PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Elicitation of Norwegian EQ-5D-5L Values for Hypothetical and Experience-based Health States Based on the EuroQol Valuation Technology (EQ-VT) Protocol
AUTHORS	Hansen, Tonya; Helland, Ylva; Augestad, Liv; Rand, Kim; Stavem, K; Garratt, Andrew

VERSION 1 - REVIEW

REVIEWER	Reiner Leidl Munich School of Management, Ludwig-Maximilians-University, Germany
REVIEW RETURNED	19-Oct-2019

GENERAL COMMENTS	There are many EQ-5D-5L valuation studies. This protocol is yet an especially interesting one as it offers two conceptual approaches, the valuation of hypothetical and of experienced health states. In total, the protocol reads well and presents the relevant major steps. However, there are also some gaps, specifically with regard to its two concepts: - The protocol does not provide much rationale for the more unusual approach of elicitating preferences for experienced health states; do authors intend to measure experienced utility (Dolan & Kahneman Economic Journal 2008), or do they follow another aim in the wide range of experience-based measurements (cf. Cubi-Molla et al. Patient 2018)? Authors should consider to compare their approach to existing value sets which use both choice-based approaches and experience-based health. - line 123: It would be helpful to read a comment on how the state of being dead relates to the concept of measuring experience-based health. Minor points - line 73-5: a health state for which preferences do not change with additional time is that of being dead – I think the mathematical logic of this statement reflects QALY calculation but not preferences: consider an individual being indifferent between being unconscious for 13 or for 14 days. - Elaborations on QALYs (lines 78-82) and on NICE (104-106) could be shortened. - line 114 the references of the recommendations mentioned are missing, please list. - line 169 Please mention in case the location strata are intended to be analysed.

	- line 264 is it possible to quote the national reference data? - line 326 ff some of these arguments might help the reader even in the beginning of the text.

REVIEWER	Paula Lorgelly University College London I am a member of the EuroQol Group
REVIEW RETURNED	04-Jan-2020

GENERAL COMMENTS	As this is a protocol, below I provide suggestions for clarification rather than revisions. Particularly as this follows the EuroQol EQ-VT protocol for the elicitation of societal preferences for the EQ-5D-5L. line 32: this reads like it is missing something "statistically comparing values given". line 55: The Article Summary refers to representativeness according to region, the abstract and text on line 130 only refers to age, sex and education level. line 73: offers an unusual/uncommon definition of 0 on the HRQoL scale. It is not incorrect, but more common to refer to it as being dead. line 96: should this read "with the EuroQol Foundation, the organisation that owns the ..." lines 99-106: I think it should just read 3L not -3L and 5L not -5L, although EQ-5D-3L and EQ-5D-5L should be given in full earlier. line 103: I think it would be helpful to expand on the serious deficiencies explaining that the 5L tariff was elicited using the earliest available protocol, and this has now been updated including improving the interviewing/data quality. line 112: missing fullstop. line 142: give EQ-VT in full. line 195: microdata.no needs to look like a website, please add http:// line 219: QC in full line 241: should read worse than dead. line 235: does the completion of the EQ-5D also include the EQ-VAS? line 254: clarify if the DCE includes an own health state? I guess it could by chance. I also guess one of the 10 cTTO states could also be a respondent's health state by chance. Worth adding some text to say what happens in this instance. line 263: how will health status of respondents be compared to national data? what do you mean by health status? EQ-5D profile or other health measures? line 276: will the experience-based values be excluded from the modelling in the first round? that is will it just include 10 responses from each respondent or 11 from some? line 276: how will the experience based valuations be compared? I'm assuming it will be on a profile level first, i.e. if someone is a 12111 then their value for 12111 will be compared with others hypothetical 12111 responses. Or will it be a case of looking where a valuation fits within a distribution? Can you add more detail please? Additionally what happens with the own EQ-VAS data? If this is <100 does that provide another indicator of less than perfect health? line 306: what is NIPH? line 343: do you need to say that the scoring algorithm will be publicly available in say R, STATA, etc?

	line 456. note the position statement has subsequently been updated November 2019. note that the instrument is not reproduced in the paper, should it be?

