
Article:
Brennan, V.K., Jones, G., Radley, S. orcid.org/0000-0002-3980-7712 et al. (1 more author) (2021) Incorporating process utility into cost-effectiveness analysis via a bolt-on domain to the SF-6D : an exploratory study. Applied Health Economics and Health Policy, 19 (5). pp. 747-756. ISSN 1175-5652

https://doi.org/10.1007/s40258-021-00646-4

Incorporating process utility into cost-effectiveness analysis via a bolt-on domain to the SF-6D; an exploratory study.

Short running title: Incorporating process utility into cost-effectiveness analysis via a bolt-on domain to the SF-6D.

Authors: Brennan Victoria K[1] vkbrennan@hotmail.com (corresponding author)

Jones Georgina[2] G.L.Jones@leedsbeckett.ac.uk

Radley Stephen[3] Stephen.Radley@sth.nhs.uk

Dixon Simon[1] s.dixon@sheffield.ac.uk

Affiliations:

1. Health Economics and Decision Science, School of Health & Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, UK
2. Faculty of Health & Social Sciences, Leeds Beckett University, Calverley Building, City Campus, Leeds, UK
3. Jessop Wing, Sheffield Teaching Hospitals NHS Trust and Honorary Senior Lecturer, University of Sheffield, Jessop Wing, Tree Root walk, Sheffield, UK

Abstract

Background: Within the UK standard methods of economic evaluation centre on the maximisation of the quality-adjusted life-year (QALY). However, preference-based measures used to estimate QALYs may not suited to all economic evaluations, as they routinely measure only health outcomes.

Aim: This study used an economic evaluation alongside a clinical trial (EEACT) comparing patients' preferences for a telephone versus face-to-face consultation to incorporate process utility into cost-effectiveness analyses.

Methods: An EEACT is described which generates QALYs using SF-6Dv1 responses. These results exclude specific consideration of process utility. A health state valuation study is then reported which bolts a process domain onto the SF-6Dv1 using data obtained from the EEACT. These results therefore include the consideration of process utility. The results of the EEACT with and without process utility are then compared.

Results: This study shows that the QALY, in its current form, does not capture patient benefits associated with the process of receiving health care. The EEACT illustrates this, showing a statistically significant difference between control and intervention groups for the PEQ communication domain, indicating that patients preferred the intervention. This preference was not identified in the cost-effectiveness outcomes and the point estimates lie in the North West quadrant of the cost-effectiveness plane. The preference is captured after adding a communication domain. The point estimate moves to the North East quadrant, where the intervention is more effective and more costly than the control.

Conclusion: This study indicates that it is possible to capture patients' preferences for processes associated with care, in a format compatible with the QALY.

Key words: Process utility, Bolt-on, QALY

Key points for decision makers:

- Within the UK standard methods of economic evaluation focus only on health effects and fail to consider the processes associated with receiving care.
- It is however possible to capture both these health and process outcomes and to incorporate them into cost-effectiveness analyses.
- Including both health and process can impact on the cost-effectiveness outcomes of interventions.

Manuscript

1. INTRODUCTION

Within the UK, standard methods of economic evaluation centre on the maximisation of the quality-adjusted life-year (QALY). This has led to the use of generic preference-based measures (PBM), which can provide utilities required for calculating QALYs [1]. However, the PBMs most commonly used (E.g. EQ-5D, SF-6D) may not be suited to all economic evaluations, as routinely, these measures focus purely on health effects, and do not consider non-health attributes. They are therefore suited only to studies expected to impact on health outcomes and fail to consider non-health attributes, such as those associated with the processes of receiving care [2].

Separate to this problem, health economists tend to focus on a valuation approach that is limited to individuals deriving utility only from the consequences of actions or processes, and not from the actions or processes themselves. This assumption of consequentialism suggests that health care consumers gain no benefit from its

consumption, but that its value is determined solely by the health outcomes it generates. It has been suggested that the processes which patients go through to receive the health outcomes should also be incorporated into the utility function [3,4].

This exploratory study investigated this concept further, using an economic evaluation alongside a clinical trial (EEACT) comparing patients' preferences for telephone versus face-to-face consultation as a vehicle for exploring process utility. Initially the EEACT described generates QALYs using Short-form 6-dimension version 1 (SF-6Dv1) responses, then a health state valuation study which bolts a process domain onto the SF-6Dv1 is reported. Finally, the results of the EEACT with and without the consideration of process utility are reported.

2. METHODS

The electronic personal assessment questionnaire for pelvic floor problems (ePAQ-PF) is a questionnaire providing a detailed evaluation of women's pelvic floor symptoms and their impact on her quality of life [5] that can be used by physicians as an adjuvant to consultations. An online version of ePAQ-PF allows women to complete the questionnaire prior to clinic appointments, facilitating pre-assessment and triage.

