Author’s response to reviews

Title: Exploring the Efficacy of an Electronic Symptom Assessment and Self-Care Intervention to Preserve Physical Function in Individuals Receiving Neurotoxic Chemotherapy

Authors:

Robert Knoerl (robert_knoerl@DFCI.Harvard.edu)

Edie Weller (edie.weller@childrens.harvard.edu)

Barbara Halpenny (Barbara_halpenny@DFCI.Harvard.edu)

Donna Berry (donna_berry@DFCI.Harvard.edu)

Version: 1 Date: 30 Oct 2018

Author’s response to reviews:

See attached word document in supplementary material as well.

Thank you for providing us the opportunity to revise and resubmit our manuscript. We have appreciated all the comments and questions that the reviewers provided to help us make our manuscript stronger.

Major Comment (DL Hertz)

1. One potential concern with any tool that enhances CIPN monitoring is that it will lead to earlier and more frequent treatment disruptions. Side-effect related dose disruption and, separately, neuropathy-related dose disruption should be secondary endpoints and their rates compared between control and intervention arms. The definition of chemotherapy dose reduction (Table 1) should not include disease progression or poor response as those are not side effect related and are not relevant to the intervention. Lines 378-380 mention this as a limitation but it's really a relevant outcome that is critical in understanding the benefits and harms of more intensive CIPN monitoring. The argument can be made that these dose disruptions are appropriate and are themselves evidence of a benefit, not a harm, but this issue must be directly addressed.

Author’s Reply: We have revised our definition of chemotherapy dose reduction to only reflect treatment side effect-related reasons for chemotherapy dose reduction (Table 1, lines 693-707,
2. Looking at table 2, the T3/T4 difference in physical function, the primary endpoint, between the two groups is only 2.7 points. This is less than the pre-specified threshold for clinical relevance (5 points). The abstract results should include the actual scores for physical function in each group, so a reader can judge for themselves the clinical relevance of this finding. I disagree that this finding supports the first sentence of your discussion (line 256) that "ESRA-C preserved physical function" (repeated in your abstract conclusion line 51) or conclusion (line 385) that "ESRA-C result in greater physical function preservation."

Author’s Reply: We have added the baseline and end of treatment physical function scores for each group in the abstract (Abstract, lines 44 – 46, Page 2). We agree with the reviewer that the differences in physical function are modest. We have softened the language related to the findings of our primary aim throughout the paper to not overstate our results (Discussion, lines 273 – 280, Pages 12-13; Conclusion, lines 381-387, Page 17; Abstract, lines 51-54, Page 3). We have added that the mean change scores for physical function reduction in the intervention and control groups were below the minimal clinically significant difference threshold (Discussion, lines 273 – 280, Pages 12-13).

3. Leaving aside the modeling, essentially the main finding of the primary analysis (Table 2) is that the difference in the intervention group (3.0-point reduction in physical function) is statistically significantly smaller than the decrease in the control group (8.5 points). However, 2.8 points of that difference is due to differences at baseline. One would expect the baseline physical function to be similar between the two groups in a randomized trial. Is this baseline difference statistically significant, and if so, is there any explanation for this? Looking at the baseline scores in table 2, for every reported measure the control group is "better" at baseline (higher physical function, less neuropathy, less pain, less depression, less fatigue, less insomnia),
biasing the results toward finding that the intervention prevented worsening for all these measures. This strongly suggests that there's an issue either with the lack of study blinding, both of the patients and presumably of the clinicians who instructed the patient of how to use the tool to complete the assessments at baseline, or with decisions that were made when filtering out patients to be included in this analysis (See comment 6).

Author’s Reply: The full sample in the original trial was well balanced except for age. We are not surprised that there were slight differences in baseline values between groups as the analyzed sample was a subset of the original trial, thus practically all the benefits of randomization are gone. We have compared the baseline scores for all variables in the analyzed sample and there were no significant differences between groups (p > 0.05 via two-sample t test) (Statistical Analysis, lines 208-209, Page 10). While we agree that the lack of study blinding is a limitation and may contribute to the baseline differences, we disagree that the baseline differences inherently bias the results toward finding that the intervention prevented worsening for all outcomes because the differences at baseline were not statistically significant. Nevertheless, we have added the lack of patient and clinician blinding as a study limitation (of note, clinicians were not intentionally told of patient’s treatment assignment) (Limitations, line 378-379, Page 17).