VERSION 1 – AUTHOR RESPONSE

Reviewer reports:

Reviewer 1: Reiner Leidl

In total, the protocol reads well and presents the relevant major steps. However, there are also some gaps, specifically with regard to its two concepts:
- The protocol does not provide much rationale for the more unusual approach of elicitating preferences for experienced health states; do authors intend to measure experienced utility (Dolan & Kahneman Economic Journal 2008), or do they follow another aim in the wide range of experience-based measurements (cf. Cubi-Molla et al. Patient 2018)? Authors should consider to compare their approach to existing value sets which use both choice-based approaches and experience-based health.
- line 123: It would be helpful to read a comment on how the state of being dead relates to the concept of measuring experience-based health.

We have re-phrased the paragraphs in the introduction concerning experience-based valuation (lines 122-142) and added a paragraph in the discussion (see lines 355-368). We have also added the suggested references.

Minor points
- line 73-5: a health state for which preferences do not change with additional time is that of being dead – I think the mathematical logic of this statement reflects QALY calculation but not preferences: consider an individual being indifferent between being unconscious for 13 or for 14 days.

We agree that this sentence may for some seem illogical, and have chosen to rephrase (line 74-76): "QALY takes the integral of health-related quality of life (HRQoL) over time, with HRQoL represented on a scale where 1 indicates a preference equal to that for full health and 0 implying a health state equal to that of being dead."

- Elaborations on QALYs (lines 78-82) and on NICE (104-106) could be shortened.

We agree in part and have removed elaborations on QALYs, but have chosen to keep comments on NICE. We believe that the data quality challenges uncovered in the England EQ-5D-5L valuation study and later position statement from NICE are important background information to this study.

- line 114 the references of the recommendations mentioned are missing, please list.

This important reference has been added (line 124).

- line 169 Please mention in case the location strata are intended to be analysed.

Results, both regarding recruitment, response rates and data quality, will be compared between location strata when evaluating the data collection. The following text has been added: "Response

rates, recruitment and data quality will be assessed for the different location strata and compared across catchment areas." (line 195-197)

- line 264 is it possible to quote the national reference data?

A table with national reference data from microdata.no has been included (see Table 4).

- line 326 ff some of these arguments might help the reader even in the beginning of the text.

We agree and have moved some of these arguments to the introduction (see lines 109-120)

Reviewer 2: Paula Lorgelly

line 32: this reads like it is missing something "statistically comparing values given".

The text has been revised: "Using a sampling strategy supporting the collection of values for both hypothetical and experienced health states, the study will have the additional aim of assessing the feasibility of collecting experience-based values, and comparing these with those given for hypothetical health states." (line 29-32)

line 55: The Article Summary refers to representativeness according to region, the abstract and text on line 130 only refers to age, sex and education level.
line 73: offers an unusual/uncommon definition of 0 on the HRQoL scale. It is not incorrect, but more common to refer to it as being dead.
line 96: should this read "with the EuroQol Foundation, the organisation that owns the ..."
lines 99-106: I think it should just read 3L not -3L and 5L not -5L, although EQ-5D-3L and EQ-5D-5L should be given in full earlier.
line 103: I think it would be helpful to expand on the serious deficiencies explaining that the 5L tariff was elicited using the earliest available protocol, and this has now been updated including improving the interviewing/data quality.
line 112: missing fullstop.
line 142: give EQ-VT in full.
line 195: microdata.no needs to look like a website, please add http://
line 219: QC in full
line 241: should read worse than dead.

All these suggested corrections have been addressed in the text.

line 235: does the completion of the EQ-5D also include the EQ-VAS?

Yes, respondents complete both the EQ-5D and the EQ-VAS. This is now clarified in the text (line 256).

line 254: clarify if the DCE includes an own health state? I guess it could by chance. I also guess one of the 10 cTTO states could also be a respondent's health state by chance. Worth adding some text to say what happens in this instance.

Since we are using the portable version of the EuroQol Valuation Technology (EQ-PVT), the TTO and DCE tasks are not responsive to previous responses; that is we cannot adjust TTO and DCE tasks to include or exclude certain states. The following text has been added to clarify: "The randomized TTO and DCE tasks do not explicitly include a valuation of the respondents own health state, however

respondents can by chance be presented their own health state as a choice, in which case the task will be completed as normal." (line 274-277)

line 263: how will health status of respondents be compared to national data? what do you mean by health status? EQ-5D profile or other health measures?