2.1 Trial analysis

The EEACT was conducted in England and recruited patients between June 2008 and September 2010. It compared the use of ePAQ-PF in combination with a telephone consultation to standard care (online/clinic-based questionnaire and a face-to-face consultation). Differences in health outcomes were not anticipated, with the focus of the trial being to evaluate differences in patient experience, measured using the patient experience questionnaire (PEQ) and client satisfaction questionnaire (CSQ). Health outcomes were measured using the SF-12. [6]

Costs associated with ePAQ-PF included staff time for the consultation, equipment, software, and overheads. Visits to health professionals, inpatient and outpatient visits over six months' follow-up were included.

The SF-12 was administered to patients at baseline and 6-months post-randomisation. SF-12 results were converted into utilities using the SF-6D UK tariff [7]. Cost-effectiveness was determined by assessing the probability that the intervention was cost-effective at willingness to pay threshold of £20,000 per QALY gained. The results showed that the intervention had higher mean costs (+£15.10, p-value=0.918) and lower

mean QALYs per patient (-0.003, p-value=0.582). There was a 35% probability that the intervention was cost-effective at a £20,000 per QALY threshold. [6]

2.2 Process utility estimation

Whilst no statistically significant difference between the groups was identified for the SF-6D, and hence QALYs, significant differences were found for the PEQ communication domain, indicating that patients preferred the telephone consultation. However, considering the domains included within the SF-6D (physical functioning, role limitations, social functioning, pain, mental health and vitality) which consider only health effects, it is unlikely that the value of this preference associated with a process of receiving care (a telephone versus face-to-face consultation) was captured in SF-6D values based on 6-month follow-up, and therefore, this benefit was effectively excluded from the cost-effectiveness analysis through the lack of a relevant domain [7].

Based on this, a study was undertaken using ePAQ EEACT data to generate utilities relating to consultation type, which allowed incorporation of these patient experience benefits into the QALY. This was achieved through the addition of a process domain to the SF-6D descriptive system. 'Bolt-on' studies have been undertaken previously and for this study, we adapted the methods from Krabbe et al (1999) [8].

The bolt-on study involved 4 key steps:

2.2.1 Classification and description of "SF-6D-process"

2.2.2 Valuation survey

2.2.3 Survey analysis

2.2.4 Application to ePAQ trial

2.2.1 Classification and description of "SF-6D-process"

The process domain was developed using the EEACT findings. Responses to the CSQ and PEQ were explored. As there were significant differences for the PEQ communication experiences domain this was selected to represent process. The communication domain of the PEQ is shown in Figure 1 [9].

A pilot study tested two approaches to the formatting of this domain:

- *Method A* used one question from the communication experiences domain taken directly from the PEQ "The clinician understood what was on my mind" and added it onto the SF-6D health states. The

respondent could select one of the following PEQ ratings: “agree completely”, “agree”, “so-so”, “disagree” and “disagree completely”. If this method was preferred, each patient could be assigned to a “SF-6D-process” health state at baseline and follow-up based on their responses to the SF-12 and PEQ from the trial. This approach would include a domain in a different format to the current SF-6D domains.

- *Method B* used a description of the communication experiences domain as a whole developed by ourselves to match the format of the other SF-6D descriptors. It stated “You have just had a consultation with the doctor and you evaluated your communication experience as …. “very good”, “good”, “average”, “poor”, “very poor”. These descriptors were developed by ourselves to correspond with the existing 5 level responses for the SF-6D domains. However, if this alternative format was preferred, a further step would be required which rates the patients responses for the PEQ domain (patients completed Figure 1) with the descriptors above (very good, good…etc).

The pilot study was performed on a convenience sample within a health economics consultancy company using self-completed questionnaires. Respondents were asked to rate 3 health states using a visual analogue scale (VAS). Each health state was presented twice, once using Method A and once using B. Respondents were asked which approach they preferred. 23 individuals returned questionnaires. All valued health states using both approaches and 15 preferred Method B, which was adopted for the valuation study. A valuation study was used to link the process utility domain descriptors (very good, good, average, poor and very poor) as outlined below.

2.2.2 Valuation survey

Health state values were elicited using a VAS with end-point descriptors of “best imaginable health” and “worst imaginable health”. Respondents were asked where on the scale they would place being dead, and then asked to value selected health states. Obtaining their valuation of “dead” allowed their responses to be transformed, onto a scale where 0 represents “dead” and 1 represents “full health” [10].

The health states were initially selected based on 2 criteria: states reported within the ePAQ trial; states providing an even spread of good, moderate and poor levels. Two additional states differing only in the severity of the process domain were added to test for response rationality. 21 SF-6D states and 23 “SF-6D-process” health states were selected. In a valuation pilot study, respondents found the states very similar, and difficult to value, therefore, less severe health states were replaced with more severe states.

An online survey was developed including three parts:

I. *SF-6D/"SF-6D-process" health states:* Respondents were asked to value either 21 SF-6D health states, or 23 "SF-6D-process" health states, using the VAS.

II. *PEQ communication scenarios:* 25 PEQ communication domain scenarios were selected using an orthogonal fractional factorial design. These were essentially 25 versions of the questionnaire shown in Figure 1 each with a different combination of responses completed. Each respondent was shown the 25 scenarios and they were asked to consider the following statements and select one of the five response options:

"You have just had a consultation with your doctor. You valued your communication experience as: 1) Very poor, 2) poor, 3) fair, 4) good, 5) very good. From these responses, a multinomial regression was used to predict an overall communication domain rating from any combination of individual PEQ communication domain questions.

