



This is a repository copy of *Development of the CFQ-R-8D: estimating utilities from the cystic fibrosis questionnaire-revised*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/194247/>

Version: Published Version

Article:

Acaster, S., Mukuria, C. orcid.org/0000-0003-4318-1481, Rowen, D. orcid.org/0000-0003-3018-5109 et al. (8 more authors) (2023) Development of the CFQ-R-8D: estimating utilities from the cystic fibrosis questionnaire-revised. *Value in Health*, 26 (4). pp. 567-578. ISSN 1098-3015

<https://doi.org/10.1016/j.jval.2022.12.002>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Patient-Reported Outcomes

Development of the Cystic Fibrosis Questionnaire-Revised-8 Dimensions: Estimating Utilities From the Cystic Fibrosis Questionnaire-Revised



Sarah Acaster, MSc, Clara Mukuria, PhD, Donna Rowen, PhD, John E. Brazier, PhD, Claire E. Wainwright, MD, Bradley S. Quon, MD, Jamie Duckers, MD, Alexandra L. Quittner, PhD,¹ Yiyue Lou, PhD, Patrick R. Sosnay, MD, Lisa J. McGarry, MPH

ABSTRACT

Objectives: Cystic fibrosis (CF) limits survival and negatively affects health-related quality of life (HRQOL). Cost-effectiveness analysis (CEA) may be used to make reimbursement decisions for new CF treatments; nevertheless, generic utility measures used in CEA, such as EQ-5D, are insensitive to meaningful changes in lung function and HRQOL in CF. Here we develop a new, CF disease-specific, preference-based utility measure based on the adolescent/adult version of the Cystic Fibrosis Questionnaire-Revised (CFQ-R), a widely used, CF-specific, patient-reported measure of HRQOL.

Methods: Blinded CFQ-R data from 4 clinical trials (NCT02347657, NCT02392234, NCT01807923, and NCT01807949) were used to identify discriminating items for a classification system using psychometric (eg, factor and Rasch) analyses. Thirty-two health states were selected for a time trade-off (TTO) exercise with a representative sample of the UK general population. TTO utilities were used to estimate a preference-based scoring algorithm by regression analysis (tobit models with robust standard errors clustered on participants with censoring at -1).

Results: A classification system with 8 dimensions (CFQ-R-8 dimensions; physical functioning, vitality, emotion, role functioning, breathing difficulty, cough, abdominal pain, and body image) was generated. TTO was completed by 400 participants (mean age, 47.3 years; 49.8% female). Among the regression models evaluated, the tobit heteroscedastic-ordered model was preferred, with a predicted utility range from 0.236 to 1, no logical inconsistencies, and a mean absolute error of 0.032.

Conclusion: The CFQ-R-8 dimensions is the first disease-specific, preference-based scoring algorithm for CF, enabling estimation of disease-specific utilities for CEA based on the well-validated and widely used CFQ-R.

Keywords: cost utility, cystic fibrosis, Cystic Fibrosis Questionnaire-Revised, patient-reported outcomes, quality of life.

VALUE HEALTH. 2023; 26(4):567–578

Introduction

Cystic fibrosis (CF) affects > 80 000 people worldwide and is the most common life-threatening autosomal recessive disorder in populations of Northern European ancestry, with an overall incidence of 1 in 3500 in European countries.^{1–3} CF has a reduced life expectancy, with median survival estimates of 41 to 57 years across North America and Europe, depending on genotype and sex.^{4,5} CF affects multiple organ systems, but for most people with CF, the largest health impact is progressive respiratory impairment.³ Symptoms commonly include cough, shortness of breath, chest congestion, lack of energy, sinus discharge, and difficulty sleeping.² People with CF also commonly report psychological symptoms.^{6,7}

Health-related quality of life (HRQOL) is a multidimensional concept that includes dimensions related to physical, mental,

and social functioning.^{8–11} Many clinical and demographic characteristics have been associated with HRQOL in CF, including occurrence of pulmonary exacerbations, disease severity, sex, and socioeconomic status.^{8–11} The Cystic Fibrosis Questionnaire-Revised (CFQ-R) is a validated instrument widely used to assess HRQOL in CF studies and clinical trials.^{12,13} The CFQ-R has 12 dimensions: physical functioning, emotional functioning, social functioning/school functioning, body image, eating problems, treatment burden, respiratory symptoms, digestive symptoms, vitality, health perceptions, weight, and role functioning.^{12,13} Three versions of the CFQ-R have been developed: 1 for parents/caregivers to proxy report for children aged 6 to 13 years, 1 that can be interviewer administered for children aged 6 to 11 years or self-completed for children aged 12 or 13 years, and an adolescent/adult version for those aged ≥ 14 years.^{12–14} Although there is some overlap in the items

¹Dr Quittner was an employee of Behavioral Health Systems Research at the time the study was completed, but is currently an employee of The Cystic Fibrosis and Pulmonary Center at Joe DiMaggio Children's Hospital, Hollywood, FL, USA.

and dimensions included in each version, the total number of items differs between each of the 3 versions.¹³ As the version with the broadest target population, this study focuses on the adolescent/adult version.

Cost-effectiveness analyses, a framework used by some health technology appraisal agencies to evaluate novel healthcare interventions, require measures of HRQOL in the form of health state utilities to generate quality-adjusted life-years, which combine the value of HRQOL with the length of life into a single index number. Health state utilities are typically generated from preference-based measures that use preference-elicitation techniques such as time trade-off (TTO), standard gamble, or discrete choice experiment to assign a value, anchored at 0 for dead and 1 for perfect health, to health states described by the underlying classification system.¹⁵ The National Institute for Health and Care Excellence strongly prefers utilities generated using the EQ-5D,^{16,17} a generic preference-based measure of HRQOL, comprising 5 dimensions (mobility, self-care ability, ability to undertake usual activities, pain and discomfort, and anxiety and depression).¹⁸ There are 2 versions of the EQ-5D: one with 3 severity levels in each dimension (EQ-5D-3L) and one with 5 levels (EQ-5D-5L).¹⁹ The EQ-5D-3L lacks sensitivity to meaningful differences in lung function and HRQOL among people with CF, with individuals self-reporting mean utility of 0.923 for mild and 0.870 for severe lung function impairment,²⁰ which are higher than UK (0.856) and US (0.867) population norms.²¹ Although the EQ-5D-5L was developed to increase sensitivity,¹⁹ it has also been shown to lack sensitivity to changes in lung function among people with CF during pulmonary exacerbations.²² Relatedly, a mapping study found that the respiratory dimension of the CFQ-R was not a significant predictor of EQ-5D-3L utility,²³ and utilities estimated from mapping to the EQ-5D-3L showed limited ability to discriminate between groups classified based on lung function in a disease largely characterized by respiratory symptoms.

Given this observed lack of sensitivity of EQ-5D in CF, an alternative approach to estimating utilities is required. Utilities generated from disease-specific measures that are sensitive to change, such as the CFQ-R, have the potential to effectively capture disease-relevant concepts. Nevertheless, given that the CFQ-R is not preference based, it cannot be used directly to generate utilities. Here we derive the first preference-based scoring algorithm to generate utilities from CFQ-R data.

Methods

The study was conducted in 5 stages using methods previously described to estimate a preference-based measure from the Short Form-36 dimension survey^{24,25} and other condition-specific, preference-based measures from existing patient-reported measures of HRQOL.²⁶ The 5 stages were (1) assessing the dimensional structure of the CFQ-R using factor and Rasch analyses, (2) identifying suitable items for the health state classification system using classical psychometric analyses, (3) using clinical and participant input to assess the face validity of the CFQ-R items and dimensions selected in stage 2, (4) valuation of the health states by members of the general public, and (5) developing the scoring algorithm for the classification system using regression modeling. The first 3 stages used existing clinical trial data (described below), whereas the latter 2 stages used primary data collected for this study. Rasch analysis was conducted using RUMM2030 (RUMMlab Pty Ltd, Perth, Australia)²⁷; all other analyses were conducted using Stata 14.2 (StataCorp, College Station, TX).²⁸

Cystic Fibrosis Questionnaire-Revised

The CFQ-R (adult and adolescent version) includes 50 items assessing 12 dimensions scored on 4-point Likert scales, including frequency (always to never), intensity (a great deal to not at all), difficulty (a lot of difficulty to no difficulty), and true-false (very true to very false). Dimensions are scored as the average across all items and rescaled on a 0 to 100 scale where higher scores indicate better HRQOL.

Clinical Trial Data

To facilitate selection of dimensions and items, we used data from 4 randomized, double-blind, placebo-controlled clinical trials designed to evaluate novel CF medications that included the adult and adolescent version of the CFQ-R (EVOLVE [NCT02347657], EXPAND [NCT02392234], TRAFFIC [NCT01807923], and TRANSPORT [NCT01807949]); full descriptions of trial designs have been published.²⁹⁻³¹ In brief, the trials enrolled participants aged ≥ 12 years with CF homozygous for the *F508del-CFTR* mutation or heterozygous for *F508del-CFTR* and a residual function mutation, who were randomized to active treatment versus placebo. The primary outcome was percent predicted forced expiratory volume in 1 second (ppFEV₁), a measure of lung function. Only participants who were administered the CFQ-R adult and adolescent version (ie, those aged ≥ 14 years) were included in this analysis. Three trials had a 24-week intervention period and were used for the main analysis; the EXPAND trial, a crossover trial with 2 intervention periods of 8 weeks, was used to replicate the main-item selection analysis. All analyses were conducted by analysts blinded to treatment assignment to ensure item selection was driven by item performance independent of treatment effect. Data included clinical outcomes and other patient-reported measures, such as ppFEV₁, number of pulmonary exacerbations, and the patient-reported Cystic Fibrosis Respiratory Symptom Diary (CFRSD).³²

Assessment of the Dimensional Structure

The dimensional structure of the CFQ-R was assessed using exploratory factor analysis with the principal-components method and Rasch analysis³³ to identify potential health dimensions and their associated items. Factor analysis can be used to identify dimensional structures, whereas Rasch models allow unidimensional estimates of item location and ability to be made. Results from factor analysis were assessed based on eigenvalues > 1 (including review of scree plots), assessment of contribution of items to each factor and whether they contributed > 1 factor (range 0-1 with higher values indicating greater contribution), and assessment of measurement error based on uniqueness where a value > 0.6 indicated that an item may reflect other information not captured in the dimension (see [Supplemental Methods](https://doi.org/10.1016/j.jval.2022.12.002) found at <https://doi.org/10.1016/j.jval.2022.12.002> for further details). Rasch analysis was undertaken for identified items in each factor to assess whether all items fit based on assessment of the residuals to identify potential divergence and assessment of local dependency (ie, where there was > 1 item measuring the same construct in the factor) (see [Supplemental Methods](https://doi.org/10.1016/j.jval.2022.12.002) found at <https://doi.org/10.1016/j.jval.2022.12.002> for further details). Items excluded at this stage were those that did not contribute to the identified factors or that showed evidence of local dependency or divergence in the Rasch analysis. Items that were optional and those relating to general health were excluded.

Item Selection

To identify items best representing each dimension, a combination of classical psychometric analysis and Rasch analysis was used (see [Supplemental Methods](https://doi.org/10.1016/j.jval.2022.12.002) found at <https://doi.org/10.1016/j.jval.2022.12.002> for further details on item selection methods). Classical psychometric criteria were applied to each item in the CFQ-R using the following metrics: level of missing data, distribution of response across categories (floor and ceiling effects), correlation of item to its own dimension, and responsiveness (standardized response mean [SRM]) to change over time based on improved ppFEV₁.³⁴ Rasch analysis was used to assess the performance of individual items. Items that did not fit the model, did not cover the full range of severity, had disordered response choices, or had differential item functioning were candidates for exclusion from the health state classification system.³⁴ Item wording was required to be suitable for TTO valuation (eg, responses such as “somewhat true” vs “somewhat false” were not concrete enough, and items that combine concepts, such as rating walking function by level of tiredness, were not sufficiently independent).

Assessment of Face Validity of the Classification System

The face validity of the proposed items and dimensions was assessed in interviews with clinicians and individuals with CF to ensure that selected items and dimensions were important, were relevant, and represented dimensions that may change after an effective treatment. Four clinicians practicing in Australia, Canada, the United Kingdom, and the United States and 5 individuals with CF (2 from the United Kingdom, 2 from Australia, and 1 from Canada) participated in the validation process.

Health State Selection and Valuation

Not all possible health states from the classification system could be valued because of the many possible combinations of items; therefore, a subset of health states was valued and used to model the utilities for the complete classification system. An orthogonal array was generated using IBM SPSS statistics version 21, which selected 32 health states for valuation, including the best state (“full health”). The “full health” state was anchored on 1 as the combination of items in the CFQ-R classification system in which no problems were recorded in any dimension, leaving 31 states to be valued. The worst health state was valued by all participants. Each state was valued by multiple respondents; nevertheless, asking respondents to value all states would be excessively burdensome. Therefore, the states were allocated to 4 sets containing mild, moderate, and severe health states, with each respondent valuing 1 set of 8 or 9 health states. To avoid bias, no reference to CF was made in the interview. Once the health states for valuation were selected, cognitive debriefing interviews were conducted with 5 members of the UK general population to evaluate face validity of the states and understanding of the task.

The valuation sample was recruited from a UK general population research panel, aiming to reflect the most recent UK census population demographics.³⁵ The interviews were conducted face to face by 20 trained interviewers across 5 regions of the United Kingdom (Birmingham, Glasgow, Manchester, London, Swansea) in 2018. At the start of the interview, participants were shown the items and an example health state. Afterward, participants completed visual analog scale tasks and TTO tasks, first for 2 practice health states and then for the assigned CFQ-R health state set. If the interviewer felt a participant did not understand or engage with the practice task, they did not continue to complete

the main study health state exercise. For the visual analog scale task, participants were asked to rate the presented health states, plus the best state and “dead,” from 0 (very worst or least preferred) to 100 (very best) to familiarize themselves with the states they were valuing. Participants then completed the TTO, a standardized interview method for valuing health states to generate utility estimates.^{36,37} The method was designed to determine the point at which participants considered 10 years in the target health state to be equivalent (or they were indifferent) to the prospect of x years in full health. Time in full health was varied between high and low values, changing by 6-month intervals, until this point of indifference was reached. If a participant indicated that they believed that being dead was preferable to any time living in a health state, the interviewer switched to a lead-time TTO exercise,³⁸ which asked participants whether they would prefer to live for 10 years in full health followed by 10 years in a health state or to live for x years of full health (where $x < 10$). This lead-time procedure allowed the participant to trade more years of life to determine how much worse than dead they considered the health state to be and to estimate a utility below zero (worse than dead). Participants also completed sociodemographic information, experience of illness (their own or family and friends), and EQ-5D-5L (scored using UK population weights³⁹).

Before conducting the valuation study, the study protocol was reviewed and determined exempt from ethical review requirements by an independent review board in the United States (Western Institutional Review Board); nevertheless, an informed consent was still collected before interview participation.

Development of Scoring Algorithm

Data were reviewed before analysis and flagged for exclusion if responses were considered to reflect either a lack of understanding or engagement or if there were inconsistencies (ie, more severe states were given higher utilities than milder states). Responses were flagged if participants assigned (1) the same value for all states (except where all values are 1 [ie, full health]), (2) the worst state the highest TTO value, and (3) all states values representing worse than dead (< 0). Estimation was undertaken with exclusions and robustness of results assessed the full data set without exclusions.

To produce utilities for every health state defined by the classification system, the utilities were modeled using regression analysis. The standard specification was:

$$y_{ij} = f(\mathbf{X}_{i\lambda}\beta) + \varepsilon_{ij}^s$$

where $i = 1, 2, \dots, n$ represented individual health states and $j = 1, 2, \dots, m$ represented respondents. The dependent variable y_{ij} was TTO disutility (1-TTO value) for health state i valued by respondent j , and $\mathbf{X}_{i\lambda}$ was a vector of dummy explanatory variables for each level λ of dimension δ of the classification system. Level $\lambda = 1$ acted as a baseline for each dimension, and ε_{ij} was the error term. The error term $\varepsilon_{ij} = u_j + e_{ij}$ could be divided into u_j the individual random effect and e_{ij} the random error term for the i th health state valuation of the j th individual. There was no constant term because full health was defined as best CFQ-R health state (ie, 11111111).

Models were estimated using tobit models with robust standard errors clustered on participants with censoring at -1 . This accounts for censored data at -1 ; respondents may want to assign a value below -1 , but this was not possible in the LT-TTO protocol used. Use of robust standard errors clustered on participants

Table 1. Baseline demographics and clinical characteristics of participants in each clinical trial (pooled treatment arms).

Sample characteristics	Trial		
	TRAFFIC and TRANSPORT (N = 1009)*	EVOLVE (N = 455) [†]	EXPAND (N = 235) [‡]
Age, mean (SD), years	26.18 (9.29)	27.75 (9.82)	35.69 (13.79)
Female, n (%)	488 (48.4)	218 (47.9)	130 (55.3)
BMI, mean (SD)	21.50 (2.88)	21.36 (2.78)	24.33 (5.04)
ppFEV ₁ , mean (SD)	59.60 (13.55)	58.50 (14.60)	61.78 (14.34)
CFQ-R dimension score, mean (SD) [§]			
Body image	77.50 (22.22)	76.75 (22.26)	83.59 (17.83)
Digestive	83.49 (16.78)	82.30 (16.52)	84.59 (16.59)
Eating problems	92.31 (15.10)	90.79 (17.25)	93.33 (14.02)
Emotion	82.73 (15.27)	82.52 (16.21)	80.71 (16.26)
Health perceptions	66.88 (20.56)	64.57 (20.83)	64.82 (21.16)
Physical functioning	79.35 (20.21)	77.24 (21.43)	71.33 (23.03)
Respiratory symptoms	68.14 (16.97)	69.37 (16.47)	67.31 (17.50)
Role functioning	84.43 (16.23)	83.92 (16.90)	82.60 (17.16)
Social functioning	73.78 (16.89)	73.27 (16.68)	68.33 (18.55)
Treatment burden	56.55 (19.27)	60.32 (19.49)	63.17 (22.11)
Vitality	64.46 (17.43)	63.40 (18.29)	59.89 (18.41)
Weight	74.54 (33.80)	75.24 (31.63)	87.38 (23.60)
SF-6D/EQ-5D score, mean (SD)	0.93 (0.10)	0.80 (0.12)	0.79 (0.12)
SF-6D/EQ-5D score, median (range)	1 (0.45-1)	0.80 (0.53-1)	0.80 (0.41-1)

BMI indicates body mass index; CFQ-R, Cystic Fibrosis Questionnaire-Revised; ppFEV₁, percent predicted forced expiratory volume in 1 second; SF-6D, Short Form-6 dimension survey.

*Up to 23 participants had missing data for some outcomes.

[†]Up to 3 participants had missing data for some outcomes.

[‡]One participant had missing data for some outcomes.

[§]Scores range from 0 to 100, with higher scores indicating a higher patient-reported quality of life with regard to the specified dimension.

accounted for multiple observations and had an impact on the standard errors but not the coefficients. Random effects (RE) tobit models were estimated to take into account differences at the individual level given that these models appropriately dealt with the structure of the data in which each respondent had multiple observations.²⁴ Mean-level models using tobit with robust standard errors clustered on participants were also estimated, given that they reduced the impact of outliers in health state utilities present in individual-level data. Tobit models that accounted for heteroscedasticity were also estimated given that TTO data typically have larger variance for more severe states. A test for heteroscedasticity in the linear model confirmed this. Inconsistent coefficients for adjacent severity levels of a dimension (eg, moving from “sometimes” to “often” experiencing a health problem) where health deterioration leads to a higher utility were contrary to expectations. To address this issue, models were also estimated that merged inconsistent adjacent severity levels to remove these inconsistencies to ensure that a health deterioration leads to a lower utility score.

Performance of regression models was assessed using the number of significant ($P < .05$) and nonsignificant coefficients, the consistency of the coefficients with the classification system, and mean absolute error (MAE) at the health state level. MAE was generated using the difference between observed and predicted utilities at the health state level, and models with a lower MAE were preferred. The Akaike information criterion and Bayesian information criterion were also examined, with lower values indicating a more preferred model. Predicted values, observed values, and errors by health states were plotted and examined for patterns that would indicate predictive inconsistency across states. The final model was selected through consideration of both logical consistency of coefficients and predictive performance.

Results

Selection of CFQ-R Classification System

Baseline participant demographics and clinical characteristics are presented in Table 1. The 3 participant data sets used for the main analysis (EVOLVE, TRAFFIC, and TRANSPORT) were similar in terms of age, sex distribution, and ppFEV₁, whereas EXPAND had a higher mean age and slightly higher percentage of female participants.

As summarized in Table 2, factor analysis (Appendix Tables 1 and 2 and Appendix Fig. 1 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002>) identified 10 independent factors. Rasch and psychometric analyses plus consideration of item wording reduced these to 8 independent factors and the item pool for consideration from 50 items to 12 items: physical functioning (1 item, phys4), vitality (2 items, vital9 and vital11), emotional functioning (2 items, emot7 and emot12), role functioning (2 items, role36 and role37), respiratory symptoms (2 items, resp41 and resp45), body image (1 item, body26), digestive symptoms (1 item, digest49), and treatment burden (1 item, treat15). Results from the psychometric and Rasch analyses are presented in Appendix Tables 3 to 13 and Appendix Figure 2 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002>.

In summary, the psychometric analysis found that the level of missing data in the data sets was low; therefore, this criterion was not used for item selection. Evidence of ceiling effects was observed for some items, but none had floor effects. Most items (all but 1) had strong correlations ($r > 0.5$) with their related dimension score across all the dimensions. SRMs were generally very small (≤ 0.2), indicating that the items were not responsive. The main exceptions were in the respiratory symptoms

Table 2. Summary of findings from factor, psychometric, and Rasch analysis and assessment of wording.

Item*	Item question	Response options	Dimension	Psychometrics	Rasch	Wording	For consideration
Physical functioning							
phys1	Performing vigorous activities such as running or playing sports	A lot, some, a little, no difficulty	Physical	üü	ü	X	X
phys2	Walking as fast as others	A lot, some, a little, no difficulty	Physical	üü	X	X	X
phys3	Carrying or lifting heavy things such as books, shopping, or school bags	A lot, some, a little, no difficulty	Physical	ü	ü	X	X
phys4 [†]	Climbing 1 flight of stairs	A lot, some, a little, no difficulty	Physical	ü	üü	üü	üü
phys5	Climbing stairs as fast as others	A lot, some, a little, no difficulty	Physical	ü	X	X	X
phys13	To what extent do you have difficulty walking?	A long time without getting tired, a long time but tired, cannot walk a long time as get tired quickly, avoid walking as too tiring	Physical	üü	üü	X	X
phys19	I have trouble recovering after physical effort.	Very true, somewhat true, somewhat false, very false	Cross-loading	üü	X	X	X
phys20	I have to limit vigorous activities such as running or playing sports.	Very true, somewhat true, somewhat false, very false	Physical	üü	üü	X	X
Vitality							
vital6	You felt well.	Always, often sometimes, never	NA [†]	NA	NA	NA	NA
vital9	You felt tired.	Always, often sometimes, never	Vitality	üü	X	üü	üü
vital10	You felt full of energy.	Always, often sometimes, never	Cross-loading	üü	üü	X	X
vital11 [†]	You felt exhausted.	Always, often sometimes, never	Vitality	üü	X	üü	üü
Emotional functioning							
emot7 [†]	You felt worried.	Always, often sometimes, never	Emotion	üü	X	üü	üü
emot8	You felt useless.	Always, often sometimes, never	Emotion	X	üü	üü	X
emot12 [†]	You felt sad.	Always, often sometimes, never	Emotion	üü	X	üü	üü
emot31	I often feel lonely.	Very true, somewhat true, somewhat false, very false	Emotion	üü	ü	X	X
emot33	It is difficult to make plans for the future (eg, going to college, getting married, getting promoted at work).	Very true, somewhat true, somewhat false, very false	Role	X	NA	X	X
Role functioning							
role35	To what extent did you have trouble keeping up with your schoolwork, professional work, or other daily activities during the past 2 weeks?	No trouble, been able but difficult, behind, not able to do at all	Role	ü	üü	X	X
role36 [†]	How often were you absent from school, work, or unable to complete daily activities during the last 2 weeks because of your illness or treatments?	Always, often sometimes, never	Role	X	X	üü	ü

continued on next page

Table 2. Continued

Item*	Item question	Response options	Dimension	Psychometrics	Rasch	Wording	For consideration
role37	How often does CF get in the way of meeting your school, work, or personal goals?	Always, often sometimes, never	Role	üü	üü	X	ü
role38	How often does CF interfere with getting out of the house to run errands such as shopping or going to the bank?	Always, often sometimes, never	Role	üü	X	X	X
Social functioning							
social22	I have to stay at home more than I want to.	Very true, somewhat true, somewhat false, very false	Role	NA	NA	X	X
social23	I feel comfortable discussing my illness with others.	Very true, somewhat true, somewhat false, very false	Social	X	NA	X	X
social27	People are afraid that I may be contagious.	Very true, somewhat true, somewhat false, very false	Social	üü	X	X	X
social28	I get together with my friends a lot.	Very true, somewhat true, somewhat false, very false	Cross-loading	üü	X	X	X
social29	I think my coughing bothers others.	Very true, somewhat true, somewhat false, very false	Social	üü	X	X	X
social30	I feel comfortable going out at night.	Very true, somewhat true, somewhat false, very false	Cross-loading	NA	NA	X	X
Respiratory symptoms							
resp40	Have you been congested?	A great deal, somewhat, a little, not at all	Respiratory	üü	ü	X	X
resp41 [†]	Have you been coughing during the day?	A great deal, somewhat, a little, not at all	Respiratory	üü	ü	üü	üü
resp42	Have you had to cough up mucus?	A great deal, somewhat, a little, not at all	Respiratory	üü	ü	X	X
resp43	Has your mucus been mostly: [optional question from 42]	Clear, clear to yellow, yellowish-green, green with traces of blood, don't know	NA [†]	NA	NA	NA	NA
resp44	Have you been wheezing?	Always, often sometimes, never	Respiratory	X	üü	X	X
resp45 [†]	Have you had trouble breathing?	Always, often sometimes, never	Cross-loading	üü	ü	üü	üü
resp46	Have you woken up during the night because you were coughing?	Always, often sometimes, never	Respiratory	üü	üü	X	X
Eating problems							
eat14	How do you feel about eating?	Thinking about food makes you feel sick, never enjoy, sometimes able to enjoy, always enjoy	Eating	X	NA	X	X
eat21	I have to force myself to eat.	Very true, somewhat true, somewhat false, very false	Eating	X	NA	X	X
eat50	Have you had eating problems?	Always, often sometimes, never	Eating	X	NA	X	X
Body image							
body24	I think I am too thin.	Very true, somewhat true, somewhat false, very false	Body	üü	ü	X	X
body25	I think I look different from others my age.	Very true, somewhat true, somewhat false, very false	Cross-loading	üü	ü	X	X

continued on next page

Table 2. Continued

Item*	Item question	Response options	Dimension	Psychometrics	Rasch	Wording	For consideration
body26 [†]	I feel bad about my physical appearance.	Very true, somewhat true, somewhat false, very false	Cross-loading	üü	üü	X	ü
Digestive symptoms							
digest47	Have you had problems with wind?	Always, often sometimes, never	Digestive	üü	ü	X	X
digest48	Have you had diarrhea?	Always, often sometimes, never	Digestive	X	üü	üü	X
digest49 [‡]	Have you had abdominal pain?	Always, often sometimes, never	Digestive	üü	ü	üü	ü
Weight							
weight39	Have you had trouble gaining weight?	A great deal, somewhat, a little, not at all	Body	X	X	X	X
Treatment burden							
treat15	To what extent do your treatments make your daily life more difficult?	Not at all, a little, moderately, a lot	Treatment	üü	üü	ü	ü
treat16	How much time do you currently spend each day on your treatments?	A lot, some, a little, not at all	Treatment/Role	üü	üü	X	X
treat17	How difficult is it for you to do your treatments (including medications) each day?	Not at all, a little, moderately, very difficult	Treatment	üü	üü	X	X
Health perceptions							
health18	How do you think your health is now?	Excellent, good, fair, poor	NA [‡]	NA	NA	NA	NA
health32	I feel healthy.	Very true, somewhat true, somewhat false, very false	NA [‡]	NA	NA	NA	NA
health34	I lead a normal life.	Very true, somewhat true, somewhat false, very false	NA [‡]	NA	NA	NA	NA

Note. X Psychometrics: ceiling effects (> 70%) and low correlation with dimensions scores.
 X Rasch: poorly fitting items/not ordered in response levels/differential item functioning in age or sex.
 X Wording: not amenable to valuation/too narrow to represent dimension.
 ü Borderline ceiling effects; differential item functioning in age or sex only.
 üü Item included.
 CF indicates cystic fibrosis; CFQ-R, Cystic Fibrosis Questionnaire-Revised; NA, not applicable.
 *Items labeled based on original measure; prefix represents original dimension, and number represents order in the CFQ-R: phys, physical functioning; vital, vitality; emotion, emotional functioning; role, role functioning; social, social functioning; resp, respiratory symptoms; eat, eating problems; body, body image; digest, digestive symptoms; weight, weight; treat, treatment burden; health, health perceptions.
[†]Included in the health state classification.
[‡]Item was not included for consideration of dimension because it is a global item and cannot be included in a classification system.

dimension, where response means were small to moderate in the confirmatory analysis conducted on the EXPAND data for all respiratory symptoms items (> 0.2 to 0.6). Correlations between the items and ppFEV₁, number of pulmonary exacerbations, and body mass index were generally weak and were therefore not used to support item selection. The CFRSD was mainly focused on respiratory symptoms, with 1 item on tiredness. Correlations between the CFRSD and CFQ-R respiratory symptoms items ranged from *r* = 0.37 to 0.63. Correlations between the CFRSD and CFQ-R vitality items ranged from *r* = 0.43 to 0.52.

Of the 12 items identified for consideration, 9 were selected for clinician and participant validation. The role functioning “impact on daily activities” item (role36) was selected over the “goals” item (role37), and vitality “exhaustion” (vital11) was selected over “tiredness” (vital9) because of the more concrete concepts referenced in the former items, despite the role36 item not having the strongest psychometric performance of all role items. Furthermore, the treatment burden item was removed because of

conceptual overlap with the role functioning item. Given that worry (emot7) and sadness (emot12) were both considered relevant and had the same response options, these items were combined to represent the emotion dimension (in line with the EQ-5D anxiety/depression dimension). Finally, the response option wording for the body image item (body26; very true, somewhat true, somewhat false, very false) was judged to be conceptually complex and was therefore dichotomized to a true or false response.

Assessment of Face Validity

The selection of the 9 of 12 items outlined earlier was endorsed by all individuals with CF and clinicians, and all proposed items in the classification outlined in Appendix Table 14 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002> (see example health state in Fig. 1) were considered valid and relevant by CF clinicians and participants interviewed. Because cough and breathing difficulties were judged to be relatively independent,

Figure 1. Example health state. Example of 1 health state; 31 health states were evaluated.

You have **a little** difficulty climbing 1 flight of stairs
 You **sometimes** feel exhausted
 You **sometimes** feel worried or sad
 You are **able** to complete daily activities
 You **sometimes** have trouble breathing
 You cough **a little** during the day
 You **never** have abdominal pain
 You **feel bad** about your physical appearance

these items were treated as independent concepts for the valuation, resulting in 8 dimensions: physical functioning, vitality, emotion, role functioning, breathing difficulty, cough, abdominal pain, and body image.

Valuation Results

After cognitive debriefing interviews (n = 5) to check the understanding and interpretability of the health states, a total of 400 TTO interviews were conducted (see Appendix Figure 3 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002>). Of the total sample, 14 TTO interviews were excluded for valuing all states as identical (but not at full health); 38 respondents were excluded for valuing the worst health state as equivalent to their highest TTO value; and 7 respondents were excluded for valuing all health states as worse than being dead. The main analyses focused on 345 respondents with robustness

analyses using the full sample; data for the practice states were not analyzed.

The analysis sample did not meaningfully differ, based on measured characteristics, from the overall sample and was comparable with the most recent UK population census data in terms of sex, age, and ethnicity (Table 3^{21,40}).

The TTO values for all health states that were directly valued are presented in Appendix Table 15 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002>. Mean and median values did not vary substantially across states, apart from the best and worst health states that were valued. The SDs were large, indicating variation in the individual valuations.

Development of the Scoring Algorithm

Estimates of preference weights based on tobit, RE tobit, and mean-level models using tobit with and without accounting for heteroscedasticity are presented in Table 4. Models with a constant term were also tested, but the constants were not statistically significant and did not improve the model fit statistics. The coefficients were all positive as expected, indicating that less than full health resulted in an increase in disutility. Regression coefficients were logically consistent in most dimensions (ie, disutility values increased as severity increased), but where the levels within a dimension were disordered (eg, levels 3 and 4 for role functioning in the RE tobit model), levels were combined to generate a logically consistent (ordered) model.

All models better predicted mean disutility for each health state at the more severe end than at the milder end, indicating a relationship between error and predictive ability (Fig. 2). MAEs ranged from 0.025 to 0.039. Across the 31 health states, 5 to 12 had MAE > 0.05, and only 1 was > 0.1, indicating small levels of error at the health state level and thus good predictive ability (Table 4).

Table 3. TTO sample demographics and equivalent UK general population census data.

Demographics	Full sample (N = 400)	Analysis set (n = 345)	UK general population census data ⁴⁰
Age, mean (SD), years	47.3 (16.0)	46.5 (16.1)	-
Age distribution, n (%)			
18-40 years	140 (35.3)	129 (37.7)	37.5
41-64 years	173 (43.6)	146 (42.7)	41.6
≥ 65 years	84 (21.2)	67 (19.6)	20.9
Female sex, n (%)	199 (49.8)	173 (50.1)	51.4
White British, n (%)	341 (85.3)	295 (85.5)	87.8
Employment status, n (%)			
Employed or self-employed	271 (67.8)	236 (68.6)	61.9
Unemployed	29 (7.3)	24 (7.0)	4.4
Full-time student	16 (4.0)	16 (4.7)	5.8
Retired	63 (15.8)	52 (15.1)	13.8
Highest education, n (%)			
Secondary school	150 (37.5)	127 (37.1)	28.5
Technical/vocational qualifications*	130 (33.0)	118 (34.2)	NA
Completed university	46 (11.5)	38 (11.1)	27.2
Experience of a serious illness: yes, n (%) [‡]	161 (40)	137 (40)	NA
EQ-5D-5L score, mean (SD)	0.81 (0.25)	0.81 (0.25)	0.86 [‡]

EQ-5D-5L indicates 5-level version of EQ-5D; NA, not applicable; TTO, time trade-off.

*Equivalent to secondary school or university qualification.

[‡]Yourself, a family member, or someone you provided care for.

[‡]UK population norm.²¹

Table 4. Models estimating preference weights based on disutility.

Dimension level	Tobit	Tobit ordered	RE tobit	RE tobit ordered	Mean tobit	Mean tobit ordered	Tobit het	Tobit het ordered
Disutility								
Physical functioning: reference—you have no difficulty climbing 1 flight of stairs								
A little difficulty	0.0351	0.0351	0.0458*	0.0467*	0.0324 [†]	0.0325 [†]	0.0421*	0.0409 [‡]
Some difficulty	0.0644 [†]	0.0645 [†]	0.0579 [†]	0.0588 [†]	0.0684 [†]	0.0687 [†]	0.0603 [†]	0.0593 [†]
A lot of difficulty	0.1062 [†]	0.1063 [†]	0.0962 [†]	0.0960 [†]	0.1035 [†]	0.1033 [†]	0.1044 [†]	0.1036 [†]
Role functioning: reference—you are able to complete daily activities								
Sometimes unable	0.0471*	0.0472*	0.0146	0.0164	0.0568 [†]	0.0570 [†]	0.0485*	0.0482*
Often unable	0.0955 [†]	0.0955 [†]	0.0901 [†]	0.0859 [†]	0.0937 [†]	0.0938 [†]	0.0874 [†]	0.0883 [†]
Always unable	0.1056 [†]	0.1058 [†]	0.0812 [†]	0.0859 [†]	0.0997 [†]	0.0996 [†]	0.1091 [†]	0.1081 [†]
Emotion: reference—you never feel worried or sad								
Sometimes	0.0579 [†]	0.0578 [†]	0.0676 [†]	0.0669 [†]	0.0643 [†]	0.0644 [†]	0.0619 [†]	0.0631 [†]
Often	0.0971 [†]	0.0971 [†]	0.0922 [†]	0.0912 [†]	0.1064 [†]	0.1040 [†]	0.0951 [†]	0.0960 [†]
Always	0.0991 [†]	0.0992 [†]	0.1134 [†]	0.1112 [†]	0.1018 [†]	0.1040 [†]	0.1038 [†]	0.1041 [†]
Vitality: reference—you never feel exhausted								
Sometimes	0.0434 [†]	0.0433 [†]	0.0419*	0.0423*	0.0495 [†]	0.0495 [†]	0.0409*	0.0396*
Often	0.0719 [†]	0.0720 [†]	0.0746 [†]	0.0753 [†]	0.0703 [†]	0.0705 [†]	0.0743 [†]	0.0708 [†]
Always	0.1105 [†]	0.1105 [†]	0.1075 [†]	0.1080 [†]	0.0994 [†]	0.0992 [†]	0.1074 [†]	0.1083 [†]
Breathing difficulty: reference—you never have trouble breathing								
Sometimes	0.0492 [†]	0.0490 [†]	0.0472 [†]	0.0460*	0.0632 [†]	0.0632 [†]	0.0502*	0.0515*
Often	0.0702 [†]	0.0702 [†]	0.0799 [†]	0.0793 [†]	0.0722 [†]	0.0723 [†]	0.0701 [†]	0.0700 [†]
Always	0.1317 [†]	0.1315 [†]	0.1277 [†]	0.1270 [†]	0.1254 [†]	0.1250 [†]	0.1257 [†]	0.1268 [†]
Cough: reference—you do not cough during the day								
A little	0.0204	0.0204	0.0273	0.0272	0.0260 [†]	0.0262 [†]	0.0217	0.0250
Sometimes	0.0368 [†]	0.0369 [†]	0.0280 [†]	0.0283 [†]	0.0451 [†]	0.0454 [†]	0.0407*	0.0426*
A great deal	0.0966 [†]	0.0965 [†]	0.0893 [†]	0.0887 [†]	0.0911 [†]	0.0908 [†]	0.1014 [†]	0.1003 [†]
Abdominal pain: reference—you never have abdominal pain								
Sometimes	0.0598 [†]	0.0551 [†]	0.0613 [†]	0.0594 [†]	0.0624 [†]	0.0574 [†]	0.0664 [†]	0.0586 [†]
Often	0.0506 [†]	0.0551 [†]	0.0582 [†]	0.0594 [†]	0.0523 [†]	0.0574 [†]	0.0525 [†]	0.0586 [†]
Always	0.0843 [†]	0.0843 [†]	0.1089 [†]	0.1082 [†]	0.0685 [†]	0.0682 [†]	0.0850 [†]	0.0847 [†]
Body image: reference—you do not feel bad about your physical appearance								
You do	0.0309*	0.0309*	0.0334 [†]	0.0332 [†]	0.0275 [†]	0.0273 [†]	0.0278*	0.0280*
Not significant	4	4	3	3	0	0	1	1
Inconsistent	1	0	2	0	2	0	1	0
AIC	4049	4047	2743	2739	-12038	-12003	3835	3841
BIC	4188	4180	2887	2872	-11899	-11876	4106	4099
MAE	0.033	0.033	0.039	0.039	0.025	0.025	0.032	0.032
MAE > 0.05	6	6	11	12	6	6	6	5
MAE > 0.1	0	0	0	0	0	0	1	0
Values at:								
22222221	0.687	0.692	0.694	0.695	0.645	0.650	0.668	0.673
44444442	0.235	0.235	0.242	0.242	0.283	0.283	0.235	0.236
33333332	0.483	0.478	0.486	0.489	0.464	0.461	0.492	0.486

AIC indicates Akaike information criterion; BIC, Bayesian information criterion; het, heteroscedasticity; MAE, mean absolute error; RE, random effects.

**P* < .05.

[†]*P* < .01.

[‡]*P* < .1.

The final preferred model, selected through consideration of logical consistency of coefficients, predictive performance, and the ability to reflect variation at the individual level, was the tobit heteroscedastic-ordered model, with a predicted range of health state utilities from 0.236 to 1 for the CFQ-R-8-dimension scoring algorithm (CFQ-R-8D). Based on the worst level, the dimension with the largest disutility was breathing difficulty (0.1268) whereas body image had the smallest disutility (0.028). The Stata code used to generate utility values is provided in the Supplement found at <https://doi.org/10.1016/j.jval.2022.12.002>.

Results based on the full sample had more nonsignificant coefficients (3 compared with 1 in the tobit heteroscedastic model;

Appendix Table 16 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2022.12.002>) although this was reversed for other models (eg, tobit model), and there was also some evidence of slightly more inconsistencies.

Discussion

Here we describe the development of the CFQ-R-8D, a novel preference-based scoring algorithm that allows utilities to be estimated for use in economic evaluations based on patient-reported CFQ-R data. The final classification system has 8

dimensions: physical functioning, vitality, emotion, role functioning, breathing difficulty, cough, abdominal pain, and body image, representing a broad range of the HRQOL impacts in CF included in the CFQ-R adolescent/adult version and endorsed as meaningful and relevant by people with CF and clinicians.

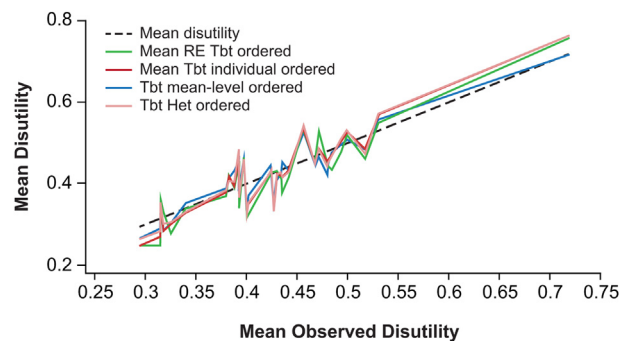
In TTO interviews, most of the states valued were considered to be better than dead, with only 5.5% of the TTO values < 0 and 6% at 0. These results are comparable with the recent EQ-5D-5L England valuation, where 5.1% were valued as worse than dead, and the US valuation where 5.1% were valued at 0.^{41,42} Unlike the EuroQol Valuation Technology protocol used to value the EQ-5D-5L,⁴² not all participants were shown both the worse than dead and better than dead procedure; participants were only shown the worse than dead procedure if their preferences took them there. It is unknown whether this may have affected responses, but it is difficult to reason why knowing how a health state is valued as worse than dead would affect participants' responses in the better than or worse than dead TTO choice. Overall, a few outliers were seen in the TTO data, but at the health state level, mean TTO values did not vary widely, with most values between 0.5 and 0.6. This lack of variability may reflect the mix of severity levels across dimensions in each health state that was valued. It may also reflect the challenge of valuing 8 dimensions using TTO; nevertheless, international protocols for longer and more complex approaches have been successfully implemented (eg, EORTC QLU-C10D).⁴³

The tobit models that were estimated all had coefficients with the expected sign and most were statistically significant at the 10% level, but all had inconsistencies. The tobit heteroscedastic-ordered model was selected because it addressed the problem of heteroscedasticity and only had 1 inconsistency. The values ranged from 0.236 to 1, which is a smaller range than the UK EQ-5D-3L (-0.594 to 1).¹⁸

Before development of the CFQ-R-8D, a study was undertaken to estimate utilities in CF based on a mapping algorithm linking CFQ-R data to the 3-level version of the EQ-5D (EQ-5D-3L).²³ Evidence from this study suggested that the EQ-5D-3L may not be sensitive to meaningful changes in health status in the CF population. The core respiratory dimension of the CFQ-R was not a significant predictor of EQ-5D-3L utility and thus not included in the mapping algorithm. Perhaps not surprisingly, utilities estimated from the algorithm showed limited ability to discriminate between groups classified based on lung function. The CFQ-R-8D reflects a broad range of CF-specific health dimensions included in the CFQ-R, which is a well-validated and widely used measure to evaluate treatment benefit in CF. Although the CFQ-R-8D uses a subset of the CFQ-R items in the scoring algorithm, as is generally necessary for scoring algorithms, this specificity does not suggest that the full CFQ-R should not be administered. Capturing the full impact of CF on HRQOL provides important evidence outside economic evaluation.

Limitations of the current work should be highlighted. Four trial data sets were used to select items for the classification system, and although the use of multiple data sets and larger combined sample size was advantageous in this context, all 4 samples included clinical trial participants for whom severity of CF may have been different from that in a typical CF population because of study inclusion criteria. In addition, most items demonstrated only small SRMs, the notable exceptions being the respiratory items, and thus may not reflect their performance in other CF populations. As such, the items selected here ideally should be validated in another setting, such as a registry or observational study. The TTO sample was drawn from the UK

Figure 2. Mean observed and predicted disutility by health state.



Het indicates heteroscedastic; RE, random effects; Tbt, tobit.

population to reflect UK societal values as recommended by agencies, including the National Institute for Health and Care Excellence.⁴⁴ To use this classification system in another country, it may be desirable to repeat the TTO valuation and algorithm estimation with a local population; nevertheless, UK valuations for utility measures may be acceptable where local weights are unavailable.⁴⁵

Interviewers were trained, but there were no specific built-in interviewer quality checks during the data collection process, and the number of interviewers and variability in number of interviews conducted may have affected data quality. Notably, 5 interviewers did not record values < 0, but because they equally did not record many values at zero (1 interviewer recorded 3 values at zero, whereas the others recorded none), this was not considered in the modeling. Therefore, data quality issues were addressed using exclusions where poor understanding in the TTO was indicated by the TTO responses. A substantial number of TTO participants (n = 55; 13.8%) were excluded from the analysis because of low data quality, suggesting some participants had difficulty with the valuation exercises that was not identified in the cognitive interviews or by the interviewer. In addition, some respondents may have found it difficult to assess 8 dimensions that may have led to heuristics such as focusing on some of the dimensions and not others. Nevertheless, the full sample and analysis TTO set were comparable, based on measured demographic characteristics, and the analysis set was comparable with the most recent UK population census data. Given the number of health states required for the TTO valuation, each participant valued only 1 of the 4 blocks of health states, which may have added systematic variability; nevertheless, having participants value all states would have risked data quality, and the approach taken here reflects that which was used in previous valuations (eg, EQ-5D). Finally, evidence is required about how the new measure compares to measures such as EQ-5D in the context of economic evaluation and quality-adjusted life-year estimation.

Conclusions

The CFQ-R-8D allows direct estimation of CF-specific utilities from the CFQ-R, a well-validated measure that is used widely in CF clinical trials and clinical practice, thus enabling utilities for use in cost-effectiveness analyses to be generated from any existing or future CFQ-R data set. The ability to adequately capture the HRQOL

in this population using a metric suitable for economic evaluation is essential to demonstrating the potential benefit and value of new CF treatments. An evaluation of the psychometric performance of the CFQ-R-8D compared with the generic EQ-5D-3L and Short Form-6 Dimension survey is ongoing.⁴⁶

Supplemental Material

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.jval.2022.12.002>.

Article and Author Information

Accepted for Publication: December 1, 2022

Published Online: January 16, 2023

doi: <https://doi.org/10.1016/j.jval.2022.12.002>

Author Affiliations: Acaster Lloyd Consulting Ltd, London, England, UK (Acaster); School of Health and Related Research, The University of Sheffield, Regent Court, Sheffield, England, UK (Mukuria, Rowen, Brazier); Child Health Research Centre, University of Queensland, Brisbane, QLD, Australia (Wainwright); Centre for Heart Lung Innovation, St. Paul's Hospital, and Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada (Quon); Department of Respiratory Medicine, Cardiff and Vale University Health Board, NHS Wales, Cardiff, Wales, UK (Duckers); Research, Behavioral Health Systems Research, Miami, FL, USA (Quittner); Biostatistics, Vertex Pharmaceuticals Incorporated, Boston, MA, USA (Lou); Clinical Development, Vertex Pharmaceuticals Incorporated, Boston, MA, USA (Sosnay); Health Economics and Outcomes Research, Vertex Pharmaceuticals Incorporated, Boston, MA, USA (McGarry).

Correspondence: Sarah Acaster, MSc, Acaster Lloyd Consulting Ltd, 8th Floor, Lacon House, 84 Theobald's Rd, London, England WC1X 8NL, United Kingdom. Email: sarah.acaster@acasterlloyd.com

Author Contributions: *Concept and design:* Acaster, Mukuria, Rowen, Brazier, Duckers, Quittner, Lou, McGarry
Acquisition of data: Rowen, Wainwright, Quon
Analysis and interpretation of data: Acaster, Mukuria, Rowen, Brazier, Quon, Duckers, Quittner, Lou, Sosnay, McGarry
Drafting of the manuscript: Acaster, Mukuria, Rowen, Wainwright, Quon, Duckers, Quittner, Sosnay
Critical revision of the paper for important intellectual content: Mukuria, Rowen, Brazier, Wainwright, Quon, Duckers, Quittner, Sosnay, McGarry
Statistical analysis: Acaster, Mukuria, Lou
Supervision: McGarry
Obtaining funding: Brazier, McGarry
Administrative, technical, or logistic support: Quittner
Other (provision of CFQ-R): Quittner

Conflict of Interest Disclosures: All authors reported receiving nonfinancial support (assistance with manuscript preparation) from ArticulateScience LLC, which received funding from Vertex Pharmaceuticals Incorporated. Ms Acaster reported employment with Acaster Lloyd Consulting Ltd, which received payment from Vertex Pharmaceuticals Incorporated for their contribution to the design, management, and undertaking of all aspects of the study reported in this manuscript; in addition, Dr Acaster is an associate editor for *Value in Health* and had no role in the peer-review process of this article. Dr Mukuria reported receiving grants from Vertex Pharmaceuticals and membership in the EuroQoL Group Association. Dr Rowen reported receiving grants from Vertex Pharmaceuticals Incorporated. Dr Wainwright reported advisory board membership for Vertex Pharmaceuticals Incorporated; consulting fees from Vertex Pharmaceuticals Incorporated; editor duties for *Respirology* (associate editor) and *Thorax* (deputy editor); honoraria from *BMJ*, DKBmed, Gilead, In Vivo Academy, Novartis, *Thorax*, University of Miami, and Vertex Pharmaceuticals Incorporated; research grant from Novo Nordisk; being a study investigator for Boehringer Ingelheim and Vertex Pharmaceuticals Incorporated; and travel expenses from Vertex

Pharmaceuticals Incorporated. Dr Quon reported receiving grants and personal fees from Vertex Pharmaceuticals Incorporated. Dr Duckers reported receiving personal fees from Vertex Pharmaceuticals Incorporated, Chiesi Pharmaceuticals, and Insmed. Dr Quittner reported research grants from the NIH, CF Foundation, and American Cochlear Implant Alliance and consulting fees from Insmed and Vertex Pharmaceuticals Incorporated. Drs Lou, Sosnay, and McGarry reported employment with Vertex Pharmaceuticals Incorporated and may own stock or stock options in that company. Dr Quittner was an employee of Behavioral Health Systems Research at the time the study was completed but is currently an employee of the Cystic Fibrosis and Pulmonary Center at Joe DiMaggio Children's Hospital (Hollywood, FL). No other disclosures were reported.

Funding/Support: This study was supported by Vertex Pharmaceuticals Incorporated.

Role of the Funder/Sponsor: Vertex Pharmaceuticals Incorporated contributed to the design of the study and the review of the manuscript. Data collection and analysis were undertaken by Acaster Lloyd Consulting Ltd, which received funding from Vertex Pharmaceuticals.

REFERENCES

1. Southern KW, Munck A, Pollitt R, et al. A survey of newborn screening for cystic fibrosis in Europe. *J Cyst Fibros*. 2007;6(1):57–65.
2. Bell SC, Mall MA, Gutierrez H, et al. The future of cystic fibrosis care: a global perspective. *Lancet Respir Med*. 2020;8(1):65–124.
3. Elborn JS. Cystic fibrosis. *Lancet*. 2016;388(10059):2519–2531.
4. 2018 Patient Registry Annual Report. ECFS. https://www.ecfs.eu/sites/default/files/general-content-files/working-groups/ecfs-patient-registry/ECF_SPR_Report_2018_v1.4.pdf. Accessed September 3, 2021.
5. Keogh RH, Szczesniak R, Taylor-Robinson D, Bilton D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: a longitudinal study using UK patient registry data. *J Cyst Fibros*. 2018;17(2):218–227.
6. Sawicki GS, Sellers DE, Robinson WM. Self-reported physical and psychological symptom burden in adults with cystic fibrosis. *J Pain Symptom Manag*. 2008;35(4):372–380.
7. Quittner AL, Goldbeck L, Abbott J, et al. Prevalence of depression and anxiety in patients with cystic fibrosis and parent caregivers: results of the International Depression Epidemiological Study across nine countries. *Thorax*. 2014;69(12):1090–1097.
8. Bradley JM, Blume SW, Balp MM, Honeybourne D, Elborn JS. Quality of life and healthcare utilisation in cystic fibrosis: a multicentre study. *Eur Respir J*. 2013;41(3):571–577.
9. Gee L, Abbott J, Conway SP, Etherington C, Webb AK. Quality of life in cystic fibrosis: the impact of gender, general health perceptions and disease severity. *J Cyst Fibros*. 2003;2(4):206–213.
10. Quittner AL, Schechter MS, Rasouliyan L, Haselkorn T, Pasta DJ, Wagener JS. Impact of socioeconomic status, race, and ethnicity on quality of life in patients with cystic fibrosis in the United States. *Chest*. 2010;137(3):642–650.
11. Habib AR, Manji J, Wilcox PG, Javer AR, Buxton JA, Quon BS. A systematic review of factors associated with health-related quality of life in adolescents and adults with cystic fibrosis. *Ann Am Thorac Soc*. 2015;12(3):420–428.
12. Quittner AL, Buu A, Messer MA, Modi AC, Watrous M. Development and validation of the Cystic Fibrosis Questionnaire in the United States: a health-related quality-of-life measure for cystic fibrosis. *Chest*. 2005;128(4):2347–2354.
13. Quittner AL, Sawicki GS, McMullen A, et al. Psychometric evaluation of the Cystic Fibrosis Questionnaire-Revised in a national sample. *Qual Life Res*. 2012;21(7):1267–1278.
14. Quittner AL, Modi AC, Wainwright C, Otto K, Kirihaara J, Montgomery AB. Determination of the minimal clinically important difference scores for the Cystic Fibrosis Questionnaire-Revised respiratory symptom scale in two populations of patients with cystic fibrosis and chronic *Pseudomonas aeruginosa* airway infection. *Chest*. 2009;135(6):1610–1618.
15. Brazier J, Ratcliffe J, Tsuchiya A, Solomon J. *Measuring and Valuing Health for Economic Evaluation*. 2nd ed. Oxford, United Kingdom: Oxford University Press; 2017.
16. Burgel PR, Munck A, Durieu I, et al. Real-life safety and effectiveness of lumacaftor-ivacaftor in patients with cystic fibrosis. *Am J Respir Crit Care Med*. 2020;201(2):188–197.
17. Brooks R, Boye K, Slaap B. EQ-5D: a plea for accurate nomenclature. *J Patient Rep Outcomes*. 2020;4(1):52.
18. Dolan P. Modeling valuations for EuroQoL health states. *Med Care*. 1997;35(11):1095–1108.
19. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–1736.

20. Solem CT, Vera-Llonch M, Liu S, Botteman M, Castiglione B. Impact of pulmonary exacerbations and lung function on generic health-related quality of life in patients with cystic fibrosis. *Health Qual Life Outcomes*. 2016;14:63.
21. Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, eds. *Self-Reported Population Health: an International Perspective Based on EQ-5D*. Dordrecht, NL: Springer; 2014.
22. Gold L, Patrick D, Hansen R, Beckett V, Goss C, Kessler L. Correspondence between symptoms and preference-based health status measures in the STOP study. *J Cyst Fibros*. 2019;18(2):251–264.
23. Acaster S, Pinder B, Mukuria C, Copans A. Mapping the EQ-5D index from the Cystic Fibrosis Questionnaire-Revised using multiple modelling approaches. *Health Qual Life Outcomes*. 2015;13:33.
24. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21(2):271–292.
25. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol*. 1998;51(11):1115–1128.
26. Brazier JE, Rowen D, Mavranzeouli I, et al. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technol Assess*. 2012;16(32):1–114.
27. Andrich D, Lyne A, Sheridan B, Luo G. (1997–2012) *RUMM2030: Rasch Unidimensional Models for Measurement*. Perth, Australia: RUMMLab Pty Ltd; 2010.
28. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP; 2015.
29. Taylor-Cousar JL, Munck A, McKone EF, et al. Tezacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del. *N Engl J Med*. 2017;377(21):2013–2023.
30. Wainwright CE, Elborn JS, Ramsey BW, et al. Lumacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del CFTR. *N Engl J Med*. 2015;373(3):220–231.
31. Rowe SM, Daines C, Ringshausen FC, et al. Tezacaftor-ivacaftor in residual-function heterozygotes with cystic fibrosis. *N Engl J Med*. 2017;377(21):2024–2035.
32. Goss CH, Edwards TC, Ramsey BW, Aitken ML, Patrick DL. Patient-reported respiratory symptoms in cystic fibrosis. *J Cyst Fibros*. 2009;8(4):245–252.
33. Boone WJ. Rasch analysis for instrument development: why, when, and how? *CBE Life Sci Educ*. 2016;15(4):rm4.
34. Young T, Yang Y, Brazier JE, Tsuchiya A, Coyne K. The first stage of developing preference-based measures: constructing a health-state classification using Rasch analysis. *Qual Life Res*. 2009;18(2):253–265.
35. Selvadurai H, Kennedy B, Bailey B, et al. Substrate utilisation during exercise in children with CF and healthy controls - the impact of CFTR modulator therapy. *Am J Respir Crit Care Med*. 2020;201:A4681.
36. Drummond MF, McGuire A. *Economic Evaluation in Health Care: Merging Theory With Practice*. Oxford, United Kingdom: Oxford University Press; 2001.
37. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ*. 1986;5(1):1–30.
38. Devlin NJ, Tsuchiya A, Buckingham K, Tilling C. A uniform time trade off method for states better and worse than dead: feasibility study of the 'lead time' approach. *Health Econ*. 2011;20(3):348–361.
39. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708–715.
40. Volkova N, Moy K, Evans J, et al. Disease progression in patients with cystic fibrosis treated with ivacaftor: data from national US and UK registries. *J Cyst Fibros*. 2020;19(1):68–79.
41. Mulhern B, Feng Y, Shah K, et al. Comparing the UK EQ-5D-3L and English EQ-5D-5L value sets. *Pharmacoecon*. 2018;36(6):699–713.
42. Pickard AS, Law EH, Jiang R, et al. United States valuation of EQ-5D-5L health states using an international protocol. *Value Health*. 2019;22(8):931–941.
43. Norman R, Mercieca-Bebber R, Rowen D, et al. U.K. utility weights for the EORTC QLU-C10D. *Health Econ*. 2019;28(12):1385–1401.
44. Konstan MW, McKone EF, Moss RB, et al. Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-CFTR mutation (PROGRESS): a phase 3, extension study. *Lancet Respir Med*. 2017;5(2):107–118.
45. Brazier J, Ara R, Azzabi I, et al. Identification, review, and use of health state utilities in cost-effectiveness models: an ISPOR Good Practices for Outcomes Research Task Force report. *Value Health*. 2019;22(3):267–275.
46. Mukuria C, Rowen D, Brazier JE, et al. PR1 comparison of the psychometric performance of a new condition-specific preference-based measure derived from the CFQ-R (CFQ-R-8D) to EQ-5D-3L and SF-6D to evaluate health-related quality-of-life (HRQOL) in people with cystic fibrosis (CF). *Value Health*. 2020;23:S689.