Original Article

MDM
Medical Decision Making

Medical Decision Making
2019, Vol. 39(4) 393–404

DOI: 10.1177/0272989X19847983
journals.sagepub.com/home/mdm
SAGE

Exploring the Impact of Adding a Respiratory Dimension to the EQ-5D-5L

Martine Hoogendoorn, Mark Oppe, Melinde R. S. Boland, Lucas M. A. Goossens, Elly A. Stolk, and Maureen P. M. H. Rutten–van Mölken

Objectives. To evaluate the impact of adding a respiratory dimension (a bolt-on dimension) to the EQ-5D-5L health state valuations. **Methods.** Based on extensive regression and principal component analyses, 2 respiratory bolt-on candidates were formulated: R1, limitations in physical activities due to shortness of breath, and R2, breathing problems. Valuation interviews for the selected bolt-ons were performed with a representative sample from the Dutch general public using the standardized interview protocol and software of the EuroQol group. Hybrid models based on the combined time-tradeoff (TTO) and discrete choice experiment (DCE) data were estimated to assess whether the 5 levels of the respiratory bolt-on led to significant changes in utility values. **Results.** For each bolt-on candidate, slightly more than 200 valuation interviews were conducted. Mean TTO values and DCE choice probabilities for health states with a level 4 or 5 for the respiratory dimension were significantly lower compared with the same health states in the Dutch EQ-5D-5L valuation study without the respiratory dimension. Results of hybrid models showed that for the bolt-on "limitations in physical activities," the utility decrements were significant for level 3 (–0.055), level 4 (–0.087), and level 5 (–0.135). For "breathing problems," the utility decrements for the same levels were greater (–0.086, –0.219, and –0.327, respectively). **Conclusions.** The addition of each of the 2 respiratory bolt-ons to the EQ-5D-5L had a significant effect on the valuation of health states with severe levels for the bolt-on. The bolt-on dimension "breathing problems" showed the greatest utility decrements and therefore seems the most appropriate respiratory bolt-on dimension.

Keywords
EQ-5D, respiratory symptoms, bolt-on, valuation study, pilot

Date received: November 1, 2018; accepted: April 5, 2019

The EQ-5D is a generic health-related quality-of-life instrument with a descriptive system that comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each of which has either 3 or 5 levels of severity. Although the EQ-5D is often used in respiratory diseases, such as asthma and chronic obstructive pulmonary disease (COPD),[1–3] important aspects of these diseases are not included in the current 5 dimensions of the EQ-5D. Shortness of breath, coughing, wheezing, and sputum production are not well captured by the EQ-5D domain pain and discomfort, nor is the impact of these symptoms and other problems frequently experienced by patients with respiratory diseases such as fatigue, poor sleep, and impaired

Institute for Medical Technology Assessment, Erasmus University, Rotterdam, Zuid-Holland, the Netherlands (MH, MR-vM); Executive Office, EuroQol Research Foundation, Rotterdam, Zuid-Holland, the Netherlands (MO, ES); and Erasmus School of Health Policy and Management, Erasmus University, Rotterdam, Zuid-Holland, the Netherlands (MB, LG, MR-vM). The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided entirely by a grant from Boehringer Ingelheim GmbH. An in-kind contribution was provided by the EuroQol Group by facilitating the use of the EQ Valuation Technology and providing modeling expertise. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Corresponding Author
Martine Hoogendoorn, Erasmus University Rotterdam, Institute for Medical Technology Assessment, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands (hoogendoorn@imta.eur.nl).

sexual activity. As a result, the responsiveness of the EQ-5D to clinically relevant within-subject change in COPD and asthma over time appears limited,[4,5] unless patients have a serious exacerbation. Several trial-based economic evaluations in COPD reported significant and relevant improvements in clinical and patient-reported outcomes, which were not reflected in significant improvements in terms of quality-adjusted life-years (QALYs).[6–10] In these studies, the 3-level version of the EQ-5D was used.

A possible solution to improve the responsiveness of the EQ-5D in respiratory diseases might be to use the recently developed EQ-5D with 5 instead of 3 severity levels for each dimension. However, a study by Nolan et al.[11] showed that the 8-week change in utility score for COPD patients receiving pulmonary rehabilitation based on the EQ-5D-5L was still weakly correlated with the clinically relevant observed changes in disease-specific outcomes. Ceiling effects (i.e., substantial proportions of patients reporting the best scores on a dimension even among severe patients) were also not improved much by using the EQ-5D-5L instead of the EQ-5D-3L, 7% versus 8.5%, respectively.[12]

Another solution for the low responsiveness of the EQ-5D in respiratory diseases would be the addition of a respiratory dimension to the current descriptive system of the EQ-5D. Such a bolt-on dimension could be relevant not only for respiratory diseases but also for other disease areas in which respiratory symptoms such as shortness of breath play an important role, like cardiovascular and metabolic diseases. Several previous studies identified gaps in the descriptive system of the EQ-5D and explored the impact of bolt-on dimensions such as cognition, concentration, energy, skin issues, vision, sleep, and so forth,[13–22] but the impact of a respiratory bolt-on has not been studied before. Results of the previous bolt-on studies varied, but in general, the impact of a bolt-on on the utility value seems to depend on the severity of the EQ-5D health state in combination with the level of the bolt-on dimension. Adding bolt-on dimensions to the EQ-5D could be a means to improve its responsiveness, but comparability of the results to other diseases/conditions is reduced.

The aim of the current study was to investigate the potential value of a respiratory bolt-on. While the value of a bolt-on ultimately depends on many criteria, in this methodological research, we focused on the development of candidate bolt-on items and the relative contribution of these bolt-ons to the overall utility of health states.

