

Allergy / Early View

ORIGINAL ARTICLE |  Full Access

Mapping the Food Allergy Quality of Life Questionnaire Parent Form onto the SF-6Dv2

Yan Watts, Élise Dufresne, Kathryn Samaan, François Graham, Roxane Labrosse, Louis Paradis, Anne Des Roches, Thomas G. Poder, Philippe Bégin 

First published: 25 November 2021

<https://doi-org.bibliosan.clas.cineca.it/10.1111/all.15190>

Abstract

Background

The Food Allergy Quality of Life Questionnaire Parent Form (FAQLQ-PF) is the most widely used quality of life questionnaire in food allergy. The objective of this study was to develop a mapping algorithm to convert FAQLQ-PF scores into health state utilities.

Methods

The Short-Form Six-Dimensions version 2 (SF-6Dv2) and FAQLQ-PF questionnaires were collected from an academic center oral immunotherapy referral cohort. Utility estimates were derived from the SF-6Dv2 using the food allergy preference set. Candidate mapping algorithm models were developed using seven regression methods starting from either the total average score, the average scores of each of the three domains or the individual item scores of FAQLQ-PF. The process was repeated twice, including only section A, common to all age groups, or including all age-applicable sections of the FAQLQ-PF. The mean absolute error (MAE) and root mean squared error (RMSE) were used to select the best fitting model. An independent cohort from a previous national online survey was used for external validation.

Results

In the index cohort, 1000 of 1257 respondents had completed both questionnaires. The lowest MAE (0.0791) and RMSE (0.1020) were recorded when entering individual item scores in a categorical regression model. The model including only FAQLQ-PF section A

was found to be most consistent when tested in the external validation cohort ($n = 248$) (MAE of 0.0898).

Conclusion

The FAQLQ-PF was mapped onto SF-6Dv2 utilities with good predictive accuracy in two independent cohorts. This will enable calculation of health utility for cost-effectiveness analyses in food allergy.

Abbreviations

CATREG

Categorical regression with optimal scaling

CCC

concordance coefficient correlation

CLAD

Censored least absolute deviations

FAQLQ-PF

Food Allergy Quality of Life Questionnaire Parent Form

HRQoL

health-related quality of life

IQR

Inter-quartile range

MAE

mean absolute error

MFP

Multivariable fractional polynomials

MM

Robust MM-estimator

OLS

Ordinary least squares

RMSE

root mean square error

SD

Standard deviation

SF-6Dv2

Short-Form Six-Dimension version 2

1 INTRODUCTION

The global burden of food allergy is largely driven by the associated decrease in health-related quality of life (HRQoL).¹ While the associated risk of death and direct cost to the healthcare system are relatively low compared to other diseases, patients and families with food allergy can nevertheless be severely impacted by social and diet limitations and food-related anxiety.² Our ability to advocate for these patients and justify investment in new

therapies is inextricably linked to our capacity to measure those intangible costs.

The interpretation of HRQoL data from clinical studies can be challenging. Most studies rely on the food allergy quality of life questionnaires (FAQLQ), which were designed and validated to capture the most relevant aspects of food allergy quality of life.³⁻⁵ The questionnaires consist in series of questions selected according to age and answered on 7-point Likert scales covering three domains (anxiety, limitations, emotional impact). A common practice has been to use the general average score or the average scores per domain to assess change with interventions. However, treating ordinal data as numerical data creates significant methodological problems affecting the validity of presented estimates.⁶ For one, the value given to the interval between each possible answer in an ordinal scale cannot be assumed to be the same. Second, the weight of each question in the questionnaire is expected to differ based on their relative importance and impact (e.g., limited choice of restaurants vs. emotional impact). Finally, even if these basic limitations were accounted for, the value of the estimate relative to other comparable disease states would remain unknown, precluding the proper valuation of interventions.

In health economics, HRQoL is measured using health utility values and the validity of cost-utility analysis is directly dependent on their proper estimation.^{7, 8} Utility expresses the value given to life lived in a given health state from 0 to 1, using perfect health (utility of 1) and death (utility of 0) as anchor values (Table 1). Individual patient utility values are typically derived from generic health-related questionnaires that have been previously validated in the population of interest using direct elicitation methods (e.g., standard gamble, time trade-off, discrete choice experiment).^{7, 9}

TABLE 1. Glossary

Health-related utility: expresses the value given to life lived in a given health state from -1 to 1, using perfect health (utility of 1) and death (utility of 0) as anchor values.
Heteroscedasticity: refers to data with unequal variability across a set of predictor variables. In other words, it means that the precision of the algorithm varies depending on the observed score.
Homoscedasticity: refers to data with equal variability across a set of predictor variables. In other words, it means that the precision of the algorithm remains the same across the full spectrum of possible scores.
Quality-adjusted life-year: The quality-life-adjusted-years combines the health-related utility value of a given health state and the time spend in that health state into a single score, generating a simple, universal measure of health outcome facilitating the comparison of interventions across disciplines, following an utilitarian approach.
Questionnaire mapping: refers to the development of an algorithm to predict health-related utility values by using data from another measure of health outcomes.
Utility set: refers to a set of values attributed to each possible answer to a health-related quality-of-life questionnaire,

allowing to estimate the utility score (*indirect elicitation*). Utility set values are themselves calculated through the direct elicitation of patient preferences, which can vary depending on the population.

