Author's response to reviews

Title: The use of the Godin-Shephard Leisure-Time Physical Activity Questionnaire in oncology research: A systematic review

Authors:

Steve Amireault (steve.amireault@utoronto.ca)
Gaston Godin (gaston.godin@fsi.ulaval.ca)
Jason Lacombe (jason.lacombe@mail.utoronto.ca)
Catherine M Sabiston (catherine.sabiston@utoronto.ca)

Version: 3
Date: 30 April 2015

Author's response to reviews: see over
We thank the editor and the reviewers for their careful reading and constructive criticism of the original submission. We have made numerous changes to the manuscript on the basis of the reviews, and believe that the revised manuscript is considerably stronger as a result of these changes. Our responses (plain text) to each of the specific comments of the reviewers (italicized text) are described below. The revision and the ensuing modifications are identified in the text using a red font color.

**Reviewers’ comments: Siobhan Phillips**

*General Summary*

This study examines the use of the GLTEQ in cancer survivors including the purpose for usage, item content, scoring methods, and validity evidence. While this is an interesting paper that makes a contribution to the literature, I think it could be improved by making the purpose clearer, making it more concise and keeping the language consistent throughout.

**Minor Revisions**

**General**

1. *Please carefully proofread the manuscript for typos as several were found throughout.*

   **Response 1:** The final version of this manuscript was proofread and the necessary corrections have been made.

2. *Please be consistent with your terminology throughout. I had to re-read and refer to earlier sections quite often to make sure I fully understood what was being discussed.*

   **Response 2:** The terminology has been revised and is now consistent throughout the manuscript.

**Introduction**

3. *Page 3, 1st paragraph: I think the last 2 sentences could be switched around to make it flow better.*
Response 3: Done; the last 2 sentences of the 1st paragraph have been switched (p. 3, lines 6-10).

4. Page 4, Line 2: “do” should be “does” and “have” should be “has.”

Response 4: This has been corrected.

5. Page 5, line 9: researcher’s and clinician’s should be researchers’ and clinicians’.

Response 5: This has been corrected.

Methods

6. Page 7, first paragraph: Please make it clear that line 1-11 are the intended scoring and cut-off values, but that many studies alter these items or scoring methods and that you were evaluating these variations. Please also include the exact classifications used for the measurement unit scores instead of adding etc., to the description.

Response 6: We have clarified what are the GSLTPAQ intended scoring and cut-off values. Also, we outline more clearly that we are evaluating variation to these features.

p. 7, lines 2-11: “The GSLTPAQ intended scoring is the LSI, which is obtained using the following formula: \((\text{frequency of mild} \times 3) + (\text{frequency of moderate} \times 5) + (\text{frequency of strenuous} \times 9)\). The intended cut-point values for the classification scoring are based on the North American public health PA guidelines are defined as follows: individuals reporting moderate-to-strenuous LSI \(\geq 24\) are classified as active whereas individuals reporting moderate-to-strenuous LSI \(\leq 23\) are classified as insufficiently active (estimated energy expenditure < 14 Kcal/kg/week) [12]. In order to evaluate the variations in the use of the GSLTPAQ, information concerning the item content (frequency items only vs. frequency and duration items), the recall period (typical or last week vs. other recall timeframes), and scoring methods (LSI, frequency/week, minutes/week, METs \(\times\) hours/week, percentage meeting PA guideline, and other measurement units) were retrieved.”

7. Page 7, lines 11-15: Please be consistent with your terminology. Rather than using type of questions, please say item content and instead of measurement unit use scoring. It would be helpful to the reader to pick one term and stick to it throughout.

Response 7: The terms “type of question” and “measurement scoring” have been respectively replaced by “item content” and “scoring methods” (p. 7, lines 7-11).
8. Page 8, lines 3-6: I would delete this sentence unless the criterion classification is assessed. If yes, please revise so that it is apparent why this information is included. Otherwise, it is distracting.

Response 8: The sentence has been moved to the discussion section (p. 15, lines 15-16).

9. Page 8, lines 10-11: Instead of saying “were retrieved”, I might say that only intervention studies that randomly assigned cancer survivors... were included in this classification.