III. *Demographic questions:* Including: age; gender; ethnicity; marital status; education.

A sample size of 200 members of the UK general population was selected based on Krabbe et al. (1999). [8] Participants were randomly selected through an internet panel company. The sample generated by the company used quota sampling on age and gender to match the UK general population. Respondents were emailed a survey link. Ethical approval for the study was granted by The University of Sheffield (0642/KW).

2.2.3 Survey analysis

Descriptive analyses were undertaken on demographic characteristics. Means and standard deviations for the SF-6D and "SF-6D-process" values for each health state were calculated. Two sets of statistical comparisons were undertaken with the health state valuations. First, differences between SF-6D health states and their analogous "SF-6D-process" health states were tested, (E.g. 123123 vs 1231235). Second, differences between the "SF-6D-process" health states which were identical except for their process level were tested. Both statistical comparisons were performed using Mann-Whitney tests. Descriptive analysis was used to explore the relationship between the level of the added domain and decrement. Initially the mean decrement for each of the 5 process domain levels was reported. In order to further explore the relationship between the decrement and the severity of the health states, the number of domain levels was

reduced from 5 to 3 (poor, fair and good) by combining the two upper domains, and two lower domains, and maintaining the middle domain.

2.2.4 Application to ePAQ trial

Results from multinomial regression of PEQ responses to the online survey were used to predict an overall communication domain rating for each set of trial PEQ responses. This predicted rating was combined with the estimated process utility decrement produced from the multiple linear regression coefficients described in (2.2.3), above. This was used to adjust the SF-6D values produced following the ePAQ consultation to incorporate process and produce an "SF-6D-process" value. The number of QALYs gained based on the SF-6D values, and "SF-6D-process" values was then calculated. An incremental cost-effectiveness ratio (ICER) and cost-effectiveness acceptability curve (CEAC) were reported using SF-6D and "SF-6D-process" values.

Further analysis was undertaken to account for the handling of missing data. The MI command within Stata® was used, with the number of imputations set equal to the rate of missing data within the least complete variable [11,12]; This led to 70 imputations being used. Costs and change in QALYs from baseline were jointly imputed using patient age, parity and treatment group as explanatory imputation variables. A seemingly unrelated regression (SUR) was then undertaken across the 70 resultant datasets with the cost-effectiveness confidence ellipses and CEAC being generated parametrically using means, correlations and standard errors generated from the output of the SUR estimated across the 70 datasets [13].

3 RESULTS

3.1 Survey Results

Of the 609 respondents who began the survey, 268 completed, resulting in a dropout rate of 56%. Of the 268 who completed, 8 were excluded for the following reasons: 3 provided responses deemed to be irrational:1 valued all health states as 100; 1 valued all health states as 0; 1 valued dead as 100; 2 valued dead as 0 and all other health states as 100.

Of the remaining 260, 130 completed valuations of the 21 SF-6D health states, and 130 completed valuations of 23 "SF-6D-process" health states. Table 1 presents the background characteristics of the 260 respondents. Mean health state valuations are shown in Figure 2.

Mean process decrements for the 5 process domain levels indicated that as the rating declined the value of the decrement did not reduce monotonically but suggested there may be a trend of greater decrements for worse process domain levels (Table 2). When the number of domain levels was reduced from 5 to 3, as the rating level declined (from 1 to 3) the value of the decrement reduced monotonically (Table 2).

3.2 Application to ePAQ trial Results.

Multinomial regression was run on the rating responses from the online survey to determine what combination of PEQ scores would relate to which rating score. Each PEQ health state was rated between 269 and 272 times. The chi-square statistic for the main effects model was significant (<0.001), indicating that the factor-by-factor interactions between the PEQ responses had a significant effect on predicting the rating of the health state. Likelihood ratio tests were significant for all predictors (PEQ responses) indicating a good ability to predict the overall rating of the health states. The observed and predicted rating scores for the PEQ health states were largely similar; again indicating a good model fit. The co-efficients from the multinomial regression model were applied to the ePAQ trial data to estimate the probability of membership into each rating score group based on combinations of participants PEQ responses.

The analysis applying the multiple linear regression coefficients to the ePAQ trial data included all ePAQ trial patients with complete cost, SF-6D and PEQ data, and had a rating score calculated for the communication domain based on their PEQ responses. This resulted in 26 patients from the intervention and 27 from the control group (Table 3).

As shown in Table 3, Mean total costs were calculated for each trial arm. Costs were slightly higher in the intervention than the montrol (£15.17). Mean SF-6D values were calculated at baseline (t_1) and 6 months (t_2). The mean process level was calculated based on the application of the multinomial regression. The mean severity was calculated based on patient's baseline SF-6D. The use of the "SF-6D-process" reduced the QALY loss associated with the intervention group from 0.0097 to 0.0070. Both analyses resulted in negative ICERs.

With missing data imputed and analysis undertaken within a SUR model to account for the correlation between costs and QALYs, QALYs and costs are slightly higher in the intervention (0.0026 and £4.31 respectively (Table 4) resulting in an ICER of £1657.69/QALY).

