Author’s response to reviews

Title: Developing a coordinated Canadian post-secondary surveillance system: A Delphi Survey to identify measurement priorities for the Canadian Campus Wellbeing Survey (CCWS)

Authors:

Guy Faulkner (guy.faulkner@ubc.ca)
Subha Ramanathan (subha.ramanathan@ubc.ca)
Matthew Kwan (kwanmy@mcmaster.ca)

Version: 1 Date: 12 Apr 2019

Author’s response to reviews:

Thank you for the engaged and constructive feedback on the manuscript. Feedback really helped us address areas where further clarification was required.

Reviewer reports:

Bridgette M Bewick, Ph.D. (Reviewer 1): This is an interesting article. The following points were noted by this reviewer:

1. The title would benefit from inclusion of study design/method used. Currently the title suggests a theoretical piece and not the reporting of an empirical study.

   We have changed to

   Developing a coordinated Canadian post-secondary surveillance system: A Delphi Survey to identify measurement priorities for the Canadian Campus Wellbeing Survey (CCWS)

2. Abstract would benefit from a sentence or two on background literature before stating the aims of the manuscript.

   Context has now been provided at start of abstract

   Interventions that promote health and wellbeing among young adults are needed. Such interventions, however, require measurement tools that support intervention planning, monitoring and evaluation.
3. The appropriateness of the secondary aim is unclear. Given this is an academic piece of work suggest the second aim (as stated in the abstract) is removed.

We have removed from the abstract but it does remain a secondary objective of this work in providing a template for other jurisdictions to follow.

4. Consider re-ordering the 9 core sections of the framework. It would seem sensible to place mental health assets and mental health deficits closer together in the list.

The core sections of the framework are in their ranked order. For example, student experience was considered a higher priority than mental health deficits. We have retained the order as a consequence.

5. Not clear why the introduction begins with 'chronic diseases'. Line 29 which begins 'Poor mental health …' appears to be a more suitable introduction to the background section.

The CCWS is addressing both physical and mental health – alleviating risks of premature mortality and morbidity from chronic disease is a distal outcome of health and wellbeing monitoring systems.

6. Line 77, it is unclear what is meant by 'student-led'.

This refers to ‘student level’ – a survey completed by students. This is in contrast to an institutional level where we may be interested in assessing policies and programmes of an institution – that may be associated with student level variables related to wellbeing.

7. Line 78, the discussion of future work would be better placed in the discussion section.

As with point 6 above, we feel that making brief mention of this future work (not being conducted by us) is important to contextualize the student level nature of the CCWS.

8. Line 82 and beyond, detail would be better placed in the method section.

We agree and have moved to line 357.
9. Line 86-88. It is not clear if this is a purpose/aim of this particular manuscript. Suggest removing this secondary aim.

Yes, this is a secondary purpose of this paper – as a knowledge translation product that can be shared with institutions across Canada so they are aware of the CCWS project. We also hope to provide a template for which other jurisdictions could follow if interested in creating a similar surveillance system.

10. The methods would benefit from clear articulation of the research design. In particular a description of what makes this study a modified Delphi.

We have clarified this to denote as a Delphi Survey which included a follow up roundtable meeting for discussion.

11. Line 92 reports on ethical approval and seems out of place with a section devoted to 'expert panel participants'.

This seemed like a good a place as any to highlight ethics approval was received before describing the sample and how they were recruited.

12. It is not clear from the methods section if a total of 17 participants were invited (or if these are those who responded and a wider pool was invited). The invitation and participation numbers are confused. This section would benefit from greater clarity. I suggest that the number who eventually participated would be better presented in the results and that the methods should only report on the number of participants invited and how they were chosen.

On line 99 we clarify

All invited panelists accepted the invitation to participate.

13. Suggest words such as 'moreover' are removed from the manuscript.

This has been removed
14. It is not clear from the methods section whether or not contributions were anonymous. That is were reminders targeted at individuals? If so the limitations of this approach should be considered in the discussion.