Response 9: Thank you for this suggestion. The sentence has been modified.

p. 7, lines 18-21: “Specifically, intervention studies that randomly assigned cancer survivors to receive either a supervised and prescribed PA training program (exercise group) or a placebo/non-PA intervention (control group) were examined [38].”

10. Page 10, lines 8-13. The effect size validity estimates are described as being defined as trivial, small, medium or large, originally, but then moderate is used as the descriptor when discussing differences by frequency/duration. Please be consistent. I might also use the term effect size rather than validity estimate in this section to be clear what you are referring to.

Response 10: The term “medium” is now used (p. 10, line 1).

Results

11. Page 10, line 21: Why not just replace “it” with “the kappa coefficient” rather than using the parentheses?

Response 11: Kappa coefficient is now used (p. 8, lines 15-16). Please, note that this information is now reported in the Methods section.

12. Page 12, line 16: Inquire should be inquiry.

Response 12: This sentence has been removed.

13. Page 14, line 2: missing the word “to”

Response 13: Correction made (p. 12, line 17-19).
Discussion

14. Page 18, line 22-23: this sentence is not grammatically correct and seems to be missing words.

Response 14: This has been reviewed and corrected (p. 18, lines 5-9).

Major Revisions

General

15. Why did the authors choose to exclude all of 2014 from their review?

Response 15: We updated the literature search, so that we now include the 2014 articles in the review.

16. I think the manuscript, in general, is rather wordy and think it could be made more concise to fit in some additional important details including some brief info about the search methods.

Response 16: The Methods, Results and Discussion sections have been shortened. Excluding the Title page, the Abstract, List of Abbreviations, Competing interests, Authors’ contributions, Acknowledgements and References, the number of pages of the manuscript has been reduced. It is now 16 pages instead of 21.

17. Page 4, lines 8-9: There seems to be a word missing.

Response 17: Correction made.

p.4 lines 4-6: “However, the questionnaire was not reviewed or analysed by their technical Evaluation Committee, and suggested that it should be viewed “as starting point[s] that can be adapted or improved upon as appropriate” [21].”

Abstract

18. Please provide some clarification of what the GLTEQ is being used to rank or classify in the background section.

Response 18: This has been clarified in the abstract.

Introduction
19. Page, 4, line 3. How do the authors anticipate this “reaching” cancer survivors? Would administered be a more appropriate term?

Response 19: The word “reaching” has been changed and replaced by the word “administered” (p. 5, line 12).

20. Page 4, lines 5-6: I think it is important to clarify that the GLTEQ is recommended as one of the potential measures of PA, since several other measures are also listed on the DCEG website.

Response 20: Thank you for this comment. We clarified that the GSLTPAQ is recommended as one of the potential measures of PA.

p. 4, lines 2-4: “The GSLTPAQ is one of the potential measures of PA that the Division of Cancer Epidemiology & Genetics research program of the National Cancer Institute recommends to oncology clinicians and researchers [20].”

21. Page 4, line 12: Please provide some examples of why you might expect measurement properties to differ in cancer survivors v. general population. I think this is important to the whole purpose of the paper. Otherwise, why does it matter if it has been validated in this population?

Response 21: Examples of why measurement properties might differ between healthy adults and cancer survivors have been provided.

p. 4, lines 10-17: “Cancer survivors’ cognitive abilities (e.g., information processing, attention, concentration, memory) needed for effective recall and reporting of PA may have been impaired by the disease itself or its related treatments [9, 28]. As a result, risk of recall bias may be higher in cancer survivors than in apparently healthy individuals, especially for older and metastatic cancer survivors. Accurate reporting of PA intensity may also be more challenging for someone going through cancer treatment because the perception of PA intensity may not reflect the intensity of a given PA described in the questionnaire [26, 29].”

22. Page 5, line 5: Again, I think the authors need to define what they mean by “ranking and classification purposes.”

Response 22: The terms “ranking” and “classification” are now briefly defined in the introduction section.

p. 3, lines 20-13: “LSI scores can be used for ranking individuals from the lowest to highest PA levels [16]. Also, the score obtained from moderate and strenuous LTPA can be used to classify respondents into active and insufficiently active categories according to published PA guidelines for public health [17-19] and cancer survivors [7, 9].”
23. Page 5, lines 7-8: The authors state in lines 2-3 of this page that the measurement quality of the GLTEQ has not been specifically assessed among cancer survivors. If this is the case, then why bother with iii? Please clarify.