Figure 3 shows the "SF-6D-process" cost-effectiveness plane and CEAC alongside the SF-6D results which were estimated in the original EEACT. The "SF-6D-process" point estimate lays in the North East (NE) quadrant of the cost-effectiveness plane. The confidence ellipses also shift right, into the NE quadrant and shows there is approximately 60% probability the intervention is cost-effective at a threshold of £20,000/ QALY.

4 DISCUSSION

This paper highlights that QALY as generally constructed are unlikely to capture patient benefits associated with the processes of receiving care. In this study, there was a statistically significant difference between control and intervention groups for the PEQ communication domain, indicating that patients considered that the process was improved with the intervention. We attempted to capture preferences associated with this improved process by adding a communication domain to the SF-6D. The VAS values associated with the domain showed decrements as communication experience worsened, and when combined with the SF-6D utilities, the cost-effectiveness of the intervention improved.

4.1 Limitations and strengths

This was an exploratory study, and certain decisions were made to simplify or overcome methodological barriers to ascertain the feasibility of undertaking further analysis.

In the description and measurement of the process domain, a single PEQ question was used to capture a key difference in communication that was found to be statistically significant. Alternative methods for describing and measuring the process domain could produce different results.

The format chosen to represent the PEQ domain within the "SF-6D-process" health states aimed to make the elicitation exercise easier for respondents. However, it required an additional statistical mapping exercise converting PEQ responses into the process domain description within the "SF-6D-process" health states. This introduces uncertainty around the statistical model used for the mapping and methodological uncertainty around the validity of the new scale.

The valuation study did not use a choice-based approach, which is considered by many to be essential to derive utilities, and hence QALYs. However, this view has been disputed by some with several studies using the VAS for eliciting utilities [13, 8]. Likewise, the use of VAS-based utilities within cost-utility analysis is widespread. Given this uncertainty, the choice of VAS was pragmatic; we felt that the ease of completion would help respondents focus on the health and process aspects of the study.

It must also be recognised that adding a process descriptor to health state classification mixes two different types of information. The validity of such an approach needs further examination if this approach is to be adopted more widely. However, we note two important findings in this regard. First, we conducted a think aloud exercise after this work to assess whether respondents understood the valuation process. This found that respondents were able to provide the information required to value health states combining health and process outcomes using the VAS or time trade off [2]. Second, a previous review of process utility studies identified several studies that have included process information into elicitation tasks [19].

Certain simplifications were made in the survey analysis. The number of process domain levels was reduced from 5 to 3, so a clear monotonic relationship was observed for process levels. Keeping all five levels was possible but would be based on smaller sample sizes within each category and could lead to counter-intuitive results, whereby worse process produced increased utility. Merging of response categories to maintain monotonicity has been used in several studies estimating tariffs for instruments [15]. As this study aimed to explore the feasibility of adding and valuing a measure of process utility to the existing descriptive system, this simplification was considered appropriate.

Within the health state valuation study, the health states were initially selected to provide an even spread of good, moderate and poor levels. However, in a valuation pilot study, respondents found the states very similar, and difficult to value, therefore, less severe health states were replaced with more severe states, therefore emphasising the difference between states. Despite this amendment all health states included within the valuation study were those experienced by patients in the PAQ EEACT. As within the EEACT there was still a greater number of less severe health states. the findings would still be representative of patients included in the ePAQ EEACT.

QALYs were calculated as the difference in the change in QALYs from baseline for both arms in order to control for baseline utility. Whilst some argue for the statistical adjustment of QALYs using baseline utility within a regression, this is problematic in the context of this study. This is because the comparison between the SF-6D and "SF-6D-process" analysis was a combination of the statistical adjustments for the two regressions, plus the differences in the calculated utilities. Using change from baseline removes the effect of a differential statistical adjustment, thereby giving a clearer picture of the effect of using the two different utility calculations.

Despite the favourable ICER, there remains uncertainty due to the small cost and effect sizes, as well as the impact of missing data. There is also considerable methodological uncertainty around the "SF-6D-process" estimates. Considering these simplifications, the findings from the valuation study, and the revised cost-effectiveness estimates should be interpreted as an indicator of the presence and magnitude of process utility as opposed to absolute values. Despite this, the objective of this analysis was not to produce definitive cost-effectiveness recommendations; it was to examine the potential for differential cost-effectiveness after the inclusion of exploratory process utility estimates. In essence, this application of the revised "SF-6D-process" estimates into the ICER highlights the potential impact of adding a process domain to the SF-6D on cost-effectiveness outcomes.

There may be some double-counting of the value of the telephone or face-to-face consultation within this study. People's preferences for the alternative modes of consultation may have been captured by differential costs within the EEACT (e.g. travel costs or productivity). Although not relevant to this specific study - which aims to ascertain whether a process domain can be added to the SF-6D, if this work were taken further it should be investigated further.