Health status of respondents will be measured with the EQ-5D-5L, and compared to health status in a parallel study collecting normative data in the Norwegian population. The manuscript now states: "The demographic characteristics and health status, i.e. EQ-5D-5L profile, of respondents will be assessed and compared to national data. Parallel to this study, the Norwegian Institute of Public Health (NIPH) has initiated data collection for a postal survey assessing the health status of the Norwegian population using the EQ-5D-5L, allowing for comparison of the health status of study populations." (line 285-289)

line 276: will the experience-based values be excluded from the modelling in the first round? that is will it just include 10 responses from each respondent or 11 from some?

The Norwegian value set will be modelled with the 10 TTO tasks and 7 DCE tasks. Valuations from the additional quota used to increase the number of experience-based valuations will be excluded in the first rounds of modelling. These data, as well as the valuation of the respondents own health state (task 11) will be used to compare hypothetical and experience-based valuations. Following text added: "Modelling of values for the national value set will exclude valuations from respondents recruited from locations specifically for the collection of experience-based values and the valuations of respondents' own health state." (line 294-297)

line 276: how will the experience based valuations be compared? I'm assuming it will be on a profile level first, i.e. if someone is a 12111 then their value for 12111 will be compared with others hypothetical 12111 responses. Or will it be a case of looking where a valuation fits within a distribution? Can you add more detail please?

This part of the study will be more explorative in nature, and we are planning to make comparisons between different profiles, however this will be dependent on successful recruitment and sufficient sample sizes per profile. We have however added some detail on potential profiles that will be considered: "To assess experience-based valuations, and explore both the wider and more narrow concepts of experience-based valuations [48], three potential profiles will be assessed; 1) respondents' valuation of own health state, 2) valuations given by respondents recruited from locations specifically chosen to target those with poorer health, i.e. health institutions, 3) valuations given by respondents who have indicated that they have experience with serious illness." (line 303-308)

Additionally what happens with the own EQ-VAS data? If this is <100 does that provide another indicator of less than perfect health?

We agree that VAS data could also prove to be a potential indicator of less than perfect health. Primarily, we plan on comparing respondents EQ-VAS and TTO valuation, with the aim of assessing how respondents use the TTO task to value their own health. See added text (lines 262-264): "Lastly, respondents are administered their current health state as a cTTO task, allowing for the comparison of how respondents value their own health state with both cTTO and VAS."

line 306: what is NIPH?

NIPH is the Norwegian Institute of Public Health – now written in full in the text and included as an abbreviation.

line 343: do you need to say that the scoring algorithm will be publicly available in say R, STATA, etc?

The text now states (line 386-388): “The study results will be published in peer-review scientific journals, presented at appropriate forums, including national and international conferences, and scoring algorithms made publicly available for R, Stata and other widely used statistical software.”

line 456. note the position statement has subsequently been updated November 2019.

The reference is now updated.

Note that the instrument is not reproduced in the paper, should it be?

The self-complete paper version as a pdf download is now referenced in the paper (line 87)

VERSION 2 – REVIEW

REVIEWER	Reiner Leidl Ludwig Maximilians University
REVIEW RETURNED	22-Apr-2020

GENERAL COMMENTS	The authors have addressed all comments adequately. One exception is the issues of how the state of being dead relates to the concept of measuring experience-based health. This only seems to be touched by admitting that “patients … may be less inclined to value their current health state as a state that is worse than being dead.” Understandably, it may be difficult to bring experience and death closer together.

REVIEWER	Paula Lorgelly University College London, UK Member of the EuroQol Group
REVIEW RETURNED	10-Apr-2020

GENERAL COMMENTS	The authors have addressed all of my comments. Not really for addressing this round but an observation about valuation studies for the future is whether they need to undertake a risk assessment for something like a pandemic. I note that the authors are collecting data in the field Nov 2019 to June 2020. I'm assuming this is on hold now due to various pandemic lockdown measures. Once lifted and data collection starts again will these valuations be comparable with those pre-pandemic?