

Article details: 2021-0113

Title: Estimation of a Canadian preference-based scoring algorithm for the Veterans RAND 12-item Health Survey (VR-12): a population survey using a discrete choice experiment

Authors: Nick Bansback PhD, Logan Trenaman PhD, Brendan J. Mulhern PhD, Richard Norman PhD, Rebecca Metcalfe PhD, Richard Sawatzky PhD, John E. Brazier PhD, Donna Rowen PhD, David G.T. Whitehurst PhD

Reviewer 1: Dr. Mina Bahrampour

General comments (author response in bold)

This manuscript estimates the utility values for the VR-12 for the Canadian population using the DCEto model. This was a well-written paper and the authors have done a great job discussing it.

- Does the background accurately represent current knowledge in this field?
The background of the study is accurate and does represent knowledge in the field.

No response required.

- Do the authors explain why they conducted the study? Is there a clear research question?

The authors have stated that there is an increasing use of the instrument as the reason for valuing the VR-12, however, I would suggest the authors to add a stronger rationale on why they valued this instrument.

We welcome the opportunity to provide a stronger rationale. The development of a value set for the VR-12 means more can be done with VR-12 data, whether the instrument is used in research projects (e.g., randomized controlled trials) or in routine data collection initiatives. Our study enables Canadian preference weights to be derived from VR-12 data. Accordingly, when compared with other health-related quality of life instruments (such as the SF-36 or the Health Utilities Index), the VR-12 is no longer at a disadvantage because health utility values cannot be generated. From a practical perspective, this means that a researcher/analyst who wants to use the VR-12 and generate health utility values does not need to use a second instrument to obtain the latter. We've amended text in the introduction section to strengthen the rationale.(Page 3, para 3)

- Is the study design appropriate?

The study design was appropriate and related to what the authors were aiming for.

No response required.

- Are the methods described in enough detail? Did you find anything confusing?
In the methods the authors mentioned the definition of the classification system, the paper will benefit from a more detailed explanation about the classification system development.

Further details about the development of the classification system have been included in the resubmission. As mentioned in our response to comment #13 (in the list of Editors' Comments), our approach builds on previous work by two of the study co-authors.

- Are the results reasonable? Interesting? Surprising?

The results on anchored disutility weights were somehow surprising and there is a lot of similarity in the values for three levels in RE and RP and some other domains as well. Do you think there is a reason for this?

The values for three levels in RE and three levels in RP are identical, as are the values for adjacent levels for some of the other attributes. This is because these levels were combined to overcome (non-significant) directional inconsistencies. We believe if our sample was larger, we would still find small differences between these levels (whether directionally consistent or inconsistent) and, therefore, it is reasonable to combine them.

- For whom are these findings relevant?

This study is relevant for researchers, health economists, and health policymakers especially when they are planning to do an economic evaluation and a cost-utility analysis.

RESPONSE: No response required.

Reviewer 2: Dr. Holly Witteman

General comments (author response in bold)

Thank you for the opportunity to review this well-organized and well-written paper. The authors conducted an online survey of 3380 people across Canada who completed a discrete choice experiment designed to establish weights for a 12-item general quality of life scale. The study is strong overall and this will be a useful contribution to Canadian health technology assessment. I offer comments below in the hopes that they help the authors improve their paper.

- Sample population: Using a general population is known to undervalue the quality of life of people with health conditions and disabilities. For example, see Polsky D et al (2001) and Dijkers (2005). The authors' lack of acknowledgement of this issue implies that they either do not believe this is a problem, or do not want to step into this argument. If this were in a health economics journal, it would perhaps be understandable not to dive into a major debate, as there's less need to restate a major field issue to specialists. However, for a general readership, this is a critical issue and limitation to acknowledge.

The lack of acknowledgement was because guidelines and recommendations in Canada and beyond (US, UK, etc.) require health preference data to reflect the preferences of the general population (e.g., "Health preferences should reflect the general Canadian population" – CADTH guidelines, point 10.3). However, we agree with the reviewer about the need to highlight the issue for the general readership. We have added text to the interpretation section in the resubmission including the more recent and relevant papers on this topic.(page 7 para 3)

- French: There are three problems related to the French language. First, the survey was translated into French after cognitive interviews were already done. Translation often requires at least some cultural adaptation. Simply translating the words does not guarantee that French speakers in Canada will interpret questions about quality of life the same way as English speakers. Second, relatively few people took the survey in French (8% vs 21% of the population.) These two issues should be added as a limitation to indicate that the findings may not apply in Quebec. Third, related to the previous point, Table 1 should include data about the proportion of the Canadian population whose primary language is English or French.

These issues are all well-received and we have learnt valuable lessons for future work. They were in part due to the limited funding for this study. Problem 1: while there was room for cultural adaptation in the forward and back translation process, we agree the cognitive interviews should have included French speakers. Problem 2: we were surprised how few people completed the survey in French. The market research company was instructed to find a sample reflective of Province, but not language. The framing of the survey (started in English, with an option to take in French at the top of the front page) may have led some bilingual people to continue in English? Also, participants in market research panels in Quebec may be more likely to complete surveys in English to make more surveys accessible to them? Problem 3: we did not ask participants their primary language. All regrets on our part and we agree its important to be clear with the reader. We've added text to the limitations section in the resubmission.(page 7, para 4)

- Section 2.4: What market research company?

The name of the market research company (Ipsos) has been added to the methods section.

- Section 2.5: If the authors are comfortable doing so, it's always nice to see the statistical code in the appendix.

We have included key code in the appendix 1.

- Section 3.1: Please provide the response rate for those randomized to the DCE. The rest of the survey is irrelevant as it is not reported in this paper.

Please see our response to comment #19 (in the list of Editors' Comments). We're grateful to Dr. Witteman and the Editors for highlighting the inadequacy of our descriptions of the sample and the respondents.

- Section 3.1: 83% is an exceptionally high response rate for an online survey administered via a market research company. Are the authors certain that number is correct? Is it possible that they wrote response rate but meant completion rate?

As described in response to comment #19 (in the list of Editors' Comments), our use of 'response rate' was incorrect. We've used some of the additional word limit afforded to us to provide further details in the methods section and results section. The bottom line is that we don't know the true 'response rate' (just like it is not possible to know how many people see a poster inviting you to take part in a research study).

- Section 3.2: Why was the cut-off of 2 minutes chosen? Is this based on analysis of the data? Testing how long it ought to take?

This cut-off was chosen after reviewing a histogram of survey completion times (and discussion between the study team). Also see our response to the Statistical Reviewer, comment #37.

- Data sharing: Evidence is clear that 'data available upon request' is a highly imperfect system, with the majority of requests for data going unanswered. Please deposit data in an institutional data repository. UBC subscribes to Dataverse.

We've been in discussion with the Research Ethics Board of record, and it's been confirmed that we're not able to deposit our data in an institutional data repository

(because of the wording used in our consent forms - it said only aggregated data would be shared to protect anonymity - which is silly and not thought through at the time). We're committed to supporting others to use our data, in accordance with our REB approval. We have amended the statement to reflect how we are willing to work with interested persons (as we have done before).

- Appendix: If it's acceptable to the journal with respect to word count, I would suggest moving the whole first page from Appendix 1 into the main paper.

The allowance of an additional 500 words have been very helpful as we have tried to address the 50+ comments from reviewers. We hope the reviewers understand it was necessary to make decisions about the level of detail needed to address each comment.