There has been relatively little progress over time with incorporating process utility into economic evaluations, despite its early advocacy by influential health economists such as Gavin Mooney, whose work into proceduralism and process utility spanned over 15 years [3, 4]. This may be for practical and methodological reasons. Practically, it is difficult to develop a standard definition of process utility, particularly as it can include so many aspects of health care, ranging from treatment attributes to patient dignity, [17,18], each of which potentially require subjective judgement for their evaluation. The disparate nature of process utility therefore makes it hard to develop standardised and recognised methods for its incorporation into economic evaluations.

Further barriers to new research on process utility are thought to be the perceived relevance of process utility to decision makers and its relative importance compared to health outcomes. In terms of relevance, a narrow extra-welfarist view may consider process utility to be outside the perspective of economic evaluation [18] As for relative importance, this study shows whilst the value attributed to process may be small, it may also be sufficient to change the ICER to a significant degree. Whilst several studies have produced large estimates of process utility, the methods used have been criticised [19].

This is the first bolt-on study adding a domain to the SF-6D. Previous studies added domains to the EQ-5D [8,20] using a much simpler descriptive system. This study suggests that the 'bolt-on' approach is generalisable to more complex instruments such as the SF-6D.

This study used part of an existing validated questionnaire (PEQ), used within an EEACT to measure process utility. This approach provides a direct link between the individual patient data and the process utility decrements included in the economic evaluation. This link was not present in previous studies where bespoke process descriptions were generated and mean values applied at the patient population level [19].

This study builds on work by Krabbe et al. (1996) who also used the VAS to bolt on a cognition domain to the EQ-5D. [8] Where Krabbe et al. (1996) used a convenience sample this study used a market research company to obtain a sample which was broadly reflective of the general population. [8]

4.2 Comparison with other studies – bolt- on and process utility

The results of this study display some face validity, with overall worse health states having lower utility values. The results also indicate a greater decrement with worse process, although it failed to show whether the decrements were greater in more or less severe patients, as found by Brazier et al. (2010). [21] The findings do however support the findings by Brazier at al. (2010) whereby the addition of a domain, whether good or bad, lead to a reduction in the health state valuation in their study, which added a pain dimension onto an asthma questionnaire. They suggested this may be due to a "focusing effect", however, the utilization of PEQ data collected from with the EEACT within our study aimed to reduce this focusing effect, and increase the internal validity, there may therefore be alternative reasons for this trend which we are unable to identify within this study. [21]

4.3 Recommendations for future research

This study suggests it is possible to capture process through a bolt-on study. However, given its exploratory nature there are areas requiring further work. In common with many bolt-on studies, there is an assumption that the additional dimension does not impact on the tariff relating to the other dimensions. However,, this has been questioned by some, with evidence suggesting that the full utility tariff needs re-estimating. Our study lends some support to this by indicating an interaction between the bolt-on and 'core' dimensions. If bolt-on instruments, in general, are to become widely used, this assumption needs to be assessed in greater detail.

In addition to this quantitative work, further qualitative work would be valuable. The number of irrational responses identified in this study highlighted a potential concern. Insight into respondents' decision-making processes, gained through qualitative research, may provide important insight into the underlying reasons for this, and allow us to understand more fully the utility estimates derived.

In addition, this study was based on VAS valuations of state/process description, when typically, choice-based valuations of states are used. In theoretical terms, this casts doubt on the validity of the utilities derived and their ability to generate QALYs. In practical terms, we do not know whether people will be willing to trade-off health for different levels of process. A further examination of these issues is required.

5 Conclusion

This study showed it is possible to capture both health and process outcomes, in a format which is compatible with estimating QALYs. It indicated that elements of a patient's healthcare utility function are not captured by conventional health utility measures. The approach used allows a direct link to individual patient data, which is different from other methods employed to date. The study also showed the feasibility of bolt-on studies using the SF-6D descriptive system. Further research is required to identify the best way for capturing process utility for cost-effectiveness evaluations.

Declarations

Funding: This work was partially funded by RTI Health Solutions through a supporting PhD fund. Funding to the online survey was obtained from The University of Sheffield, School of Health and Related Research data collection fund.

Conflict of Interest: S Radley is an unsalaried director and shareholder in ePAQ Systems Ltd, an NHS spin-out technology company; the largest shareholder in which is Sheffield Teaching Hospitals NHS Foundation. VK Brennan, S Dixon, G Jones and S Radley declare that they have no conflict of interest.

Ethics approval: The EEACT study protocol was approved by the North Sheffield ethics committee and was registered at www.clinicaltrials.gov (identifier NCT02176330). The process utility valuation study gained ethics approval from The University of Sheffield Ethics committee.

Consent to participate: Informed consent to participate in the study was obtained from all participants.

Availability of data and material (data transparency): Information on the availability of data from the ePAQ EEACT is reported within the relevant publication (Jones et al, 2018)[6] . For the bolt-on study, ethics approval and patient consent were gained only for results to be shared through publications in academic journals and presented at conferences and can only be published in aggregate form. These data sets are therefore not publicly available. Data enquiries may be directed to Victoria Brennan: email vkbrennan@hotmail.com.