4. Line 165-175: Please define more clearly what features were available to only the intervention group and which to the control group also. Did the control group also receive the self-care messages? Did they also have access to symptom graphs? The way this is described it seems like the only difference between the group is that the intervention group could use the tool at home. It's critical to understand what variables were different between intervention and control to understand exactly what was being tested in this study.

Author’s Reply: We have revised this area of text to more clearly state the differences between the intervention and control group (Procedures, lines 175 – 193, Page 8).

Minor Comments

5. Line 106: It's not obvious how patient self-report and patient-doctor communication will preserve physical function. Is the implication that this tool will detect CIPN earlier, causing more treatment disruptions, and preservation of physical symptoms? That should be stated more
clearly as that may or may not be seen as a positive end-result of use of this tool (related to comment 1). I recognize there is evidence that clinical use of PRO generally has been shown to preserve treatment and improve overall survival (e.g. Basch JAMA) but that seems to be due to earlier recognition of symptoms that are manageable by effective supportive care. This is not currently the case for peripheral neuropathy; earlier detection of CIPN will primarily lead to treatment disruption.

Author’s Reply: The premise of the intervention’s effect on physical function is that the earlier identification of CIPN would allow for earlier management, albeit chemotherapy dose modification or supportive care such as duloxetine or physical rehabilitation. At the time when the original trial was conducted (2009-2011), chemotherapy dose modification was a primary treatment for CIPN. At present day, while treatment disruption due to CIPN is not ideal, it is sometimes a necessary choice to prevent worsening CIPN and physical function. We have revised our description of how we think the intervention may prevent worsening physical function in the introduction (Background, lines 91-94, Pages 4 - 5). Also, like our response to Comment 1, we have added discussion surrounding how the intervention may have worked to improve physical function and/or CIPN (Discussion, lines 294-336, Pages 13-15).

6. Line 122: Please specify what is meant by "all study-related activities." It's important to know who was filtered out of this analysis to determine whether this process was appropriate or could have introduced any bias. A CONSORT diagram of patient flow from enrollment into this analysis would be very helpful, perhaps as a supplementary figure (with reference at line 216). In the CONSORT checklist this is listed as "N/A Secondary Analysis," but this is particularly important in a secondary analysis when only a subset of the primary cohort is included, particularly in this case where there seems to be some bias at baseline in the outcome measures.

Author’s Reply: We have clarified our statement regarding, “all study-related activities” in the Methods section (Subjects, lines 120-122, Page 6). Participants were included in this analysis if they were receiving neurotoxic chemotherapy and completed the baseline (T1) and end-of-treatment (T3/4) outcome measures. We added information to describe the number (frequency and percentage) of participants that were removed from the original trial based on the eligibility criteria of this analysis (Sample Characteristics, lines 224-231, Pages 10 – 11). Due to the simplicity of the eligibility criteria of this exploratory analysis, we believe a text description is sufficient. However, we are happy to add a CONSORT diagram if the reviewer thinks it would add additional information. The CONSORT diagram of participant flow from the original trial is reported along with the results of the primary aim in a separate publication.
7. Line 170: Please provide more detail regarding #2: how the symptom can be managed. There are not effective means of treating peripheral neuropathy, so what was recommended? This also relates to the paragraph starting at line 291, and essentially to the entire objective of this project. If there’s no way to treat or prevent neuropathy, aside from disrupting treatment, how would you expect this intervention to prevent severe CIPN?