Response 23: Validity is a continuous process that requires accumulating various and appropriate sources of evidence with respect to the intended use of a score among a given population. We acknowledge that the validity of GLTPAQ has not been specifically assessed among cancer survivors prior to 2014, and the third objective of this review is relevant for the following two reasons. First, we anticipated that some studies designed to answer a different research question than one specifically related to validity would report correlation coefficient reflecting the association between the GSLTPAQ and a device-based (accelerometer, pedometer) score, as the use of such devices is increasing. Second, we used the systematic review setting to report sensitivity to change validity evidence for the LSI. This type of validity estimate is rarely reported for most PA questionnaire, in any given population; this is one of the key findings of our study.

24. I know space is limited, however, I think one important point to highlight in the introduction is the importance of this measure in terms of assessment in clinical practice. If the GLTEQ is a widely used, valid, reliable measure, it could result in a number of opportunities for data collection, patient monitoring/surveillance and survivorship care planning as well as outcomes and practice-based/pragmatic research. This may be particularly important given the research on PA and cancer survivorship as it could provide a unique opportunity to provide rich, reliable data on PA in treatment trials and at the population-level throughout the survivorship continuum. I know the authors mention the potential benefits of this measure briefly on page 4, lines 2-3, but feel it could be expanded on to highlight the importance of this paper. I think some of the information in the first full paragraph on page 4 could be condensed as it seems to almost answer one of the research questions (see comment #8, above) and may be more appropriate for the discussion.

Response 24: Thank you for this relevant comment. We now highlight in the introduction the importance of the GSLTPAQ in terms of assessment in clinical practice.

p. 5, lines 11-19: “Since the GSLTPAQ is inexpensive, does not require specific skills for completion or interpretation, and can be administered to a large number of cancer survivors quickly and efficiently, identifying validity evidence can help facilitate opportunities for data collection, patient monitoring and survivorship care planning, as well as outcomes and practice-based research. This may be particularly important for research on PA and cancer survivorship as it could provide a unique opportunity to provide valuable reliable data on PA treatment trials and at population level throughout the survivorship continuum. As such, this study provides research and practical recommendations that will facilitate the researchers’ and clinicians’ decision to use and interpret the GSLTPAQ among cancer survivor populations.”
Methods

25. Page 6, lines 8-12: I find the parentheses a bit distracting. I think the authors could remove them from the first sentence and then use them after and think the authors could move it after the description of ranking and classification to describe the specific purposes.

Response 25: The description of ranking and classification purpose is now briefly presented in the introduction section of the manuscript (see Comment #22). In addition, parentheses have been removed.

26. Page 6, lines 17-20: I find these sentences a bit confusing. Did the authors extract this information to create subclasses for the classification purpose? If yes, how did they determine whether the interpretation was relative or absolute? Was this defined by the authors of this study or those of the studies reviewed? Please clarify.

Response 26: We acknowledge that these sentences were confusing. We did not aim to create subclasses for classification purpose; therefore, these sentences were removed from the manuscript.

Results

27. Page 11, line 13: Is the past behavior used to predict current LTPA? Please clarify.

Response 27: In some studies, past LTPA behavior (assessed at baseline) was used to predict future LTPA behavior (assessed at a subsequent time point). This is now clarified.

p.10, lines 19-21: “Specifically, the GSLTPAQ was most frequently used (frequency ≥ 10%) for identifying correlates/determinants of LTPA barriers, motivation towards LTPA or future LTPA behavior (32.1%)…”

28. Page 11, lines 11-20: I find this very confusing and think it should be broken up into multiple sentences. Additionally, I think it would be better to put the type of study first and then how it is used and would be helpful to include the percentage for each of the categories. (e.g. In studies aimed at identifying the determinants of LTPA, the GLTEQ was most frequently used as a measure of past behavior (XX%), …)

Response 28: Thank you for this suggestion. This sentence has been modified.

p. 10, lines 17-23 and p. 11 lines 1-4: “Specifically, the GSLTPAQ was most frequently used (frequency ≥ 10%) for identifying correlates/determinants of LTPA barriers.
motivation towards LTPA or future LTPA behavior (32.1%); examining the association between LTPA and health-related outcomes (e.g., quality of life, fatigue; 29.7%); evaluating the effectiveness of an intervention (18.4%), comparing baseline levels of LTPA of cancer survivors randomly allocated to one of the experimental conditions (13.2%), reporting and describing PA prevalence (10.9%); or evaluating changes in LTPA levels across the cancer experience (i.e., before diagnosis, during treatment, and after treatment; 10.4%). It is worth noting that a single study can used the questionnaire to more than one purpose.”