Code availability (software application or custom code): Coding from the analyses can be provided by the authors on request. This includes STATA files for the SUR analyses and production of the cost-effectiveness planes and CEACs, and in SPSS format for the regression analyses run. Coding enquiries may be directed to Victoria Brennan: vkbrennan@hotmail.com

Authors' contributions: Victoria Brennan, was involved in study conception and design, data collection, analysis and interpretation of results and drafted the manuscript. Simon Dixon was involved in study conception and design and advised on the analysis and interpretation of results. Georgina Jones and Stephen Radley lead the EEACT. All authors have reviewed the results and approved the final version of the manuscript.

Acknowledgments: We also thank the reviewers for their comments on the draft manuscript.

References

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. Oxford Medical Publications; 2015.
2. Brennan VK. A Study of Process Utility. 2016. http://etheses.whiterose.ac.uk/id/eprint/15202
3. Mooney, G. (1998) The benefits of health care. Health Care Analysis, 1998;6(2):99-101.
4. Gerard K and Mooney G. QALY League Tables: handle with care. Health Economics, 1993:2,59-64.
5. Jones GL, Radley SC, Lumb J, Farkas A. Responsiveness of the electronic Personal Assessment Questionnaire-Pelvic Floor (ePAQ-PF). Int Urogyecol J Pelvic Floor Dysfunct 2009; May20(5)557-64.
6. Jones G, Brennan V, Jaques R, Wood H, Dixon S, Radley S. Evaluating the impact of a 'virtual clinic' on patient experience, personal and provider costs of care in urinary incontinence: A randomised controlled trial. January 18, 2018 ttps://doi.org/10.1371/journal.pone.0189174.
7. Brazier, JE. Roberts, J. The Estimation of a preference-based measure of health from the SF-12. Medical Care, 2004;499:851-859.
8. Krabbe P, Stouthard M, Essink-Bot M, Bonsel G. The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. J clin Epidemiol 1999; 52(4):293-301.

9. Steine, S. Finset, A. Laerum, E. (2001) A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. Family Practice 18,410-418.
10. Torrance, GW. (1986) Measurement of health state utilities for economic appraisal. Journal of Health Economics, 1986; 5:1-30.
11. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics, 2014; 32(12):1157-70.
12. Gomes M, Grieve R, Nixon R, Ng ES, Carpenter J, Thompson SG. Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. Health Econ. 2012;21(9):1101-18.
13. Hoch J, Briggs A, Willan A: Something old, something new, something borrowed, something blue: A framework for the marriage of health econometrics and cost-effectiveness analysis. Health Economics. 2002;11,415-430.
14. Parkin D, and Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis. *Health Economics*, 2006; 15, 653-664.
15. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; Mar;21920:271-92.
16. Donaldson C and Shackley P. Does "process utility" exist? A case study of willingness to pay for laproscopic cholcystectomy. Soc Sci Med, 1997; 44, 699-707
17. Swan J and Sainfort F. Process Utility for Imaging in cerebrovascular disease. Acad Radiol, 2003:10(3), 266-74.
18. Birch S, Donaldson C. Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? Soc Sci Med. 2003;56(5):1121-33.
19. Brennan VK. and Dixon S. Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies, Pharmacoeconomics, 2013;31(8):677-91
20. Yang, Y. Brazier, J. Tsuchiya, A. The effect of adding a 'sleep' dimension to EQ-5D. Health Economists' Study Group meeting January. 2008.
21. Brazier, J. Rowen, D. Tsuchiya, A. Yang, Y. Young, T. (2010) What a pain: adding a generic dimension to a condition-specific preference-based measure. *HESG Abstract*.

Table 1. Background characteristics of survey sample

Background characteristics		National census data (2011)[1]	SF-6D (n=130)	"SF-6D-process" (n=130)
		%	n(%)	n(%)
Age (years)	18-39	40.5	34(26.1)	38(29.2)
	40-59	32.4	33(25.4)	32(24.6)
	60-75+	27.1	63(48.5)	60(46.2)
Gender	Male	49	71(54.6)	63(48.5)
	Female	51	59(45.4)	67(51.5)
Education level	O levels/GCSCs and equivalent		56(43.1)	60(46.2)
	Degree and higher degree		43(33.1)	37(28.5)
	Professional qualifications/other vocational related qualifications		22(16.9)	23(17.7)
	Foreign qualifications		2(1.5)	0(0)
	No qualifications		7(5.4)	10(7.7)

Notes: [1] Calculated for age 18-24 based on 2011 Census data. Based on 2011 Census data tables: Usual resident population by 5-year age group and sex. SF-6D = short form six dimension; "SF-6D-process" = short form six dimension plus process.

Table 2: Mean Decrements for the 3 and 5 process domain levels

5-level domain description	n	Mean	SD	3-level domain description	n	Mean	SD
Very good	4	-4.73	2.82	Good	9	4.53	2.45
Good	5	-4.37	2.43				
Fair	4	-5.78	2.40	Fair	4	-5.78	2.40
Poor	4	-7.90	4.01	Poor	9	-7.10	3.34
Very poor	5	-6.46	3.01				

Table 3: Application of process utility to the ePAQ trial data.

