open-ended response for item II with children (i.e., patients would enter the exact number of times they pooped) to evaluate whether they can accurately remember the exact number of times they pooped in the last 24 hours. If they do not believe they can do so, probe children on whether they would rather select a range of frequency from a pre-specified number of response options (i.e., as is currently displayed in item II) and what they believe the range should be based on their experience. Regardless of the response options finalized for item II, we recommend you add a stand-alone "0 times" response option separate from the other response options.

10. Submitter's response: The suggested instructions have been added to the item.

QRT reply (item II-updated user guide): In the instructions for the item assessing stool frequency, currently, it appears that children should count going to the "bathroom" several times over a short period of time (a few minutes) as only one poop. We believe this is confusing and that children may read this to mean that even separate bathroom "visits" over a short period of time should be counted as one poop. We do not agree that this would truly and adequately capture the disease severity.

In addition, the existing item stem does not capture stooling accidents that patients may have experienced (e.g., in their underwear). Therefore, we suggest that you include the following instructions as part of the item stem: "Please count the number of all visits to the bathroom when you pooped, including accidents (e.g., in your underwear or pants). If you poop a few times during one bathroom visit or during one accident, count them all as one time." Please consider adding a stand-alone response option for when patients have not had a bathroom visit or accident during the past 24 hours (i.e., "0 times" as a stand-alone response option).

See our comments above for response #9.

Weakness domain

11. Submitter's response: As detailed in the phase 2B report- The vast majority (82%) of 34 children interviewed in phase 2B could easily distinguish between disease-related fatigue and sleep-deprivation tiredness, saying that disease-related fatigue is different: "Barely able to walk"(child aged 14 years), "Hard for me to breathe"(child aged 10 years), "You can't do even the things you really want" (child aged 11 years). Therefore, we have added this clarification within the question to increase precision of capturing disease – related symptoms. We have reviewed again the previous interviews and found that some of the children are using the term "loss of energy" and thus this has been incorporated in the explanation to the item (see revised TUMMY-UC).

Quotes from children replies at phases 1 (describe weakness in your own words)

Term	Quotes from children replies to the question "describe weakness in your own words"

Loss of energy	Loss of energy (#4001, 14YO); (#8002, 12YO); don't have much energy (#4007, 15YO); Energy goes down. (#4009, 13YO); No energy at all (#4010, 14YO); Put my energy level down because of loss of blood (#4011, 15YO); Low energy. hard to move (#3001, 18YO); Less energy (#3002, 11YO); No energy (#5005, 8YO); decreased energy (#6002, 12YO);

Quotes from children replies at phases 2B:

Term	Quotes from children's replies to different open questions regarding weakness and tiredness
Weakness	I'd say she's more weak when her colitis is acting up (#3009, Caregiver of 6YO); I say weak as a more stronger word to not able to do much, not able to move much. Practically you're lying on the couch not able to move. (#3008, 12YO); Maybe not "energy level" but definitely "weak" and "tired (#4034, 13YO); Interviewer: Do you understand what I mean by "weak"? Child: Yeah, like not really able to walk or concentrate on stuff (#4036, 13YO); When you're tired you're able to open your eyes sometimes, but not able to concentrate. But weak when you're sick, you just can't do anything (#4036, 13YO); Feel like you're low energy, weak (#5011, 15YO); Tiredness, you can go to sleep, and in the morning, you're fine. But weakness – sleeping doesn't always help (#5017, 12YO); Feeling weak is like 'it's hard for me to walk, hard for me to breath (#6005, 10YO); I'd be weak, when my colitis is bad enough (#6005, 10YO); Would you say you are weak when you do not sleep enough and are tired? Or would you say you are weak when your colitis is bad and acting up? Child: When my colitis is bad (#5016, 10YO); I'm weak when my IBD is bad (#8005, 11YO);
Loss of energy	Feel like you're low energy, weak (#5011, 15YO); Out of energy, like you'd have to lie down. (#5016, 10YO);
Tiredness	Tired (#4033, 13YO); sometimes I just feel tired (#4040, 15YO); Drowsy and tired, not wanting to do anything, not wanting to get out of bed (#5014, 11YO); You just feel tired or more stressed (#5015, 16YO);

Other	Exhausted (#4039, 8YO);

QRT reply (item VI-updated user guide): We acknowledge your clarification. However, we assume that the weakness item is intended to be an exploratory item in this study and we believe that the current response options may be "double-barreled." We recommend exploring the following modified response options in your cognitive interviews:

a) Not at all
b) A little bit
c) A lot

12. Submitter's response: Scorings have been removed as suggested.

QRT reply: We acknowledge your changes.

Please note our additional separate current QRT comment below regarding TUMMY-UC PRO item VII:

QRT reply: We recommend changing PRO item VII to "did you wake up to go poop last night?"