29. Page 12, section starting on line 18: Did any studies use the GLTEQ score/cut-points as intended? Since this was one of the main question, I may include these results here.

Response 29: No studies have reported this information.

p. 12, lines 2-3: “There was no study specifically designed to estimate the validity of the GSLTPAQ LSI or classification scoring system in cancer survivors.”

Discussion

30. Page 16, line 14: Can the authors provide some examples of ‘well-tested criteria” for classifying participants that could be used?

Response 30: We initially referred to the GSLTPAQ classification scoring system. We however agree that the terms “well-tested criteria” are not appropriate and have consequently been removed.

31. The authors seem to take issue with the inclusion of time spent in activity. Could this not allow for a more precise estimate of activity participation given that the GLTEQ only asks for bout of 15 minutes or more which could lead people to misclassification of activity level? For example, what if a participant A puts 3 times next to mild activity when they participate in 90 minute sessions for a total of 270 minutes per week whereas participant B puts 7 next to mild when they participate in 15 minute sessions for a total of 105 minutes per week. Participant B would end up with a higher score than participant A, but actually participate in lower levels of activity. I think this is important information and rather than treat this only as a weakness, I think the authors should indicate how this information might be used to understand the measurement properties and validity of the GLTEQ.

Response 31: Public health authorities promote health behaviors, including PA, and their recommendations are mainly based on the frequency of these behaviors (e.g., to perform PA ≥ 3-5 times/week). The intensity of PA is also important, as it is known to influence fitness and health outcomes.
Considering that public health authorities promote regular PA behavior and the development of a PA habit and that many PA benefits can be accumulated in bouts of 10-15 minutes, the duration of PA may be less important. Moreover, the GSLTPAQ was originally designed to: “classify people into several activity [intensity] categories, with the view to examining this aspect of behavior in relation to psychosocial variables before and after implementation of community health and physical fitness promotion programmes” (Godin & Shephard, 1985; p. 142).

We nonetheless acknowledge that in certain circumstances, the duration of PA may be more important. For instance, metabolic variables (such as obesity indicators, risk of diabetes, and hypercholesterolemia) seem to mainly be influenced by the amount of PA. Thus, if the purpose of the study is to examine the association between LTPA and metabolic variables, the use a modified version of the GSLTPAQ and the interpretation of its score in minutes/week scoring might be preferred to the use of LSI (provided that there is validity evidence supporting the interpretation of this score; see Boyle, Lynch, Courneya & Vallance, 2015).

In sum, we believe that the inclusion of time spent in activity should be done in taking into account the research question to be answered and the fact that it may increase the complexity of recalling and reporting the average duration of LTPA. As mentioned in the manuscript (p. 15, lines 1-3), the “frequency × average duration” feature may increase the complexity of recalling and reporting the average duration of LTPA, resulting in increased measurement error.

Reference:


32. I think rather than re-iterating the results, the author could be more concise and really help the reader to understand the implications of these findings and future directions.

Response 32: The discussion section has been re-written and is now more concise. The number of pages is now lower; 5 pages instead of 7.5 pages in the initial submission.

Conclusion

33. This section really seems to encompass future directions more than conclusion and is misleading. I think this label should only be used for a concise summary of the findings and future directions and the other information should be included in the discussion.
Response 33: Thank you for this comment. This section has been re-written and is now used for a concise summary of the findings.

p. 18, lines 9-16: “This systematic review showed that the use of the GSLTPAQ for classification purpose in oncology research is common. Standardization in the use and interpretation of the GSLTPAQ in oncology research is warranted. Although limited, the current state of evidence tends to support the use of the original form of the GSLTPAQ and interpreting the LSI for ranking respondents from the lowest to highest levels of LTPA within a given sample of cancer survivors. Thus, the GSLTPAQ’s LSI may be used in cancer survivors’ studies for (i) identifying the correlates/determinants of LTPA behavior, (ii) verifying whether or not LTPA is a risk or a protective factor of relevant health-related outcomes, and (iii) evaluate the efficacy of LTPA behavior change interventions.”