	Intervention N = 26	**Control N =27**	**Difference**
Total costs	£1,169.45	£1,154.28	£15.17
SF-6D t_1	0.6445	0.6175	0.0270
SF-6D t_2	0.6334	0.6259	0.0075
Change in SF-6D	-0.0111	0.0084	-0.0195
Mean process level[1]	1.5	1.78	-
Mean severity[2]	0.5	0.52	-
Mean process decrement[3]	-4.9505	-5.5088	0.5583
Mean process utility decrement[4]	-0.0495	-0.0551	0.0056
"SF-6D-process"t2	0.5839	0.5708	0.0131
QALY Gained SF-6D	-0.0056	0.0042	-0.0097
QALY Gained "SF-6D-process"	-0.0303	-0.0233	-0.0070
IC/QALY (without process)	-£1,550.90 (Dominated)		
IC/QALY (with process)	-£2,180.09 (Dominated)		

NOTES: SF-6D = short form six dimension; "SF-6D-process" = short form six dimension plus process. QALY = quality adjusted life year; t1 = time 1; t2 = time 2. IC/QALY = incremental cost per quality adjusted life year.

[1]Based on application of the multinomial regression to the trial data.

[2]Based on patient's baseline SF-6D.

[3]The regression co-efficients applied to estimate the process decrement.

[4]Divided by 100 to transform VAS calculations to process utility

Table 4: NHS Costs and "SF-6D-process" QALYs

	Group	Co-efficient	Std.err	P>	95% CI	
NHS Costs	Intervention[1]	4.31	148.60	0.98	-287.27	295.89
	Control[2]	704.71	100.61	0.00	507.40	902.03
SF-6D-process QALYs	Intervention[1]	.0026	.0064	0.687	-.0101	.0153
	Control[2]	.0028	.0043	0.000	.01435	.0312

[1]Mean incremental value (compared to control). [2]Mean value for control

Figure 1: The communication domain of the PEQ [9]

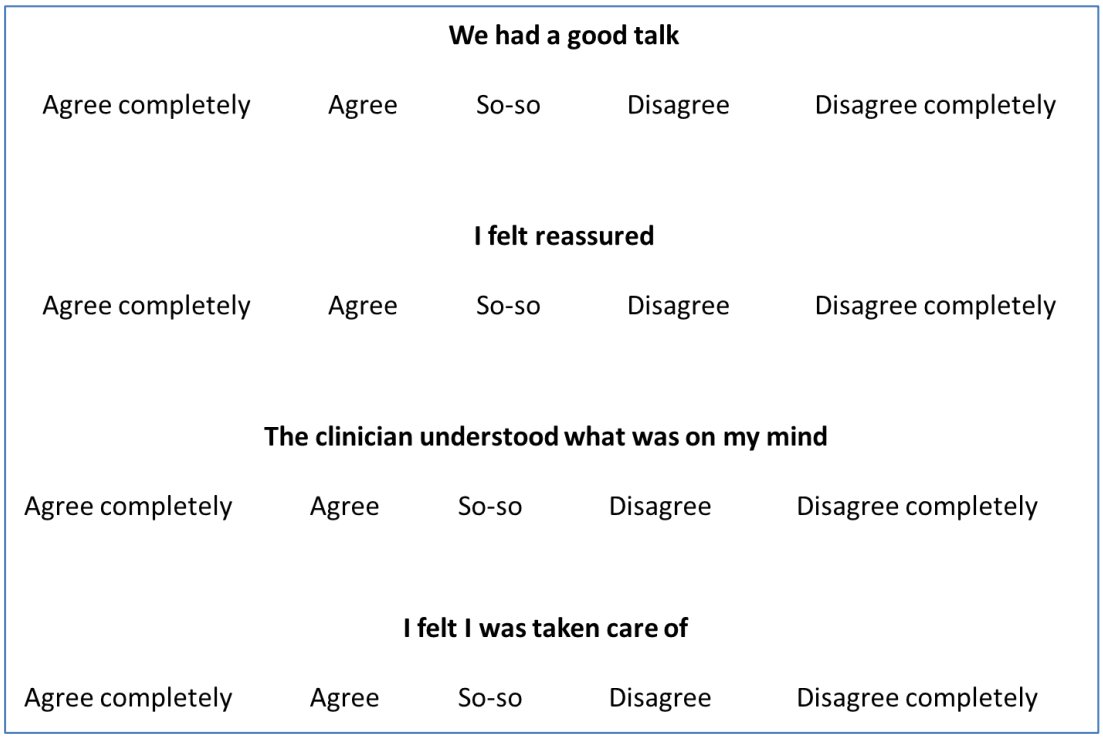
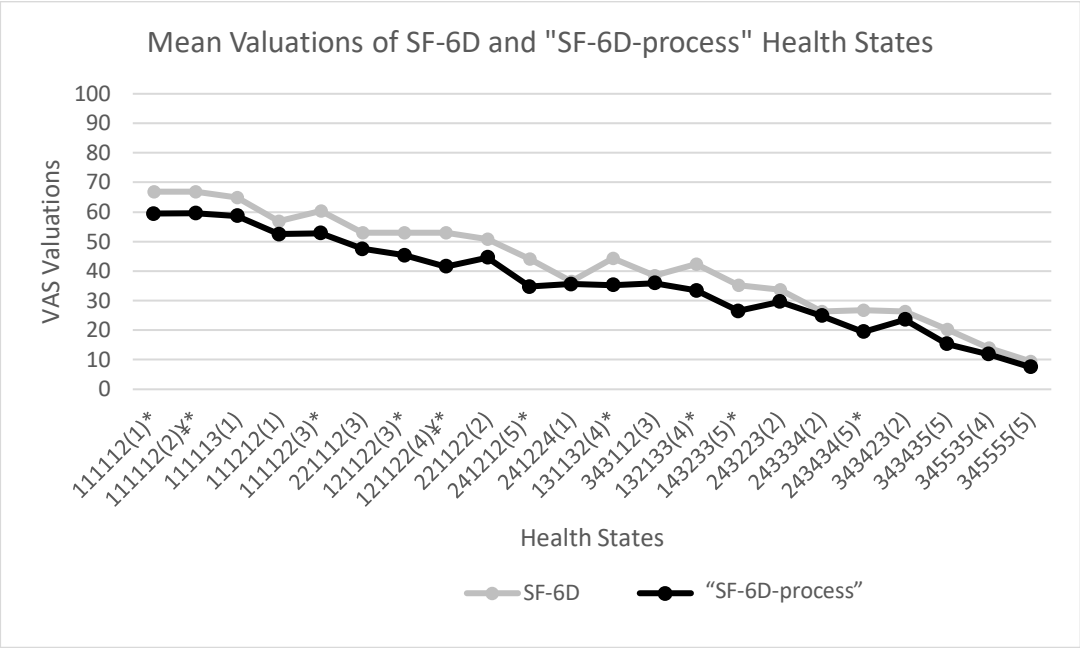