Author’s Reply: For all symptoms, the self-care messages contained links from national cancer organizations about self-management strategies for specific symptoms. Related to neuropathy, at the time when the original trial was conducted (2009-2011), there were few recommendations regarding neuropathy prevention and management. Thus, the website links generally contained information about staying safe with neuropathy (e.g., wearing gloves when working with sharp objects, non-slip socks/shoes, be careful with hot liquids). We have updated this description in the Methods section (Procedures, lines 184-187, Page 9).

Like our response to Comments 1 and 5, we have revamped our discussion related to how the intervention worked for CIPN (paragraph that use to begin at line 291) (Discussion, lines 294-336, Pages 13-15). Yes, an intervention like the ESRA-C may work to prevent worsening of physical function due to the early identification and management of cancer treatment-related symptoms such as neuropathy. Of which, neurotoxic chemotherapy dose alterations may be the correct course of action in specific cases, but, additional supportive care interventions may be offered (e.g., balance/strength training, duloxetine). We disagree with the reviewer’s comment that there are currently no effective treatments for neuropathy besides treatment disruption. The American Society of Clinical Oncology Clinical Practice Guideline (Hershman et al., 2014) recommends duloxetine 60 mg as a first line treatment.

8. Line 188: If I understand it correctly, the power calculation does not represent the primary analysis and is therefore not relevant. The study had 80% power to detect an absolute difference of 5 points between what? Is this a five-point drop from baseline to T2 (or T3/T4) in a single group? Is this a 5-point difference in the difference from baseline to T2 (or T3/4) in the intervention group compared with the control group? The primary statistical analysis plan seems to have been to model the difference in the change in physical function from baseline to T2 (or T3/T4) between the groups, but that’s not what was used to estimate power. I would recommend
not including a power analysis as this was a secondary analysis with a fixed cohort size that was going to be conducted regardless of a power calculation.

Author’s Reply: Per the reviewers’ recommendation, we will remove power analysis as this analysis was going to be conducted regardless of a power calculation (Statistical Analysis, lines 199 – 221, Pages 9 – 10).

9. Line 190: It seems odd, in a study focused on patients on neurotoxic chemotherapy, to choose a general physical functioning scale as the primary endpoint instead of the neuropathy-specific scale (CIPN20) that was also available. Was this decision made prior to the analysis, as suggested within the manuscript? If so, please provide justification as it's important to know this to determine whether to consider these results confirmatory or more hypothesis-generating.

Author’s Reply: The choice of using the physical functioning scale (QLQ-C30) instead of the QLQ-CIPN20 motor subscale as the primary physical function measure was determined prior to the analysis. We decided to use the QLQ C-30 Physical Functioning Scale as the primary measure of physical function because this scale is the most commonly used measure of physical function in cancer clinical trials and a myriad of symptoms related to neurotoxic chemotherapy administration may lead to reductions in physical function, not CIPN alone (Measures, lines 134-139, Page 6). Importantly, as suggested in major comment #2, we have toned down the enthusiasm of our results in several locations as there were no differences in QLQ-CIPN20 motor subscale scores between groups and we thought that improvements in CIPN may lead to physical function preservation (Discussion, lines 273 – 280, Pages 12-13; Conclusion, lines 381-387, Page 17; Abstract, lines 51-54, Page 3). As this was an unplanned exploratory analysis of previously collected data, the results are hypothesis-generating and not confirmatory.

10. Table 1: Please describe what is meant by "cumulative dose category," to what "risk" refers, and how these risk categories were defined in the methods. I'm not aware of any established risk categorization for neuropathy, like there is for emetogenicity. In Table 2 the terms "high dose" are used in subgroup analyses of CIPN20 scores. Is this the same categorization? If so, please specify what was considered "high dose" and whether these categorizations were defined prior to analysis.
Author’s Reply: Cumulative dose category refers to the amount of neurotoxic chemotherapy that the patient received over the course of the study (low, moderate, or high). These categorizations were defined prior to the analysis and were based on published literature (Measures, lines 153-158, Page 7). We have provided the range of doses that were considered low, moderate, and high dose for paclitaxel, docetaxel, oxaliplatin, cisplatin in the foot note of Table 1 (Table 1, lines 698-707, Page 32). For example, high dose indicates that individuals were receiving cumulative neurotoxic chemotherapy dosages associated with more severe CIPN (in comparison to the low or moderate dose categories as CIPN severity increases with dose received).