The second author handled the Delphi survey collection process and targeted reminders to those who had not responded the day before and on the day that each survey was due. This is clarified on lines 116 and we acknowledge as a limitation on line 385.

Although panelist contributions were not anonymous to the second author, the indicator rankings, comments and feedback circulated in the second and third rounds of the survey were collated and anonymized so that experts could consider fellow panelists’ perspectives without bias (lines 127).

15. Detail is required on how the four population surveys used to generate the initial items were chosen.

We have removed reference to ‘4’ as we were not restricted to the four mentioned given as examples. We certainly first looked at section headings from the large scale US college surveys (National College Health Assessment, Healthy Minds Study), and the two well-established Canadian population frameworks and surveys - Positive Health Surveillance Indicator Framework from the Public Health Agency of Canada, and the Canadian Health Measures Survey.

16. It is not clear what is meant on line 141 by 'discuss considerations for implementing the CCWS'.

We have provided some examples for clarity on line 144-145 (e.g., dealing with privacy and data sharing; improving response rates).

17. Line 157 suggests that the content included in the final core section framework was 'nearly identical' to the 9 sections generated at the start of the exercise. It is therefore not clear how the rankings were used to build a consensus? The aim of the Delphi exercise is not clear. Is it the case that although the sections were nearly the same the content of each section altered? Greater clarity is needed as to the impact of the Delphi exercise on the generation of content.

The aim of the Delphi exercise was to identify post-secondary institutional priorities for core sections and indicators that may be assessed within a 20-minute student survey (lines 106). As
large scale studies have been deployed in North America with post-secondary students, we chose to draw upon this existing research instead of starting the Delphi exercise with an open brainstorming round for sections and indicators (lines 119-120). We circulated this list for panelists to prioritize and modify to reflect their respective institutional needs.

Given that the nine core areas were derived from existing longitudinal post-secondary surveys, it is not surprising that the framework resulting from the Delphi exercise was nearly identical to the initial list generated by the facilitation team. At the same time, there were important modifications (lines 162-167). The Delphi exercise resulted in changes to the core section labels and content (i.e., indicators) and priority sequence. These directly informed the next stage of developing the CCWS (i.e., identifying instruments that can assess priority sections within a 20-minute timeframe).

18. The results section is difficult to follow. It would be helpful if the authors took sections one at a time and explained how consensus was reached (or not) and what this meant for the generation of content.

We have carefully reviewed each section to ensure the reader sees the link between columns 4 and 5 in table 1. Column 5 identifies the content produced after both Delphi and the roundtable meeting.

19. It would be useful for the reader to understand where content was decided by the remote Delphi exercise and where decisions were largely based on the face-to-face meeting.

The results of the Delphi survey were the basis for all discussion at the face to face meeting. In the results section we narratively describe how the final indicators were decided (in column 5 of table 1) and highlight when salient issues emerged from the roundtable discussions.

At the bottom of table 1 we provide information to help readers understand which indicators received clear consensus among the 3 groups

* all three roundtable discussion groups agreed on the inclusion of this indicator in the final core survey
† one or two roundtable discussion groups agreed on the inclusion of this indicator in the final core survey

20. Can the authors explain what is meant by 'a nuanced approach' (line 182).
We have modified this on line 187 to clarify that a range of considerations will be important during the next stage of CCWS question development for this surveillance tool to be applicable and useful for diverse institutions and populations (e.g., staff; colleges).

21. From line 184 provides a more coherent articulation of the content of the survey. It is not clear how the material before this (lines 156 to 183) sits within the sections detailed. The results section would benefit from a more coherent structure.

This information is necessary to contextualize the discussion and outcomes of the roundtable meeting – the stimulus for that meeting was the results of the Delphi Survey.