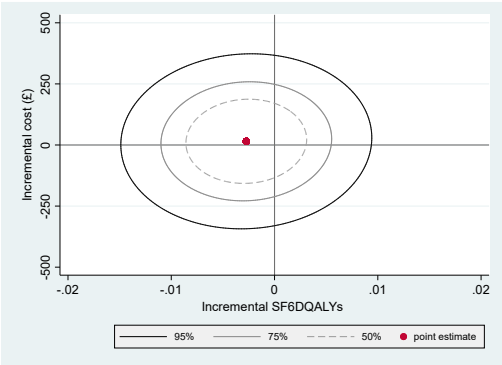
Figure 2: Mean Valuations of SF-6D and "SF-6D-process"(N=260)



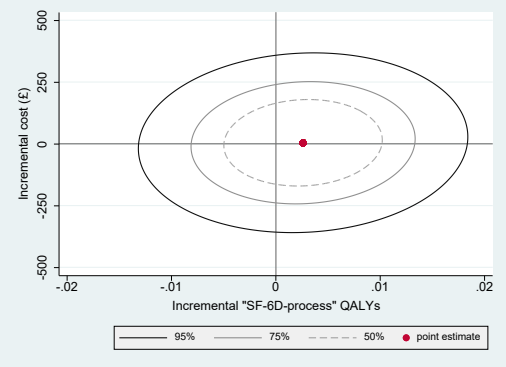
NOTES: *Significant difference($p<.05$) as measured by Mann-Whitney test; ¥ health states differ only in the severity of the process domain; VAS: Visual analogue scale.

Figure 3: Cost-effectiveness planes, and CEACs for SF-6D and SF-6D-process analyses.

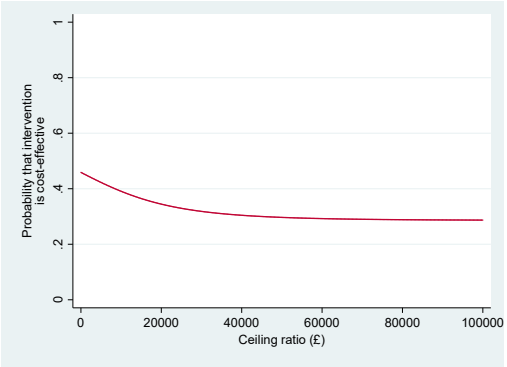
Cost-effectiveness plane (SF-6D)



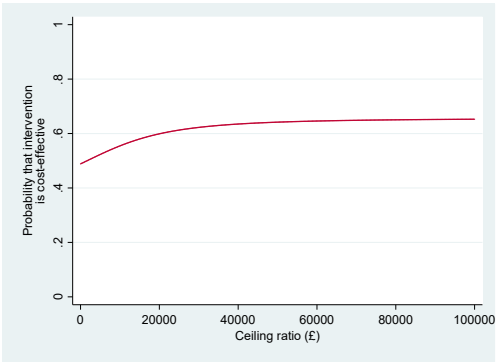
Cost-effectiveness plane (SF-6D-process)



CEAC (SF-6D)



CEAC (SF-6D-process)



NOTES: CEAC = cost effectiveness acceptability curve; SF-6D = short form six dimension; SF-6D-process = short form six dimensions plus process.