Methods

Development of the Respiratory Bolt-on

To identify which of the respiratory complaints of asthma and COPD patients was the most promising candidate for a respiratory bolt-on, 3 different approaches were used. First, a content review of the EQ-5D and 6 disease-specific health-related quality-of-life measures for asthma and COPD was performed. Disease-specific questionnaires included in this content review were the Asthma Quality of Life Questionnaire (AQLQ),[23] the Asthma Control Questionnaire (ACQ),[24] the St. George's Respiratory Questionnaire (SGRQ),[25] the COPD Assessment Test,[26] the Clinical COPD Questionnaire (CCQ),[27] and the Modified Medical Research Council (mMRC) Dyspnea Scale.[28] Second, a post hoc analysis of data from 3 large clinical studies in asthma and COPD, which included both the EQ-5D-3L and a disease-specific quality-of-life measure, was undertaken.[29–31] Multivariate regression analyses were performed to investigate how much the proportion of explained variance in EQ-5D visual analog scale (VAS) scores (i.e., the R^2) increased when respiratory items of the different disease-specific questionnaires were added to the EQ-5D dimension scores. Linear regression was performed using the EQ-5D VAS score as a dependent variable and 2 dummy variables for each of the 5 different EQ-5D dimensions as independent variables (i.e., the standard model). After that, the same regression model was run plus 1 respiratory item/question. The R^2 of the model, including the respiratory item/question, was compared with the R^2 of the standard model to explore how much additional variance was explained by the respiratory item. The third approach to identify promising candidates for a bolt-on consisted of a principal component analysis (PCA) for the disease-specific health-related quality-of-life questionnaires in combination with the EQ-5D to investigate which items from the questionnaires form different constructs apart from the EQ-5D domains.

After selection of the 2 most promising respiratory candidate items (labeled R1 and R2), the items and levels were phrased such that they were in line with the wording of the EQ-5D-5L items. The final wording of the 2 respiratory candidate items was tested for clarity, comprehensiveness, and relevance in think-aloud sessions with patients with asthma ($n = 5$) and COPD ($n = 6$) and healthy members of the general population ($n = 5$).

Valuation Interviews

To explore the impact of the 2 bolt-on candidate items on health state valuations, a pilot valuation study was performed. Valuation interviews were performed with a representative sample from the Dutch general public aged at least 18 years. Quota sampling was applied to ensure that the recruited panel was representative of the Dutch population in terms of age, gender, and educational level. During individual interviews, participants were asked to value different health states described in terms of the dimensions of the EQ-5D-5L plus the respiratory dimension. A split-sample design was used in which half of the respondents were asked to value the EQ-5D-5L + respiratory item 1 (EQ-5D-5L + R1) and the other half of the respondents were asked to value the EQ-5D-5L + respiratory item 2 (EQ-5D-5L + R2). Respondents were randomly assigned to 1 of the options. During the valuation interviews, both the composite time-tradeoff (TTO) and a discrete choice experiment (DCE) were used to assess the impact of adding a respiratory bolt-on to the EQ-5D-5L on health state valuations.

The pilot valuation study used the EuroQol Valuation Technology (EQ-VT) version 2.0, an updated version of the EQ-VT protocol described by Oppe et al.[32,33] The valuation protocol was embedded in the EQ-VT platform to create the computer-assisted personal interview tool for use during the interviews and was accompanied by standardized interviewer training materials and interviewer scripts. The updated version of the EQ-VT included 3 changes compared with the original version aiming to improve the quality of the data: 1) 3 practice c.q. test states were added to the instructions of the TTO; 2) a feedback module was included after the TTO, to allow respondents to verify their answers and indicate any errors they made; and 3) a quality control (QC) process was included, using the QC tool, a software package purposely developed for use alongside EQ-VT studies.[34] The EQ-VT platform also included the assignment of participants to sets of states from the underlying blocked design, randomization procedures, the iterative process in the TTO, and capturing and time stamping the participants' responses to all tasks.

The interviews were performed by 5 different interviewers who received training on the EQ-VT software and the valuation protocol in advance. In addition, the performance of the interviewers was continuously assessed throughout the entire period of data collection using the QC tool.

Recruited participants were invited to come to 1 of the 2 interview locations for an individual interview. After a brief introduction of the purpose of the research, respondents were asked to complete the descriptive part of the EQ-5D-5L with the respiratory bolt-on and the EQ VAS as well as standard questions regarding age, sex, and experience with illness included in the EQ-VT. The third part consisted of valuation of the selected health states using the composite TTO technique.[35] After extensive explanation of the TTO task, respondents practiced 3 health states, a mild, severe, and moderate state, to get familiar with the range of severity of the health states included in the task. Respondents were asked to consider 2 alternatives: 1) having a particular health state for 10 years or 2) living in perfect health with a shorter duration than 10 years, both followed by immediate death. The duration in perfect health was decreased or increased until the respondent indicated he or she was indifferent between the 2 alternatives. The utility for the health state at the point of indifference between the 2 alternatives was given by $x/10$, where x is 10 minus the maximum tradeoff. When respondents exchanged all 10 years in perfect health (i.e., the health state is worse than death), the "trading time in perfect health" was increased to 20 years. The health state being evaluated still lasted for 10 years but was preceded by a period of 10 years in full health (so-called lead time TTO).[36] Respondents could indicate how much time of the 20 years in full health they were willing to trade so that when valuing very poor states, respondents could trade off more than 10 years and express negative utility values. The end of the TTO task included a feedback module in which the rank order of the valued health states was presented to the respondents. Health states were presented in descending order, starting with the highest valued health state. Health states with the same valuation were presented next to each other. Respondents were asked to check the rank order. If the respondent indicated that a health state was positioned incorrectly in the rank order, that health state was "flagged." Because it was not possible to correct the order of the "flagged" states in the feedback module, the "flagged" states were excluded in the analysis phase of the study. The fourth part was the valuation task using DCE questions. The health states that were valued with DCE questions were different from the states that were valued with the TTO. For the DCE task, respondents had to indicate which health state, state A or state B, they considered the better health state. A number of cognitive debriefing questions were completed both after the TTO as well as the DCE part. For the current study, a separate questionnaire including detailed background questions on education, income, marital status, and experience with lung diseases was also completed. Participants received a 25 Euro compensation for their participation.

