Clinical Outcome Assessments (COA) Qualification Program
DDTCOA #000041: TUMMY-UC
July 15, 2018 Update

July 15th, 2018

Dr. Jessica J. Lee
Dr. Elektra Papadopoulos

FDA

RE: Draft questionnaire and interview script (DDT COA 000041 TUMMY-UC Scale) -**point-by-point response to your email dated June 8th, 2018**

Subject: DDT QUALIFICATION
DDT Type: CLINICAL OUTCOME ASSESSMENT
DDT Tracking Record Number: DDT COA 000041
Submission Type: point-by-point response to your email dated June 8th, 2018
Principal investigators: Dan Turner, Anne Griffiths
Working group members: Anat Nissan (Social Worker, Jerusalem), David Mack (Ottawa, Canada), Tony Otley (Halifax, Canada), Nick Croft (Barts and the London School of medicine), Richard Russell (Edinburg), Michael Kappelman (USA), Seamus Hussey (Ireland)

We would like to thank you for the thoughtful review of our submission of the draft TUMMY index. Please see below our point-by-point replies to your letter trying to address your comments.

Although this has been a long process, we remain committed to improving the application as suggested.

With this letter please find below and attached:

1. **A point by point reply to your comments below. We have bolded and underlined within the replies additional interviews to be completed to meet specified requirements.**
2. **A revised TUMMY-UC based on your comments (and the EMA's)**
3. **A suggested script for the requested limited phase 2C**

Sincerely,

Dan Turner MD, PhD
Director, Juliet Keidan Institute of Pediatric Gastroenterology and Nutrition
Shaare Zedek Medical Center,
The Hebrew University of Jerusalem
P.O.B 3235 Jerusalem 91031; Israel
email: turnerd@szmc.org.il

Tel: +972-2-6666482, +972-50-8685841
Fax: +972-2-6555756

Anne M. Griffiths MD
Co-Lead, IBD Centre,
Northbridge Financial Corporation Chair in IBD
SickKids Hospital
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

POINT-BY-POINT RESPONSE

1. "Our review concludes that you have identified the important and relevant concepts for assessment based on qualitative research. However, the evidence submitted is not sufficient to warrant proceeding to the quantitative psychometric evaluation phase for the TUMMY-UC PRO and TUMMY-UC ObsRO because of concerns around the construction of the instruments (including instructions and items). We are also concerned that the patients interviewed do not represent a diversity of population with UC in the United States. Based on the need for revisions, you should conduct a limited number of additional cognitive interviews to confirm appropriateness of the final measures prior to large scale quantitative testing. The work should preferably be conducted in a more diverse patient population to include patient subsets that were underrepresented (e.g., greater racial diversity including African American children).

   If pooling of data from the PRO and ObsRO is planned, we recommend that you establish conceptual equivalence among the two instruments based on both qualitative and quantitative research. Once the instruments are ready for quantitative testing, you may consider evaluating a subsample of patients who complete the PRO and whose parents complete the ObsRO for comparison of consistency of results."

   Reply: 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.

TUMMY-UC PRO Instrument

Instrument instructions:

2. "The instrument instructions for investigators emphasize completion of the diary at a consistent time daily, but this is not emphasized in the instructions to patients. Please clarify how you plan to address this inconsistency. For example, you may revise the instructions to ensure that the diary is completed at a consistent time of day each day. You may ask for the patient to put the time of day it was completed on the form. We recommend that you develop an electronic diary with reminder functions and time and date stamps for use in the clinical trial context."

   Reply:

   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.

3. The wording, "last 24 hours since yesterday at this time", may be improved by revising to "since yesterday at this time (last 24 hours)".

   Reply: Wording has been improved as suggested.

Abdominal pain item

4. "Your qualitative data showed that some children selected discordant faces and numbers when completing item 1 of the TUMMY UC PRO. We also note that a different item was shown to children in the cognitive interviews-differing faces as well as numbers- from that shown in the final instrument proposed to take forward in psychometric evaluation studies. In general, a pictorial scale is sufficient as it could be understood across a wide range of children capable of providing self-report and could be administered without numbers. However, the fifth face in the TUMMY PRO pain item shows tears, and some children stated that they don't cry even with the most severe pain and, therefore, would not select this response option. Another concern is that some of the faces look very similar. Please address whether children can tell the differences among the faces and whether they can put the faces in the intended order during a card sorting exercise."

   Reply:

   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**).

Blood in stool domain

5. "There is a possibility of inconsistent responses between the two items assessing blood in stool (i.e., items 2 and 3). Ultimately, it is possible that one or the other item may be sufficient to adequately capture this concept. Additionally, it may be possible to adequately assess blood in stool using a binary response ("present vs. absent"). Please comment."

   Reply: 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.

6. “For item 2, there is no response option if the child did not pass stool that day.”

   Reply: This option has been added, as suggested.

7. “The stem for item 3 asks “how many times did you see blood in your stool?”. However, the response options query patients about the proportion of stools containing blood as follows: a. There was no blood in any of my poops (or I did not poop at all); b. There was blood in only some of my poops (half, or less than half of my poops); c. There was blood in most of my poops. This means a patient who had one stool with blood and no other stools that day would score the same as a patient with 10 stools, all having blood.”

   Reply: 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.

Stool consistency domain:

8. “For item 4, there is no response option if the child did not pass stool that day.”

   Reply: The option of “no-poop” has been added to the first category, as suggested.

Stool frequency domain:

9. “We encourage you to consider using a stool event-based log to minimize recall error when reporting on stool frequency, rather than a 24-hour recall item. Particularly, if the frequency is very high, individuals reporting may lose count or may double count from one 24-hour period to the next. Particularly if the reporting period is limited (e.g., 4-7 days), it does not appear to be overly burdensome to complete a stool event-based log.”