22. It is not clear how the Delphi informed decisions. For example Line 187-188 suggests that the lowest ranked item 'sense of meaning or purpose' was included despite the low rankings. It would be helpful if the method section could detail how consensus would be measured and decisions taken so the reader can understand how the final items were decided upon. Is it the case that the rankings put forward in the Delphi were essentially 'overturned' at the roundtable event?

We took a consultative approach and used the Delphi survey as a prompt for discussions during the roundtable meeting. This has been clarified in line 174. Largely, final indicators (column 5) were in line with priorities identified in the Delphi survey (column 4). Not surprisingly when you got everyone in a room to discuss the Delphi survey results, this sparked ideas and conversation! When considering “sense of meaning,” there was consensus that it should receive a higher priority. This was the only clear case of an “overturn” although only in the sense that panelists requested it was given higher priority.

23. Line 203 suggests it is the roundtable discussion that was most important for reaching consensus (and it is not clear how this fits or not with describing the study as a modified Delphi). Without a clear description of the study design it is difficult to comment on the intended and implemented design.

We feel that the Methods section clearly describes the study design: a Delphi survey was conducted; the results of the Delphi survey was used as a prompt for further in-depth discussion; at the end of the meeting, prioritized indicators were identified collaboratively (see column 5, table 1).
24. It would be helpful to have more precise figures on the proportion/number of participants who agreed (or not). For example, line 207-208 'some panellists felt that this indicator should include …' and lines 227-228 'Some but not all groups retained anxiety and depression …'. More precise language when reporting is required throughout.

This is not possible. Individual votes were not taken at the meeting and we are providing an aggregated reporting of whether perspectives were in the majority or the minority. We did vote by group however. In table 1 we provide information to help readers understand which indicators received clear consensus among the 3 groups or just by 1-2 groups.

* all three roundtable discussion groups agreed on the inclusion of this indicator in the final core survey

† one or two roundtable discussion groups agreed on the inclusion of this indicator in the final core survey

25. It is not always clear when results are based on the Delphi and when external factors influenced the decisions taken (e.g. Line 210 'Furthermore, many provinces and post-secondary institutions have recently developed policies and offices around these issues'). A clearer articulation of how decisions were reached is required throughout.

It is not clear what is meant here by decisions. Line 220 provides context as to why possibly there was no consensus among all 3 groups (as indicated in table 1). There was no consensus in the necessity to include questions on experience of sexual assault, for example, because many institutions now have data collection methods for addressing this. However, ‘†Negative experiences (e.g., sexism, racism, violence, discrimination)’ is still retained as an indicator.

26. How consensus was reached it not clear. Nor is it clear how consensus was quantified (e.g. Line 212-215).

We have clarified on lines 189-191

Column 5 of Table 1 (‘Roundtable meeting indicators’) identifies the final indicators as a result of both the Delphi Survey and the roundtable discussions. During the roundtable, participants were invited to vote as a group as to the final inclusion of an indicator (see Column 5, Table 1).

27. It would be helpful if the manuscript could provide an indication of the number of indicators retained or not (instead of a narrative which is difficult to follow Lines 217 onwards). Placing the detail in a table may help the reader to maintain a focus on the key
points being articulated. It is not clear from the text which indicators were kept and why (and whether or not there was consensus about the content to be maintained).

All indicators in column 5, Table 1 were retained.

28. It would be unusual to start a discussion with a suggestion that a paper is the 'first step'. I suggest the authors would be better to state what this individual study has achieved. The discussion of future plans would be better placed later in the discussion.

This paragraph helps contextualize where the study fits within a broader program of surveillance development.

29. The summary of results starting line 301 provides a clear statement of the main findings. The reviewer suggests that a similar summary would be helpful in the results sections. While the reviewer acknowledges that the results will then need to go into further detail it is this overarching structure to the results which is missing and made the section difficult for the reader to follow.

This section is a summary and discussion of the results.

30. Table 1 includes some information of use to the reader but it is not clear what it's purpose is. From the current structure of Table 1 it is not easy to see which indicators were included or not (and which evolved over time).