Definition of the Reference (= Comparator)

Because for the current valuation study, the same methodology was used as in the Dutch EQ-5D-5L valuation study and respondents were sampled in a similar way, it was planned to compare the results of the current study to the results of the Dutch EQ-5D-5L valuation study without the respiratory dimension.[37] The subset of TTO states and DCE pairs that matched those in the current bolt-on valuation study (see below) was selected from the total data set of the Dutch EQ-5D-5L valuation study and used for comparison with the results of the valuation study, including the respiratory bolt-on. A caveat of this comparison is that the Dutch EQ-5D-5L study was done with version 1.0 of the EQ-VT, which did not include the 3 practice states for the TTO, the feedback module for the TTO, and the QC tool.

Selection of Health States

The health states that were included in the current study were a subset of health states included in the Dutch EQ-5D-5L valuation study (Dutch EQ-VT states). In this way, it could be guaranteed that the descriptive results of the current bolt-on study could be compared directly with those of the Dutch EQ-5D-5L valuation study.[38] For the current study, we aimed for 200 respondents for each of the 2 arms (i.e., R1 and R2). Each arm included the same 2 blocks of 10 states for the TTO task and 4 blocks of 10 pairs for the DCE task. Each individual respondent completed 10 TTO questions and 10 DCE questions. To achieve the recommended 100 observations per state,[39] in total 200 respondents were needed for each bolt-on.

For the selection of the TTO states, 3 criteria were taken into account: utility balance, reasonable level balance, and a fractional addition of the respiratory dimension. Like in the EQ-VT TTO design, each block of 10 questions contained 1 very mild state (with only 1 domain at level 2, all others at level 1) and the worst possible health state (i.e., 55555 + R). These states were selected manually, as was the addition of the sixth domain (the Respiratory dimension) for these states. The 16 remaining TTO states were selected as a subset of the TTO design of the EQ-VT complemented with the respiratory dimension. To increase the utility balance, the Dutch EQ-VT states were first stratified by utility in 7 strata, and mild, moderate, and severe levels of the bolt-ons were added to the states selected in each stratum.

The 40 pairs for the DCE were chosen as a subset of the 196 pair design of the EQ-VT complemented with the respiratory dimension. A Bayesian D-efficient design algorithm was used to generate the design for the DCE, using the Dutch EQ-VT data as priors for the 5 dimensions of the EQ-5D. The prior for the respiratory bolt-ons was based on the average of the first 3 EQ-5D dimensions (i.e., mobility, self-care, and usual activities) but with increased uncertainty surrounding the priors.

Statistical Analysis

The first part of the statistical analyses consisted of the descriptive analyses, including characteristics of respondents, mean TTO-based utility values, and choice probabilities based on the DCE. For each health state, the mean scores were compared between the 2 different bolt-ons and between the bolt-ons and the subset of the Dutch valuation study (= comparator) using t tests to see whether differences were statistically significant. DCE results were presented as the probability that a certain health state was chosen, and chi-square tests were used to test differences between probabilities.

In the second part of the analyses, so-called hybrid models based on the combined TTO and DCE data were estimated. For the 2 bolt-on arms, basic hybrid models were estimated in which a normal distribution was assumed for the TTO data and a conditional logistic distribution for the DCE data, following Ramos-Goñi et al.[38] The hybrid models were estimated using Stata 14 and the hyreg command.[40] For all analyses, a P value less than 0.05 was considered statistically significant.

Results

Development of the Respiratory Bolt-on

Extensive results of the content review and the post hoc regression analyses to determine the most promising respiratory item to include in a bolt-on are presented in detail in an online supplement. In summary, results of the content review showed that all disease-specific questionnaires include questions about the impact of asthma/breathlessness on physical activities (Supplementary Table A1). The EQ-5D includes an item that describes the impact on physical activities in terms of the impact on mobility. Five of the disease-specific questionnaires address asthma or COPD-specific symptoms, while the EQ-5D addresses symptoms more broadly in terms of pain/discomfort. Impact on daily functioning was included in 4 disease-specific questionnaires as well as the EQ-5D. Dimensions that were included in several disease-specific questionnaires, but not in the EQ-5D, were impact on sleep and use of medication. The post

hoc regression analyses showed that the explained variance in the EQ-5D VAS score improved most when items on the impact of asthma or breathlessness on physical activities from the ACQ, AQLQ, CCQ, or mMRC were added to the EQ-5D dimension scores. For the SGRQ, adding items about the impact of the chest condition to the EQ-5D dimension scores resulted in the highest increase in the explained variance (Supplementary Table A2 to A6). Results of the PCA showed that for multiple questionnaires, separate constructs were found on which none of the EQ-5D items loaded. These constructs related either to symptoms or to the impact of breathlessness on activities. Based on the combined results of the post hoc regression analyses and the PCA, we concluded that symptoms and the impact of shortness of breath on physical activities were the 2 most appropriate items/domains for inclusion in the respiratory bolt-on.

Several possible formulations of the new respiratory bolt-on were tested for clarity, comprehensiveness, and relevance, and the final wording of the 2 bolt-ons items were as follows:

R1: Limitations in physical activities due to shortness of breath (e.g., climbing stairs, going for a walk, carrying things, gardening):

- I have no problems with physical activities due to my shortness of breath.
- I have slight problems with physical activities due to my shortness of breath.
- I have moderate problems with physical activities due to my shortness of breath.
- I have severe problems with physical activities due to my shortness of breath.
- I am unable to do physical activities due to my shortness of breath.

R2: Breathing problems (e.g., shortness of breath, wheezing, coughing, sputum):

- I have no breathing problems.
- I have slight breathing problems.
- I have moderate breathing problems.
- I have severe breathing problems.
- I have extreme breathing problems.

With regard to the relevance of the items, patients considered a question on the impact of shortness of breath on their ability to perform physical activities highly relevant and important. However, they thought that the added value of such an item to the existing questions of the EQ-5D was limited. A question about symptoms was regarded to have less importance but was considered to have no overlap with the existing questions of the EQ-5D. Based on these findings, both candidate respiratory bolt-ons were included in the pilot valuation study to determine which of the 2 items had the most impact on the valuation.

Valuation Study of the Respiratory Bolt-on

In total, 430 valuation interviews were completed: 221 for bolt-on R1 (i.e., limitations in physical activities due to shortness of breath) and 209 for bolt-on R2 (i.e., breathing problems). Data were collected in the spring and summer of 2015. In total, 989 completed interviews from the Dutch EQ-5D-5L valuation study were available to use as reference (=comparator). Table 1 shows the characteristics of the respondents. Half of the respondents were male. The mean age was about 47 years, and the mean EQ-VAS score for own health was about 80. About a quarter of the respondents had experienced a severe disease themselves. The characteristics of the respondents valuing R1 and the respondents valuing R2 were not statistically significantly different, although respondents valuing R2 tended to have slightly more health problems themselves and seemed to have somewhat more experience with severe disease. Participants in the current bolt-on study were comparable to the participants in the Dutch EQ-5D-5L study with respect to percentage of males, mean age, and EQ-5D VAS score. During the interviews, a QC report was produced on a weekly basis to check whether all interviewers were using the interview protocol correctly and to standardize the interviews as much as possible between different interviewers. The final report showed that the quality of the interviews can be considered as "good" according to EQ-VT standards.

TTO Values

TTO data of 4 respondents, 2 for each bolt-on, were excluded from the analyses because respondents did not understand the TTO task and the task was prematurely terminated. During the feedback module, 8.4% of the TTO observations were marked as "incorrectly positioned" in the rank order, and these observations were excluded from the analyses. Results including the flagged states were very similar to the ones excluding these flagged states.

Table 2 shows mean TTO values for the bolt-ons and, as a reference, the values for the EQ-5D-5L states in the Dutch EQ-5D-5L valuation study. For very mild health states with no problems on the respiratory dimension (e.g., state 112111), utility values for both bolt-ons were higher compared with the reference values. Health states

Table 1 Baseline Characteristics of the Different Groups of Respondents

	Bolt-on R1	Bolt-on R2	Dutch EQ-5D-5L Valuation Study
Total *N* for analysis	221	209	990
Men, %	49.8	47.8	49.0
Age, mean (SD)	46.8 (15.9)	45.1 (15.2)	47.4 (16.8)
Problems with mobility,[a] %	19.9	23.9	25.6
Problems with self-care,[a] %	2.7	2.4	3.9
Problems with usual activities,[a] %	25.3	29.2	27.6
Pain/discomfort,[a] %	42.1	45.9	49.1
Anxiety/depression,[a] %	14.9	17.7	20.5
EQ VAS score, mean (SD)	81.5 (12.4)	79.2 (15.8)	80.6 (14.7)
Experience with severe disease, % yes			
• Yourself	23.5	28.2	23.6
• Relatives	73.8	80.9	72.5
• Caring for others	38.9	47.4	33.6
Experience with lung disease, % yes			
• Yourself	18.6	24.4	-
• Relatives	50.2	52.2	-
• Caring for others	24.9	31.1	-

[a]Level 2 to 5.

Table 2 Differences between Mean TTO Values

				P Value		
Health State	**Bolt-on R1[a]**	**Bolt-on R2[b]**	**Reference (EQ-5D-5L without Bolt-on)**	**Bolt-on R1 v. Reference**	**Bolt-on R2 v. Reference**	**Bolt-on R1 v. Bolt-on R2**
112111	0.958	0.969	0.925	0.03	0.002	0.34
133131	0.805	0.845	0.759	0.30	0.03	0.24
351431	0.275	0.229	0.330	0.50	0.26	0.63
315251	0.093	0.191	0.311	0.02	0.20	0.29
532441	–0.141	–0.065	0.059	0.02	0.18	0.42
111112	0.954	0.971	–	–	–	0.19
251222	0.643	0.694	0.601	0.55	0.18	0.41
441252	0.118	0.095	0.124	0.95	0.77	0.81
445532	–0.217	–0.205	–0.112	0.26	0.35	0.90
423213	0.679	0.682	0.689	0.85	0.90	0.96
323143	0.370	0.414	0.423	0.48	0.87	0.59
145543	–0.177	–0.130	–0.150	0.76	0.84	0.63
555553	–0.509	–0.580	–0.314	0.002	<0.001	0.40
534124	0.410	0.154	0.466	0.44	0.001	0.004
224344	0.120	0.082	0.305	0.02	0.008	0.66
214444	–0.052	–0.195	0.139	0.03	0.001	0.15
244454	–0.278	–0.423	–0.152	0.15	0.004	0.11
131225	0.678	0.418	0.808	0.01	<0.001	0.001
342325	0.440	0.193	0.603	0.02	<0.001	0.006
555555	–0.581	–0.581	–0.314	<0.001	<0.001	0.99

States are ordered by increasing level of the bolt-on dimension and within that by decreasing mean utility when including the bolt-on R1. Gray cells indicate statistically significant differences.
[a]R1, limitations in physical activities due to shortness of breath.
[b]R2, breathing problems.

with levels 4 and 5 for the respiratory dimension resulted in significantly lower utility values. Utility values did not differ much between the 2 different bolt-ons. Only health states with a level 5 for the respiratory dimension utility values were significantly lower for bolt-on R2, the bolt-on of breathing problems. Although the differences were

not statistically significant at the 95% level, it seemed that bolt-on R2 resulted in a wider range of utilities compared with bolt-on R1. That is, the very mild states showed systematically higher values, while the very severe states showed systematically lower values than bolt-on R1 (apart from the worst state 555555).

Choice Probabilities

Table 3 shows the choice probabilities for different pairs of health states. In general, states with a level 1, no problems, for the respiratory dimension had a higher probability to be chosen as the better state than the same state without respiratory dimension (=comparator). For states with a level 4 or 5 for the respiratory dimension, the probability to be chosen was lower compared with the reference. That applied to both bolt-ons, but more so for R2.

Hybrid Models

Table 4 shows the results of the hybrid models in which both the TTO data and DCE data were combined. The intercept shows the difference between the predicted value for the mildest state defined as 111111 and the theoretical value for this state, equal to 1. It was not statically significant. Results of the hybrid models showed that, in general, for both respiratory bolt-ons and the reference, coefficients decreased as the level increased. For both respiratory bolt-ons R1 and R2, coefficients for the levels 3, 4, and 5 for the respiratory dimensions were statistically significant, although for bolt-on R2 (on breathing problems), the decrements were greater. For bolt-on R2, but not for bolt-on R1, the utility decrements were statistically significantly greater than the decrements of the previous level from level 3 onward. The models for bolt-on R1 and R2 resulted in 2 and 6 inconsistencies, respectively. All inconsistencies except 1 were for levels 2 or 3, and the inconsistent coefficients were not significantly different from zero. For bolt-on R2, the coefficient for usual activities level 5 was nonsignificantly lower than for level 4.

Discussion

The current study explored the potential room for adding a respiratory bolt-on dimension to the EQ-5D-5L. Based on combined results of quantitative analysis and interviews with patients with a respiratory disease, 2 different bolt-on candidates were identified and included in the valuation study: R1, limitations in physical activities due to shortness of breath, and R2, breathing problems.

The largest utility decrements were associated with the second bolt-on candidate. For bolt-on candidate R1, the utility decrements were –0.055 for level 3, –0.087 for level 4, and –0.135 for level 5, whereas for R2, they were –0.086, –0.219, and –0.327, respectively. The decrements compared with the previous level were statistically significant for bolt-on R2 but not for bolt-on R1. However, the overall model for R2 showed more inconsistencies on the other dimensions compared with the model for R1. The inconsistencies mostly resulted from level 2 or 3 not being different from level 1.

Results of the valuation part of this study seem to be in line with previous bolt-on studies. In the current study, mean TTO utility values for very mild states with a level 1 for the respiratory dimension were slightly higher compared with the Dutch EQ-5D-5L study, while states with a level 4 or 5 for the respiratory dimension were lower. Similar results have been found by other researchers. A study by Krabbe et al.[13] that explored the impact of adding a cognitive dimension to the EQ-5D-3L showed that adding a good level of the bolt-on to a worse health state did not result in an increase of the utility values, while a bad level of the bolt-on added to worse health states resulted in an additional reduction of the utility value. Swinburn et al.[17] showed that adding dimensions on skin irritation and self-confidence to the EQ-5D-5L resulted in significant utility decrements for levels 4 and 5 of both bolt-ons. The overall conclusion of a study of Yang et al.[20] exploring the impact of bolt-ons on sleep, hearing problems, and vision problems was that addition of a level 1 for the bolt-on did not change or slightly increased utility values. For severe levels of the bolt-on, utility values were reduced compared with the same health state without bolt-on dimension.[20]

A strength of the current study was that the valuation study was performed using the standardized EQ-VT software and interview protocol. During the interview period, weekly QC reports were made for the first 70% of the data, to ensure that the interviewers followed the protocol and to see whether there were large differences in outcomes between interviewers. In this way, changes to the interview techniques of the different interviewers could be made in an early stage to make the interview approach as uniform as possible.

A limitation of the current study was that we did not create a completely new design for the study (i.e., the set of TTO states and DCE pairs) but selected the TTO states and DCE pairs based on subsets of those included in the EQ-VT complemented with the respiratory dimension. This was done to allow for maximum comparability of obtained responses with data from the Dutch EQ-5D-

Table 3 Choice Probabilities for the Discrete Choice Experiment

Pairs	Bolt-on R1[a]	Bolt-on R2[b]	Reference (EQ-5D-5L without Bolt-on)	*P* Value		
				Bolt-on R1 v. Reference	Bolt-on R2 v. Reference	Bolt-on R1 v. Bolt-on R2
State A v. state B	Probability state A was chosen					
511311 - 353534	96.4	92.5	93.8	0.58	0.82	0.38
121511 - 355432	89.3	84.6	77.6	0.10	0.36	0.47
443231 - 215254	82.1	82.7	63.2	0.04	0.04	0.94
112121 - 221123	81.5	90.7	66.7	0.11	<0.01	0.16
235511 - 431355	73.2	69.2	43.8	<0.01	0.02	0.65
445211 - 411532	71.4	68.0	67.6	0.56	0.83	0.70
311351 - 114442	66.1	74.0	82.9	0.06	0.31	0.38
515521 - 355134	23.6	30.2	15.0	0.30	0.09	0.44
222222 - 255145	100	98.1	100	1.0	0.39	0.30
353212 - 532155	96.4	96.2	90.6	0.27	0.29	0.97
134322 - 132455	92.9	96.0	100	0.15	0.28	0.49
252122 - 324433	92.7	96.1	100	0.12	0.27	0.43
213352 - 445514	74.1	90.7	78.8	0.62	0.12	0.02
225442 - 354523	53.6	63.5	62.5	0.42	0.93	0.30
442342 - 334413	35.7	54.0	38.2	0.81	0.16	0.06
343452 - 513254	22.2	61.1	21.2	0.91	<0.001	<0.001
431413 - 255541	98.1	98.1	95.1	0.40	0.40	0.99
234423 - 254145	63.0	68.5	40.7	0.06	0.02	0.54
343333 - 331422	55.4	57.7	71.9	0.13	0.19	0.81
332253 - 533141	34.5	32.1	39.0	0.65	0.48	0.79
232353 - 111415	30.9	58.5	13.3	0.07	<0.001	0.004
152443 - 442415	30.4	62.0	34.1	0.69	<0.01	0.001
213543 - 413212	12.7	5.7	9.5	0.62	0.47	0.21
342553 - 352211	1.9	3.7	—	—	—	0.56
424214 - 542552	100	90.7	97.1	0.21	0.25	0.02
413124 - 242535	90.9	94.3	93.9	0.57	0.92	0.50
231224 - 124153	89.3	68.0	80.5	0.22	0.18	<0.01
511234 - 434512	69.6	61.5	88.9	0.06	0.01	0.38
541214 - 443223	46.3	24.1	60.7	0.22	0.001	0.02
135534 - 312343	27.8	5.6	36.7	0.40	<0.001	<0.01
344424 - 152141	10.9	9.4	37.0	<0.01	<0.01	0.80
414314 - 242122	8.9	8.0	11.4	0.70	0.59	0.86
225125 - 553133	67.9	38.0	86.0	0.04	<0.001	<0.01
155345 - 434541	67.3	58.5	90.7	<0.01	<0.001	0.35
231245 - 143141	33.9	17.3	68.6	0.001	<0.001	0.049
422555 - 555244	32.1	24.0	33.3	0.91	0.37	0.35
332245 - 421131	12.5	7.7	22.2	0.25	0.07	0.41
455155 - 344334	5.6	3.7	11.8	0.30	0.14	0.65
455425 - 421334	5.4	5.8	2.0	0.38	0.34	0.93
552445 - 535311	1.8	2.0	13.2	0.03	0.04	0.94

Pairs are ordered by increasing level of the bolt-on dimension in state A and within that by decreasing choice probability of the bolt-on R1. Gray cells indicate statistically significant differences.

[a]R1, limitations in physical activities due to shortness of breath.

[b]R2, breathing problems.

5L valuation study (i.e., comparing mean TTO values, DCE choice probabilities, as well as the regression models). However, other designs potentially would have been more efficient with respect to model identification and estimation. Also, the fact that we collected TTO values only for 20 states limited our modeling options, ruling out the possibility of estimating a main effects TTO model including parameters for all levels of the 6 dimensions separately. This sufficed for the current study, in which we could implement the hybrid model instead of a TTO-only model, and with respect to our main aim to see whether addition of the bolt-on had a significant

Table 4 Coefficients Obtained from a Hybrid Model with Intercept

	Bolt-on R1[a]			Bolt-on R2[b]		
	Mean	**Sig.**	**Sig. Inc.**	**Mean**	**Sig.**	**Sig. Inc.**
Intercept	–0.019			–0.025		
Mobility 2	–0.044			–0.087	[c]	[d]
Mobility 3	–0.134	[c]	[d]	–0.120	[c]	
Mobility 4	–0.194	[c]	[d]	–0.213	[c]	[d]
Mobility 5	–0.258	[c]	[d]	–0.290	[c]	[d]
Self-care 2	–0.024			–0.056		
Self-care 3	*–0.012*			*–0.035*		
Self-care 4	–0.112	[c]	[d]	–0.135	[c]	[d]
Self-care 5	–0.141	[c]		–0.135	[c]	
Usual activities 2	*0.013*			*0.014*		
Usual activities 3	–0.009			*0.056*		
Usual activities 4	–0.161	[c]	[d]	–0.177	[c]	[d]
Usual activities 5	–0.187	[c]		*–0.175*	[c]	
Pain/discomfort 2	–0.064	[c]	[d]	–0.085	[c]	[d]
Pain/discomfort 3	–0.105	[c]		*–0.039*		
Pain/discomfort 4	–0.353	[c]	[d]	–0.330	[c]	[d]
Pain/discomfort 5	–0.415	[c]	[d]	–0.434	[c]	[d]
Anxiety/depression 2	–0.081	[c]	[d]	–0.085	[c]	[d]
Anxiety/depression 3	–0.153	[c]	[d]	–0.160	[c]	[d]
Anxiety/depression 4	–0.393	[c]	[d]	–0.325	[c]	[d]
Anxiety/depression 5	–0.493	[c]	[d]	–0.390	[c]	[d]
Respiratory dimension 2	–0.020			*0.018*		
Respiratory dimension 3	–0.055	[c]		–0.086	[c]	[d]
Respiratory dimension 4	–0.087	[c]		–0.219	[c]	[d]
Respiratory dimension 5	–0.135	[c]		–0.327	[c]	[d]

Values in italic indicate a logically inconsistent parameter (e.g., self-care 3 problems result in a higher utility than self-care 2 problems).
[a]R1, limitations in physical activities due to shortness of breath.
[b]R2, breathing problems.
[c]The decrement of the dummy is significantly different from zero at the 95% level.
[d]The dummy is significant compared with the previous level (i.e., the increment of the dummy is different from zero at the 95% level). Level 2 is compared with level 1 (by definition 0 in regression), level 3 with level 2, level 4 with level 3, and level 5 with level 4.
Note: Sig.: Significance of the coefficient; Sig. Inc.: Significance of the increment.

impact on the valuations and whether there were differences between the 2 selected bolt-ons. However, to come up with a value set for the EQ-5D-5L + R, a design created for that specific purpose should be used, which will lead to better properties of the regression model (e.g., fewer inconsistent parameters). Depending on the type of model used (i.e., TTO only or a hybrid), more health states for TTO and/or pairs of health states for the DCE are needed, in addition to a larger sample size.

In both the development and valuation part of the study, we faced a number of issues. During the development phase, we observed a discrepancy between the qualitative and quantitative data about the relative importance of the 2 bolt-on candidates. Patients clearly indicated that an item on the impact of shortness of breath on the ability to perform physical activities was highly relevant and more important than an item on symptoms, even while they recognized that an item on symptoms could have more added value compared with the existing dimensions of the EQ-5D. The valuation study confirmed the latter. This raises the question: what is most important in the development of a bolt-on, relevance or impact on valuation?

One could also argue that respiratory symptoms might already be captured by the EQ-5D pain/discomfort dimension. However, in the PCA, items related to respiratory symptoms did not load on the same construct as the EQ-5D pain/discomfort item, and none of the patients reported seeing an overlap with respiratory symptoms and the pain/discomfort dimension of the EQ-5D.

Furthermore, caution is warranted in comparing TTO values for the bolt-ons with the reference scores from the Dutch EQ-5D-5L valuation study, because the

comparison is made on a between-study basis, and the current data were collected using a later version of the EQ-VT protocol. In the updated EQ-VT protocol, greater attention is given to the QC. It has been suggested that QC reduces interviewer effects and clustering of responses on certain values and increases protocol compliance and the proportion of negative values.[36] Although the impact of using the QC tool was found to be less pronounced in the Netherlands compared with Spain,[36] the difference between protocols might have had an impact on the comparison of the TTO values for states with and without bolt-on.

Although adding a bolt-on/disease-specific dimension to the EQ-5D-5L might be an option to improve the responsiveness of the EQ-5D for certain diseases, the main disadvantage of such an approach is that the comparability of the outcomes of the EQ-5D with other diseases is reduced. Comparability between different diseases is currently considered one of the strengths of the EQ-5D, which raises the question as to whether it is desirable to sacrifice comparability for better responsiveness. Instead of being an alternative to the standard EQ-5D-5L, the EQ-5D-5L plus bolt-on could be used in addition to the standard EQ-5D-5L to calculate so-called "bolt-on" QALYs next to the calculation of the standard QALYs to show the potential change in treatment impact when a disease- or symptom-specific domain is included in the evaluation.

In conclusion, the current pilot valuation study showed that the addition of a respiratory bolt-on about either breathing problems or limitations in physical activities due to shortness of breath to the EQ-5D-5L had a significant effect on the valuation of health states for severe levels of the bolt-on. The bolt-on dimension "breathing problems" showed the greatest impact on health state valuations and therefore seems the most appropriate respiratory bolt-on dimension. However, before this bolt-on can be included in studies to measure the health status of patients, additional valuation interviews including more health states need to be conducted to estimate a consistent valuation set. Further research is also needed to investigate whether the EQ-5D-5L + R is indeed more responsive to change in health status of patients over time.

Authors' Note

This work was presented at the ISPOR Annual European Congress 2015: "Development of an EQ-5D respiratory bolt-on ISPOR Annual European Congress 2016: "Exploring the impact of adding a respiratory dimension to the current EQ-5D descriptive system."

Acknowledgments

We thank the EuroQol group for their permission to use and help with the use of the EQ-VT software, protocol, and QC tool. Arnd Jan Prause is thanked for assistance in adapting the EQ-VT to the current bolt-on study and additional help with technical issues during the interviews. Furthermore, we greatly acknowledge the help of Sander Arons, Rik van Dinteren, and Mathilde Berghout in conducting the interviews.

ORCID iD

Lucas Goossens https://orcid.org/0000-0003-4450-4887

Supplementary Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://journals.sagepub.com/home/mdm.

References

1. Pickard AS, Wilke C, Jung E, Patel S, Stavem K, Lee TA. Use of a preference-based measure of health (EQ-5D) in COPD and asthma. *Respir Med*. 2008;102(4):519–36.
2. Moayeri F, Hsueh YS, Clarke P, Hua X, Dunt D. Health state utility value in chronic obstructive pulmonary disease (COPD); the challenge of heterogeneity: a systematic review and meta-analysis. *COPD*. 2016;13(3):380–98.
3. Tarride JE, Burke N, Bischof M, et al. A review of health utilities across conditions common in paediatric and adult populations. *Health Qual Life Outcomes*. 2010;8:12.
4. Petrillo J, van Nooten F, Jones P, Rutten-van Molken M. Utility estimation in chronic obstructive pulmonary disease: a preference for change? *Pharmacoeconomics*. 2011; 29(11):917–32.
5. Sullivan PW, Ghushchyan VH, Campbell JD, Globe G, Bender B, Magid DJ. Measurement of utility in asthma: evidence indicating that generic instruments may miss clinically important changes. *Qual Life Res*. 2016;25(12):3017–26.
6. Gillespie P, O'Shea E, Casey D, et al.; PRINCE study team. The cost-effectiveness of a structured education pulmonary rehabilitation programme for chronic obstructive pulmonary disease in primary care: the PRINCE cluster randomised trial. *BMJ Open*. 2013;3(11):e003479.
7. Hoogendoorn M, van Wetering CR, Schols AM, Rutten-van Mölken MP. Is INTERdisciplinary COMmunity-based COPD management (INTERCOM) cost-effective? *Eur Respir J*. 2010;35(1):79–87.
8. Brown C, Austin G, McGowan J, Chakravorty I. Is the EuroQol general health status questionnaire sensitive to

the impact of pulmonary rehabilitation in COPD? *Am J Respir Crit Care Med*. 2009;179:A3852.
9. van den Boom G, Rutten-van Mölken MP, Molema J, Tirimanna PR, van Weel C, van Schayck CP. The cost effectiveness of early treatment with fluticasone propionate 250 microg twice a day in subjects with obstructive airway disease: results of the DIMCA program. *Am J Respir Crit Care Med*. 2001;164(11):2057–2066.
10. Briggs AH, Glick HA, Lozano-Ortega G, et al.; Towards a Revolution in COPD Health (TORCH) investigators. Is treatment with ICS and LABA cost-effective for COPD? Multinational economic analysis of the TORCH study. *Eur Respir J*. 2010;35(3):532–9.
11. Nolan CM, Longworth L, Lord J, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. *Thorax*. 2016;71(6): 493–500.
12. Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717–27.
13. Krabbe PF, Stouthard ME, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *J Clin Epidemiol*. 1999;52(4):293–301.
14. Wolfs CA, Dirksen CD, Kessels A, Willems DC, Verhey FR, Severens JL. Performance of the EQ-5D and the EQ-5D + C in elderly patients with cognitive impairments. *Health Qual Life Outcomes*. 2007;5:33.
15. Jelsma J, Maart S. What additional domains, if any, would increase the explanatory power of the EQ-5D? Presented at the 29th Meeting of the EuroQol Group, Rotterdam, the Netherlands, 13–15 September 2010.
16. Swinburn P, Edson-Heredia E, Boye K, Janssen B, Bowman L, Lloyd A. Estimating index values for the EQ-5D-Psoriasis. Presented at the 29th Meeting of the EuroQol Group, Rotterdam, the Netherlands, 13–15 September 2010.
17. Swinburn P, Lloyd A, Boye KS, Edson-Heredia E, Bowman L, Janssen B. Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value Health*. 2013;16(8):1156–62.
18. Yang Y, Brazier J, Rowen D, Tsuchiya A, Young T, Longworth L. Valuing a 'bolt-on' item on vision for the EQ-5D. Presented at the 29th Meeting of the EuroQol Group, Rotterdam, the Netherlands, 13–15 September 2010.
19. Yang Y, Brazier J, Tsuchiya A. Effect of adding a sleep dimension to the EQ-5D descriptive system: a "bolt-on" experiment. *Med Decis Making*. 2014;34(1):42–53.
20. Yang Y, Rowen D, Brazier J, Tsuchiya A, Young T, Longworth L. An exploratory study to test the impact on three "bolt-on" items to the EQ-5D. *Value Health*. 2015;18(1): 52–60.
21. Luo N, Wang X, Ang M, et al. A V vision "bolt-on" item could increase the discriminatory power of the EQ-5D index score. *Value Health*. 2015;18(8):1037–42.
22. Finch AP, Brazier JE, Mukuria C, Bjorner JB. An exploratory study on using principal-component analysis and confirmatory factor analysis to identify bolt-on dimensions: the EQ-5D case study. *Value Health*. 2017;20(10):1362–75.
23. Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the asthma quality of life questionnaire. *Chest*. 1999;115:1265–1270.
24. Nathan RA, Sorkness CA, Kosinski M, et al. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol*. 2004;113(1):59–65.
25. Jones PW, Quirk FH, Baveystock CM. The St. George's Respiratory Questionnaire. *Respir Med*. 1991;85(suppl B): 25–31.
26. Jones PW, Harding G, Berry P, et al. Development and first validation of the COPD Assessment Test. *Eur Respir J*. 2009;34(3):648–54.
27. van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes*. 2003;1:13.
28. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax*. 1999;54(7):581–6.
29. Kerstjens HA, Engel M, Dahl R, et al. Tiotropium in asthma poorly controlled with standard combination therapy. *N Engl J Med*. 2012;367(13):1198–207.
30. Tashkin DP, Celli B, Senn S, et al. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med*. 2008;359(15):1543–54.
31. Kruis AL, Boland MR, Schoonvelde CH, et al. RECODE: design and baseline results of a cluster randomized trial on cost-effectiveness of integrated COPD management in primary care. *BMC Pulm Med*. 2013;13:17.
32. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17(4):445–53.
33. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goñi JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. 2016;34(10):993–1004.
34. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJ, Stolk E. Quality control process for EQ-5D-5L valuation studies. *Value Health*. 2017;20(3):466–73.
35. Janssen BM, Oppe M, Versteegh MM, Stolk EA. Introducing the composite time trade-off: a test of feasibility and face validity. *Eur J Health Econ*. 2013;14(suppl 1):S5–13.
36. Attema AE, Versteegh MM, Oppe M, Brouwer WB, Stolk EA. Lead time TTO: leading to better health state valuations? *Health Econ*. 2013;22(4):376–92.

37. Versteegh M, Vermeulen K, Evers S, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19(4):343–52.
38. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care*. 2017;55(7):e51–8.
39. Oppe M, van Hout B. The 'power' of eliciting EQ-5D-5L values: the experimental design of the EQ-VT. 2017. EuroQol Working Paper Series 17003. Available from: https://euroqol.org/wp-content/uploads/2016/10/EuroQol-Working-Paper-Series-Manuscript-17003-Mark-Oppe.pdf. Accessed 18 February 2019.
40. Ramos-Goñi JM, Craig BM, Oppe M, van Hout BA. Combining continuous and dichotomous responses in a hybrid model. EuroQol Working Paper 16002. Rotterdam, the Netherlands: EuroQol Research Foundation; 2016.