34. The paper ends abruptly. I think this can be fixed by reconfiguring to put the summary at the end and including the rest in the discussion section.

Response 34: The conclusion section has been reviewed and re-written (see Response #33).

Reviewers’ comments: Lila J Rutten

The use of the Godin-Shephard Leisure-Time Physical Activity Questionnaire in oncology research: A systematic review.

The authors describe a comprehensive review of the literature spanning 1985-2013 examining use and validity of the Godin-Shephard Leisure-Time Physical Activity Questionnaire (GSLTPAQ) in cancer-related research. While in need of some revision and tightening up, overall, the manuscript describes an appropriately conducted review of the literature and offers a nice summary of supported uses for the GSLTPAQ.

Please consider the following observations and suggestions as means for improving the quality of this manuscript:

35- The results summarized in the abstract do not flow with the conclusions given the abstract. The abstract summarizes the lack of evidence gleaned from the review around the validity of the GSLTPAQ while concluding that the tool is appropriate for use in cancer patients. This needs to be rectified.

Response 35: This has been fixed (see p. 2).
36- The introduction does a fine job reviewing the relevant literature and building the case for the need for review. Attention should be paid to the writing in this section as several typos, some grammatical issues, and inconsistencies in use of language were observed.

Response 36: The final version of this manuscript was proofread and the necessary corrections were made.

37- The methods section is far too lengthy. Where possible (i.e., the GSLTPAQ’s Item Content and Scoring Methods section) rely on reference to prior studies for fine details.

Response 37: We have reviewed the writing of the Methods section. The number of pages of this section has consequently decreased from 5 to 4.5 pages.

38- The literature search section belongs in the Methods, not the results.

Response 38: The following paragraph has been moved to the Methods section:

p. 8, lines 15-19: “Intraclass coefficient for study sample characteristics ranged from .98 to 1.00, whereas, for study design and GSLTPAQ-related items, kappa coefficient ranged from .03 to .79. After discussion between reviewers, items for which kappa coefficient was unsatisfactory (< .41) were extracted again from all the included articles by one of the reviewers (SA) and corrections were made as needed.”

39- No need to restate objects throughout the results section.

Response 39: Change made.

40- Tighten up/reduce the length of the results section by relying more on reference to your tables than in text descriptions.

Response 40: We have reviewed the results section. The number of pages of this section is now 3 pages long instead of 4 pages.

41- Be careful throughout the discussion to summarize (but not repeat) the findings.

Response 41: Agree. We have re-written some parts of the discussion and we now summarize, without repeating, key findings. The number of pages of the discussion section has decreased from 7.5 pages in the initial submission to 5 pages in the revised submission.
42- The real gem in this review is Table 3! Highlight the utility of this review in terms of informing future research.

Response 42: The utility of this review in terms of informing future research is highlighted in the text (p. 16, lines 16-22; p. 17, lines 1-19; p. 18, lines 13-21 and in Table 4).

43- The manuscript, with some revision, will make a nice contribution to the literature. Thank you for this opportunity to review your work.

Response 43: We want to thank the reviewer for this comment.

Editor’s comments:

44- We recommend that you copyedit the paper to improve the style of written English. If this is not possible, you may need to use a professional language editing service. For authors who wish to have the language in their manuscript edited by a native-English speaker with scientific expertise, BioMed Central recommends Edanz (www.edanzediting.com/bmc1). BioMed Central has negotiated a 10% discount to the fee charged to BioMed Central authors by Edanz. Use of an editing service is neither a requirement nor a guarantee of acceptance for publication. For more information, see our FAQ on language editing services at http://www.biomedcentral.com/authors/authorfaq/editing.

Response 44: The final version of this manuscript was read by a native English speaker and the necessary corrections were made.

45- Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that your files are correctly formatted.

Response 44: Modifications in 4 references were made to comply with the Journal reference format.