   Reply: 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).

10. “Patients are not asked to report on stool frequency until the second to last item (i.e., item 7 of 8). It is helpful to provide instructions on how to count stool frequency (e.g., count the number of visits to the bathroom where stool was passed as 1 stool, regardless of the number of times stool came out during each visit).”

    Reply: The suggested instructions have been added to the item.

Weakness domain:

11. “The item on weakness asks the patient, “we are now asking about weakness related to colitis and not being tired from not sleeping enough.” It’s unclear whether patients can understand the difference and reliably attribute weakness to lack of sleep, colitis or another cause (e.g., anemia). Clarify whether this was documented in concept elicitation or cognitive interviews. While some of the quotes in the February 16, 2015 report refer to “weakness”, many others refer to having less or no energy, feeling tired, missing school or feeling sleepy/cannot wake up. In general, we recommend having patients rate their symptom experience without the need for attribution.”

    Reply: 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, YO); don't have much energy (#4007,15 YO); 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);

All items:

12. “Generally, the scoring algorithm is a separate document outside of the form administered to patients and not embedded in the instrument completed by patients. We do not recommend inclusion of the numbers that will be used for scoring within the copy of the instrument that will be administered to patients as this may distract or confuse patients.”

    Reply: Scorings have been removed as suggested.

TUMMY UC ObsRO Instrument:

General:

13. “It appears that you are developing the instrument for children <8 years of age. Please clarify the target population for the ObsRO for clinical trial use, including whether there is a lower age bound.”

    Reply: 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).

14. “The instrument seems lengthy. We suggest you consider the need for potential item reduction for any redundant or highly correlated items measuring closely related concepts depending on results of quantitative testing.”

Reply: 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 addresses also the current concern.

15. "Please note that many of the comments regarding the TUMMY-UC PRO above are also applicable to the TUMMY-UC ObsRO instrument and are not repeated here."

    Reply: All changes embedded in the PRO have been embedded also in the ObsRO.

Instrument instructions:

16. "The instrument instructs the caregiver to complete the questionnaire while the child is asleep. However, it is more important to answer at a consistent time each day (e.g., when the child goes to sleep at night). This should be clarified in the instrument instructions."

    Reply: 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.

17. "It is unnecessary to include in the instructions "please do not answer the questions for your child as if he/she were answering the questions."

    Reply: The phrase has been deleted as requested.

18. "While not a critical flaw, the rationale for including the phrase "observable behaviors and verbalizations" as the header above the item stems is unclear."

    Reply: As suggested, the TUMMY-UC now does not include this phrase.

Weakness domain:

19. "Within the weakness domain, the phrase "not as happy as usual" is not an observable behavior or verbalization" rather it is a mental state, which is unobservable. The relationship of not being as happy as usual to the concept of "weakness" is also unclear. Please comment."

    Reply: 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.

20. “There appears to be some redundancy among items (e.g., “is not as happy as usual” and “moodier than usual”). The concepts of ‘happiness’ and ‘moodiness’ are internal states and therefore unobservable. If these items are retained, we suggest that they include examples of observable signs or behaviors that lead caregivers to conclude the child as moody or unhappy. Also, consider rewording the item, ‘Is not happy as usual’ to ‘Is not as happy as usual’.”

    Reply: See reply to comment #19. Similar to “happy”, we removed also “moodier” as suggested.

21. “For the item on pallor, the response options do not appear to match the item stem (e.g., “looks paler than usual” is graded on a frequency scale). Its relationship to the concept “weakness” is also unclear. Pallor may be difficult for caregivers to rate accurately and might be better rated on a severity scale than a frequency scale. In addition, patients will presumably have their hemoglobin levels measured as part of clinical trial participation, which would obviate the caregiver’s need to rate pallor at home. Therefore, it is unclear whether the item, ‘seems paler’, is essential to the measurement of weakness and should be retained.”

    Reply: As suggested, ‘paler’ has been removed, while reconstructing this item.

22. “The item, weak/tired/sleepy appears to be inconsistent with the PRO instrument, because patients completing the PRO instrument are instructed not to consider tiredness or sleepiness when responding. We recommend that the concepts are in alignment with the PRO, if the results from the two instruments are intended to be pooled in clinical trial analyses. Please comment on whether the items are measuring the same concept between the PRO and ObsRO. Additionally, it is less likely that a 2-year-old would be able to verbalize feeling tired, weak or sleepy. Therefore, if the caregiver responded “never,” it may be related to the child’s young age and inability to verbalize rather than being a true indicator of the child’s state”

    Reply: 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.

Urgency domain:

23. “Consider revising the following item as follows: ‘When he/she needs to go, it needs to go right now (cannot hold it)’ to ‘When he/she needs to go, he/she needs to go right now (cannot hold it)’.”

Reply: Wording has been revised as suggested as part of the reconstruction of the item to meet also the other comments.

24. "We recommend that you engage with the Critical Path Institute (C-Path) as you prepare your next submission to the FDA. Through a grant provided by FDA, C-Path has agreed to provide DDT development advice for projects referred by FDA. Given C-Path's past and present DDT development efforts and its familiarity with the qualification process, we believe there is benefit to working with them to refine your project's goals and define the necessary components to support a future qualification effort. Note that C-Path is acting in a purely advisory capacity and is not an agent of FDA. As part of this voluntary process, C-Path will not be responsible for creating or submitting regulatory submissions on your behalf. C-Path makes no guarantee of a specific outcome or result by FDA, nor does C-Path guarantee approval by FDA for your future submission. If you wish to contact C-Path with questions or to initiate the external advice process, please email QualificationAdvice@c-path.org."

Reply: 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.