TUMMY UC ObsRO Instrument:

General:

13. Submitter's response: We have now added the lower age limit to the ObsRO introduction (i.e. children with UC who are >2 but <8 years of age).

QRT reply: We acknowledge your changes.

14. Submitter's response: Following the request by the EMA and the FDA to improve the intuitive pooling of the PRO and the ObsRO, we have shortened the three subjective items and re-constructed them, so they are now similar to the PRO version. We believe this address also the current concern.

QRT reply: We acknowledge your efforts to shorten the instrument; please also see our response above to question #1.

15. Submitter's response: All changes embedded in the PRO have been embedded also in the ObsRO.

QRT reply: As previously stated, please note that all current QRT comments included above regarding TUMMY-UC PRO are also applicable to TUMMY-UC ObsRO, are not repeated below, and should be applied to the applicable ObsRO items.

Please note our additional separate current QRT comment below regarding TUMMY-UC ObsRO item I (updated user guide), stomach pain domain:

- **Before caregivers complete the item, we recommend that the descriptor (e.g., "no stomach pain", "minimal stomach pain" and etc.) should be removed and caregivers should be asked to put the response options in the order of severity. They should be asked to explain whatever order they chose. The recommended response options are as follows:**

  a) **Acts and plays normally, no complaints of stomach pain**
  b) **Complains infrequently about stomach pain but no other related behavior- acts normally, plays, does not hold stomach and does not curl up**
  c) **Still active but less than usual, holds stomach, seems less cheerful, may complain of stomach pain**
  d) **Stops playing, eats less than usual, complains a lot of pain, lies down from time to time, may hold stomach but does not curl up and no obvious crying**
  e) **Hold stomach, freezes in place or curls up, crying, lots of complains of pain, lies down often**

  After this task is complete, each caregiver should be shown the full proposed item including the underlined portion (i.e., pain severity levels) and then indicate his/her choice based on the level of pain severity indicated by the child's signs/behaviors/verbalizations and be asked whether the response options with the underlined terminology are clear and appropriate.

Instrument instructions:

16. Submitter's response: The current instructions (i.e. "when your child go to sleep") have been revised to "when the child goes to sleep at night", as suggested.

QRT reply: We acknowledge your changes.

17. Submitter's response: The phrase has been deleted as requested.

QRT reply: We acknowledge your changes.

18. Submitter's response: As suggested, the TUMMY-UC now does not include this phrase.

QRT reply: We acknowledge your changes.

Weakness domain:

19. Submitter's response: The phrase "not as happy as usual" has been selected based on the concept elicitation interviews and approved by the FDA before embarking on the cognitive interviews. Nonetheless, the phrase about happiness has been removed as suggested while reconstructing the item to address also the other comments on this item.

    QRT reply (item VI-updated user guide): We acknowledge your changes. Please refer to our response to question #15 and test this item in the same way.

20. Submitter's response: See reply to comment #19. Similar to "happy", we removed also "moodier" as suggested.

    QRT reply (item VI-updated user guide): We acknowledge your changes.

21. Submitter's response: As suggested, 'paler' has been removed, while reconstructing this item.

    QRT reply (item VI-updated user guide): We acknowledge your changes.

22. Submitter's response: We confirm that the items are measuring the same concept between the PRO and ObsRO and both will be pooled in clinical trials analyses. The item has been re-constructed to reflect that and we believe that it addresses the FDA's concern. The same reconstruction has been applied to the other two subjective items in the ObsRO (i.e. abdominal pain and urgency). All three will be now explored from conceptual equivalence during phase 2C.

    QRT reply: We acknowledge your efforts in pooling of the data; however, we refer you to our response to question #1 above.

Urgency domain:

23. Submitter's response: Wording has been revised as suggested as part of the reconstruction of the item to meet also the other comments.

    QRT reply (item VIII-updated user guide): We acknowledge your changes. Please refer to our response to question #15 and test this item in the same way.

24. Submitter's response: We have contacted them and they reviewed all the materials. We have had two teleconferences and email exchange. They have proved very useful and the current submission is a product of their input. Their comments have been embedded throughout. Thank you for the excellent connection.

    QRT reply: Noted

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=13001 70743, cn=Elektra J. Papadopoulos -S
Date: 2019.01.25 15:49:25 -05'00'

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

Jessica J. Lee -S

Digitally signed by Jessica J. Lee -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, cn=Jessica J. Lee -S, 0.9.2342.19200300.100.1.1=20005 96373
Date: 2019.01.25 15:08:36 -05'00'

Jessica J. Lee, MD, MMSc
Associate Director
Division of Gastroenterology and Inborn Errors (DGIEP)
Office of New Drugs
Center for Drug Evaluation and Research

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov