National Institute for Health Research

Draft Final Report Peer Review Form

Policy Research Programme (PRP)

Thank you for agreeing to review this draft final report for the Department of Health Policy Research Programme. The Department very much appreciates you taking the time to undertake this vital part of the research process.

If you have any questions or experience any issues with submitting this form, please contact the PRP CCF on +44 (0) 20 8843 8027

Please return this form by: **19 September 2018**

PRP Reference Number: **104/0001 – Policy Research Unit in Economic Evaluation of Health and Care Interventions**

Project Title: **Quality review of a proposed EQ-5D-5L value set for England**

Reviewer Reference Number: **04**

Reviewer Expertise

Please indicate the nature of your expertise by double clicking on the appropriate tick-box(es) below and selecting 'checked':

- ☐ Patient, service user or carer with direct experience of this field
- ☐ Member of public with a more general view
- ☒ Researcher in the same/a very similar field
- ☐ Researcher in a broadly related field
- ☐ Clinician in the same/a very similar field
- ☐ Clinician in a broadly related field
- ☒ Methodologist
- ☐ Industry professional in the same/a very similar field
- ☐ Industry professional in a broadly related field
- ☐ Other (please see below)

If the tickboxes above do not adequately capture the nature of your expertise, please briefly provide details in the box below (or use it to give us more detail about your expertise if you wish):

(Max 20 words)

Were you in involved in the peer review of the original application for funding for this piece of work?

- ☐ Yes

☒ No

Wherever possible, you will be sent a copy of both the original application form and the research specification that this application responded to, for reference.

Confidentiality

Once you have completed this form, your responses will be used by the Policy Research Programme (PRP) to review the project's findings and assess its outcomes.

Your completed responses are considered confidential, and are therefore exempt under the provisions of the Freedom of Information Act (section 41).

Receipt of this document from the PRP Programme, and your subsequent completed return, form a 'mutual confidentiality agreement' covering your completed responses. This information will not be released without prior consent (except to the parties named below) unless required by law. To read more about our commitment to data security and confidentiality, please click on this link: **http://www.nihr.ac.uk/privacy-policy.htm**

Please treat this report and the associated application form as **confidential.**

How we will use your review

This form will be passed:

(i) **(unattributed) to the Chief Investigator** of the final report you are reviewing with the ratings summary and any conflicts of interest removed (questions 7 and 8 respectively). Please ensure that you **do not include any comments which you would not want to be seen by the Chief Investigator or which could identify you as the reviewer.**

(ii) **(in its entirety) to designated individuals of the Department of Health, who will also be made aware of your name, institution and area of expertise**. These individuals will learn your name in confidence.

Using this Form

Your expertise and experience have been recognised in asking you to review for us and both the researchers and colleagues at the Department of Health would benefit from your comments.

Please rate how the report addresses each of the criteria in the relevant sections below. The prompts are intended to help you focus on the areas addressed by the criteria, but please feel free to comment on additional aspects which you consider to be relevant.

Patient and public involvement in peer review (information for lay reviewers)

Additional information and guidance for patients, service users and carers on peer review is available from the Central Commissioning Facility's website: (**http://www.nihr.ac.uk/about-us/how-we-are-managed/managing-centres/nihr-central-commissioning-facility/ccf-ppi/resources-for-public-contributors.htm**)

An introductory leaflet for patients, service users and carers, entitled 'Getting going as a research reviewer' can be viewed at the following link: (**http://www.nihr.ac.uk/about-us/how-we-are-managed/managing-centres/nihr-central-commissioning-facility/ccf-ppi/**)

We would be grateful if you would provide us with your opinion on each of the following aspects of this study:

1. Evidence that the findings/outcomes of this research are relevant to the issues specified in the original research specification and reflect the programme of work described in the application form

A) Does the report demonstrate a clear understanding of the policy context of the question(s) and are the findings/outcomes clearly linked to this?
B) Are the findings/outcomes of this research placed in context of previous relevant research in this area such that added value is immediately evident?

(max 500 words)
Yes, the report conducts analyses and presents evidence that is relevant to understanding the consistency between the EQ-5D 3L and 5L measures and value sets, and the merits of the composite time trade-off method, data and analysis. The findings are shown in context. The authors received feedback from the persons who ran the EQ-5D-5L study and authored the final value set, in case any evidence was misunderstood.

2. Effectiveness of the research design and work plan (including methods of data collection and analysis) in delivering high quality, fit-for-purpose findings/outcomes for policy-makers

A) Did the research design and methodology prove appropriate to deliver the work described in the application?
B) Were the analyses conducted appropriate in terms of being able to draw robust conclusions?
C) In light of the findings/outcomes presented, are there any particular strengths and/or weaknesses of the research design or analyses chosen?

(max 800 words)
The research team lacked clinical and psychometric experience for the analysis of responses to health-related quality of life instruments, but had a wealth of expertise in econometrics and economic evaluations. Although their study was unable to assess relevance in terms of measurement or clinical practice, the team based its conclusions on their perspectives as economists, which is a strength.

For example, their analysis of consistency did not examine the relationship between items along similar domains, account from recall periods, or use across clinical applications. Nevertheless, what the author did do, they did very well (e.g., 2.2.4 Coverage of states important to cost-effectiveness studies).

The comparison and selection of instruments is typically not based on economic evidence alone. However, the motivation to use the EQ-5D is largely to help inform economic

evaluation of health interventions, and related decision-making. Therefore, such complaints by economists create a strong signal to NHS.

3. Quality of research findings/outcomes
3a. General

A) Has this research yielded clearly defined and appropriately described findings/outcomes?
B) Are the conclusions logical and well thought out, and do they reflect the evidence presented in the report?
C) Can policy-makers have confidence in the quality of the findings presented?

(max 600 words)
Lack of consistency between the EQ-5D-3L and -5L

As the authors point out, each is a five item instrument and covers the same domains and recall period (today). The primary difference is the number and labeling of the levels, which led to three inconsistencies:

1. Slight problems were not captured by the EQ-5D-3L, and this new level diminishes the frequency of "No problems" and "Some problems" responses.
2. Respondents are often confused by the 5L label orderings, for example, "Severely anxious or depressed" and "Extremely anxious or depressed," which decreased the number of extreme responses
3. The worst possible EQ-5D-5L state includes "confined to bed" and the worst possible EQ-5D-5L state includes "unable to walk about." Therefore, the ranges differ.

The authors are correct that the 3L and 5L are not consistent. The selection between the two versions depends on the relative merits of capturing slight or serious problems, and this choice may lead to different conclusions. I agree with the authors conclusions that "the EQ-5D-3L and -5L cannot be used interchangeably if consistent decision making is required."

Deficiencies in the data and statistical methods

The composite time trade-off task was described well in their report. However, the authors did not properly compare it to the TTO task used in the EQ-5D-3L valuation study, namely:

1. Compared to the TTO, the CTTO (CTTO) is inherently different in process and theoretical assumptions. Therefore, they should give different results, which is not addressed by the authors.
(2) The discrete choice experiment was conducted using a separate descriptive system (no time attribute), and its merger with the CTTO may cause further difference.
(3) The EQ-5D-5L survey excluded parts of the UK and used computers with a strictive iterative procedure and well-known interviewer effects, which may cause differences. The differences between EQ-VT protocols across countries is particularly troublesome.

(4) The analyst of the original EQ-5D-3L study (xxxxx) intentionally changed its worst-than-death TTO responses (>25% of the data) prior to the analysis to improve the face validity of the results. This precedent has be criticised by many. The analyst of the EQ-5D-5L (xxxxxx) did not change the data, but manipulated the analysis without justification (changing sample selection, priors, formulae and computational procedures) until it produced the desired results. To this day, no one (even the EQ-5D-5L analyst) can reproduce the same EQ-5D-5L

value set using the same approach and data. These actions by the analysts are more a matter of [xxxxxxx] rather than scientific or methodologic debate.

Apart from the gaps, the report is a logical and well thought out, and policy makers can have confidence in the quality of the findings presented. Nevertheless, "our analyses do not allow us to identify the reasons why the data suffer these limitations. " It is difficult to point to a list of lessons learned from their report, so here are a few:
1. The EQ-5D-5L is inherently different than the EQ-5D-3L, and each has limitations that impede their use in health measurement (i.e., seemingly obsolete).
2. The composite TTO is inherently different than the original TTO, and each has limitations that impede their use in health valuation (i.e., definitively obsolete).
3. The protocol did not specify the quality control, data management and analysis plan properly in advance of data collection, which led to "numerous serious concerns."

Each of these is a fatal flaw. Ideally, a future health valuation study will use a large representative sample, a coherent descriptive system, a simple choice-based task, and a clear protocol. The EuroQol investigators may argue in favor their instrument and method based on precedent, comparability and convention, but this work is not acceptable due to its limitations.

3b. Patient and public involvement

A) Where applicable, was there appropriate patient and public involvement in this research?
B) Did actual levels of involvement reflect those set out in the research application?
C) Could a greater emphasis on patient and public involvement have improved this study?
D) Is there evidence that findings will be shared with patients and the public (or have findings already been shared)?

(max 500 words)
No patient and public involvement is mentioned in the report; however, this seems appropriate.

3c. Intellectual property

A) Where applicable, are you able to identify any intellectual property (IP) produced or improved on during the course of this research?
B) Is there evidence of any plans to protect or exploit any such IP?

(max 500 words)
No IP was produced or improved.

4. Likely impact of the research on policy and/or practice

A) Does the research described in this report have the potential to impact on health care and/or social care and/or public health policy and service delivery?
B) Does this research add to the knowledge base on health and well-being? Does the research add value or advance what is already known in this area?

C) Have target audiences been identified and effective methods of disseminating this research proposed to maximise impact?

(max 500 words)
Yes, if the recommendations are followed, this will be a severe loss for the EQ-5D-5L study team and the EuroQol Group more generally. Any clinical trial that used the EQ-5D-5L will not be able to submit its evidence to inform UK health policy. This report has the potential to make a large impact.

Within the field, the report adds to the knowledge base as an outside perspective (i.e., non-EuroQol members). Its dissemination within health policy is particularly relevant. The task given to the authors was challenging, and the conclusions are largely well founded. As an analogy in clinical trials, these two interventions (3L and 5L) have different mechanisms of action and both trials were conducted poorly, so the reviewers recommend a new trial.

The authors' recommendations assume that NICE and DHSC has an appetite for a new value set after this experience. Alternatively, they may choose to simply continue to use the value set for the EQ-5D-3L and wait for the next version of the EQ-5D. The authors did not provide any motivation to switch to the EQ-5D-5L (new value set or not). Why bother?

Assuming that NICE and DHSC want a new value set, its seems prudent that the new study include a valuation of the EQ-5D-3L and EQ-5D-5L (both descriptive system; head-to-head trial) and not be conducted or reviewed by members of the EuroQol Group or original study team, except in an advisory capacity. Its protocol should be stated in advance and its data and code should be distributed widely. There are many excellent research teams within England who are not members of the EuroQol Group and would be willing to take on a large discrete-choice study of this type.

5. The value of the output relative to cost (if cost of research provided)

A) To what extent is there evidence that the level of resources used to complete this study was appropriate and justified?
B) Has the research provided adequate outcomes for the funds spent?
C) If the researchers have suggested that any financial benefits may arise from their work, please indicate how realistic you feel these might be.

(max 500 words)
No cost information was provided.

6. Suggestions for improvement to the research

A) Do you have any thoughts on how this research might have been improved? If so, please indicate whether you see these as critical factors.

(max 300 words)
1. Involve psychometric and clinical expertise. This flaw is not critical, but would have added to the depth of the comparison, particularly beyond economics.
2. Involve a researcher who conducts health valuation studies regularly (apart from the EQ-5D). None of the study team members have ever produced a value set, so they are more like

hecklers. This flaw is not critical, but it impedes the dissemination of their findings. They seem disgruntled.
3. Comparing English and UK values seems inappropriate, and was not addressed. This is not a critical flaw and would not have changed the recommendations.
4. Differences between ICERs may be due to error in the models that has nothing to do with the different value sets. This was not addressed by the authors, but is not a serious flaw. The report would have benefited from a limitation section.

7. Rating Summary:

Considering your answers to the questions above, please rate how well this final report addressed each of the criteria using the drop down boxes below:

Evidence of a clear focus on the issues specified in the research specification and the work described in the application form	**Excellent**
Quality of the research design and work plan including methods of data collection and forms of analysis in terms of work delivered	**Good**
Quality of research outcomes	**Excellent**
The value of the output relative to cost	**Excellent**
Likely impact of the research on policy and/or practice	**Excellent**

8. Conflict of Interest

We should know about any competing interests that referees may have. Are you aware of any potential competing interests that you may have? If you are in any doubt about any potential competing interest then please declare it. We will not reject your opinion simply because you declare a competing interest, but we would like to know about it, please.

Do you consider yourself to have a conflict of interest with the applicant(s) or institution: No

If yes, please give details:

(max 100 words)

Thank you again for taking the time to complete this review for us.