Comparing column 4 and 5 allows for consideration of how indicators may have evolved between the Delphi and the end of the roundtable meetings. Column 5 presents the prioritized indicators.

It is difficult for the reviewer to comment further on the discussion section due to the lack of clarity around how results were arrived at (i.e. how was consensus reached/decisions made).

We hope we provided greater clarity to the methodological approach adopted.
Nichola Shackleton (Reviewer 2): Thank you for the opportunity to review this interesting manuscript. I believe this is an important paper demonstrating best practice techniques for indicator selection - this is relevant for surveys and indexes. Often, work of this nature is relegated to a technical report or an appendix. Academics and government researchers need to be exposed to these methods with real-life worked examples. The manuscript is very well written and clear, but I have a few minor suggestions for improvement.

Comment 1: More context for using Delphi

I think providing a little more context for the Delphi method might be useful for readers. I thought it might be useful to highlight how the Delphi method fits in with best practice. Highlighting how important it is to get input from subject experts and members of the target population (the student voice panel), how the Delphi method is one of a handful of techniques for synthesizing responses from multiple experts and why this was the most appropriate method to use.

We have elaborated on our choice of the Delphi method from line 106, with a new sentence added beginning at line 112.

The Delphi technique has been widely used for identifying measurement indicators in health and healthcare because it enables synthesis of knowledge from a geographically and experientially diverse group of experts with available evidence. [15,16] Comments and feedback can be shared anonymously, prompting unbiased consideration by panel experts. [15,16] The Delphi approach was also chosen to encourage a sense of ownership of the developing instrument which may promote future institutional uptake of the CCWS.

Comment 2: Discuss the limitations and give examples of lessons learned

If you could include any discussion of 'lessons learned' or things that you would do differently next time in the discussion that would be fantastic. What advice would you give to somebody about to undertake a similar study?

There needs to be some discussion of the limitations of the approach. I have put some questions/comments below here that might help:

How did you ensure that you got the 'right' experts involved? Was there adequate representation - are the 12 post-secondary institutions chosen at random or are they potentially selective in any way? Could the selected priorities be different if a different group of experts from different post-secondary institutions were chosen?
You didn't always get consensus - there is no clear instruction on what to do in this situation and you have chosen to retain some of these. For example, there was less consensus on what to include for academic achievement and nutrition. You made a decision to keep these in the framework and note the degree of consensus - what are the advantages/disadvantages to this? What is your advice?

Sometimes the experts who shout the loudest about their particular area of expertise get more input - how did you handle this? What advice can you give?

Maybe make some comment on the time it took to collate this information and again give advice for speeding up the process.

Thank you for this advice. We have included a paragraph on limitations including some of your suggestions beginning on line 372.

The methodological approach we adopted was successful as a first step in developing a common surveillance mechanism tailored to the Canadian postsecondary context. It also created a space for collegial discussion of the importance of creating a national platform for knowledge exchange and evolving a stronger community of practice. There were some limitations to the approach. First, given budgetary and time constraints, participants were selected through personal contacts of the research team and were individuals who had expressed interest in developing a new student level survey for use in Canada. It is not known how different perspectives may be from other institutions not represented. Further consultation regarding the framework is planned throughout 2019. Second, consensus was not always met on some topics for inclusion. An example was perceptions of the built environment. In such cases where there was no clear consensus, the research team erred on the side of caution in excluding the indicator with a view that future modules could be created to assess topics of interest to institutions. The goal of creating a twenty-minute survey was central to such decisions. Finally, the Delphi Survey was not anonymous as the second author was aware of individual responses throughout the survey process. At the same time, indicator rankings and comments were collated and shared anonymously with the group.

Comment 3: characteristics

It may be useful, if you have the data, to provide some information of the characteristics of the experts and/or the institutions they belonged to.

Characteristics of experts and their institutions were not collected. Experts are included as co-authors and their institutional affiliations will be recognized. No demographic data (e.g., gender) was collected on individual participants.