11. Lines 367-369: The data in Table 3 suggests that very few patients in the intervention group triggered any of the self-care messages. If it's not these messages that are causing the differences in outcomes between intervention and control, what could be causing the difference? Again, it was unclear exactly what the differences were in the intervention and control arm and there is concern about the lack of blinding, given the subjective patient-reported outcomes and their bias at baseline.

Author’s Reply: As in our response to Comments 1, 5, and 7, we have added discussion as to how the ESRA-C intervention may have worked to improve physical function and/or neuropathy in individuals receiving high cumulative neurotoxic chemotherapy dosages (Discussion, lines 294-336, Page 13-15). We have added more information about the self-care features available to both treatment groups as described in our response to Comment 4 (Procedures, lines 175–193, Page 8). We have added lack of participant blinding as a limitation as described in our response to Comment 3 (Limitations, line 378-379, Page 17).

12. Line 306-308: Please reword for clarity or remove this sentence. It's not appropriate to state that patients were "experiencing more severe CIPN symptoms" just based on their high cumulative doses when their actual CIPN data is available.

Author’s Reply: We have removed this confusing sentence (Discussion, Page 15). As suggested in Comment #10, we have made sure we are using consistent terminology throughout the paper regarding cumulative dose category.

13. Lines 311-333: While this is interesting commentary, I'm not sure it is sufficiently relevant to this study to be included in this discussion. These sentences describe limitations to the
CIPN20 that are mostly relevant to its use in the clinic. This study specifically attempted to demonstrate the benefit of using an electronic screening tool outside of clinic, so neither the complexity of hand scoring nor the length of the assessment are relevant issues. The PRO-CTCAE does not overcome either of the other two limitations, lack of cutoff scores and content validity, as noted in lines 330-333.

Author’s Reply: We agree with the reviewer and have removed the discussion related to measurement issues as this is not a focus of this study. We have removed our discussion of CIPN measurement and now speculate more on how the intervention may have improved physical function and/or neuropathy in individuals receiving high cumulative neurotoxic chemotherapy dosages. We hope this commentary also helps to address comments 1, 5, 7, and 11 (Discussion, lines 294-336, Page 13-15).

14. Line 356: It's unclear how one of the implications from this study is to use PRO-CTCAE to screen for CIPN since this was not included in this study in any way.

Author’s Reply: We agree with this comment as again, CIPN measurement was not a focus of this study. We have removed the mention of the PRO-CTCAE from the entire manuscript (Discussion, lines 358-359, Page 16).

Formatting

15. Table 1: This table is referred to as "baseline characteristics" but includes outcomes that are not known at baseline (i.e. cumulative dose, chemotherapy dose reduction). Either move these out of the baseline characteristics or change the title of the table.

Author’s Reply: The name of Table 1 has been changed to “Demographic and Cancer Treatment-Related Characteristics” (Table 1, line 693, Page 31).

16. Table 3: I recommend reformatting this table. In the first column you should have the outcome/cut point as the row headers. The next three columns should be T1, T2, and T3/4. In
each of those cells you should put the frequency (%). It's not a major difference, but there's no reason for the extra rows and columns that are currently included.

Author’s Reply: We have made the requested change. However, we created a column for the cut point used to generate the self-care messages. Thus, the columns are now “Outcome, Cut Point, T1, T2, and T3/4” (Table 3, lines 751 – 759, Page 35).

17. Figures: Please informatively label Y-axis of all figures with information as to which outcome/scale is being displayed and indicate which direction is worsening of that outcome.

Author’s Reply: We have relabeled the Y – axis in each respective figure. In addition, for each figure, we have added a note indicating which direction represent worsening of that specific outcome (Figures 1 – 5, see attached figures; Figure Legend, lines 670 – 683, Page 30).