The only questionnaire for which a utility preference set has been specifically elicited in the food allergic population is the Short-Form Six-Dimension version 2 (SF-6Dv2), which defines health states according to 6 dimensions: physical functioning, role limitation, pain, emotional impact, social limitations, and vitality.¹⁰ Its main limitation is that it is a generic questionnaire, and thus inevitably lacks the precision and sensitivity of a disease-specific questionnaire to capture subtle and relevant disease-specific changes in food allergy HRQoL.¹¹⁻¹³ The easiest solution to this common problem is to map disease-specific questionnaires to utility scores obtained from a population-validated generic questionnaires to create a conversion algorithm that can estimate utility directly from disease-specific questionnaires with both high sensitivity and high accuracy. This is the preferred approach of the National Institute of Clinical Excellence (NICE) and the International Society of Pharmacoeconomics and Outcome Research (ISPOR), which both stress the importance of validating the conversion algorithm in a second independent cohort to ensure its validity.^{14, 15}

The goal of this study was to map the FAQLQ Parent Form (FAQLQ-PF) questionnaire to the health utility values derived from the SF-6Dv2 using two large cohorts of food allergic patients.

2 METHODS

2.1 Study design

The cohort used to generate the conversion algorithm was previously described elsewhere.¹⁶ It consisted of all patients aged 0–12 years old referred to the oral immunotherapy clinic at the CHU Sainte-Justine in Montreal, Quebec, between September 2017 and March 2019. As part of the referral process, parents were required to fill the FAQLQ-PF and SF-6Dv2 questionnaires in French or English. To be eligible for referral, the food allergy needed to have been confirmed by the referring allergist by allergy testing or food challenge within the previous year.

A second validation cohort consisted of parents of food allergic patients aged 0–12 years old who participated to the online pan-Canadian survey that included the original direct elicitation experiments establishing the SF-6Dv2 food allergy utility set. The survey was available in French or English between August 2018 and August 2019.¹⁰ The invitations were distributed through patient associations and in allergy clinics. Before starting the survey, participants had to confirm that their food allergy had been previously diagnosed by a physician using allergy tests or a food challenge. This study was approved by the institutions

ethics committee and informed consent was obtained from participants.

Finally, to assess the tools' ability to examine change in utility over time, we used an opportunity sample of patients who had undergone oral immunotherapy at the CHU Sainte-Justine and who had been repeated the FAQLQ and SF-6Dv2 after treatment. This third cohort included patients treated for various types and numbers of food allergies, with or without the use of biologic therapy.

2.2 Instruments

The SF-6Dv2 questionnaire uses a standardized health state descriptive system consisting of 6 dimensions rated on a 5-point ordinal scale (except for pain, which uses a 6-point scale). The questionnaire was administered using the unitary approach, in which parents are instructed to choose the answer that best describes their health and that of their child, thus treating the parent-child dyad as a whole, as previously described.¹⁰ The FAQLQ-PF consists in 30 questions, each answered on a 7-point Likert scale. It is intended for parents of children between 0 and 12 years of age. The first 14 questions (Section A) are answered by all participants; questions 15–26 are answered only by patients aged 4 years and above (section B); questions 27–30 are answered only by patients aged 8 years and above (section C).

2.3 Data filtering

In both questionnaires, responses on the ordinal scales were coded from 1 and up, with 1 indicating a lack of impact for the corresponding item. Questionnaires with more than 20% missing values were excluded from analyses. In the SF-6Dv2, items left unanswered or marked by parents as “not applicable” were attributed a value of 1. To allow inference from available answers, missing values in the FAQLQ-PF were imputed with the median for that question in the corresponding cohort. Decimals were rounded up to the ceiling value (e.g., 4.5 → 5).

2.4 Statistical analyses

Patients' characteristics were summarized using median and interquartile range (IQR) for continuous variables and frequency (percentage) for categorical variables. A Shapiro-Wilk test was used to verify the normality of continuous responses. The relation between SF-6Dv2 utility scores and FAQLQ-PF total average scores was assessed with the Spearman correlation coefficient. The discriminative ability of the SF-6Dv2 utility score was assessed using *t*-tests comparing SF-6Dv2-derived utility score according to clinical background.

To test the various ways FAQLQ-PF is reported in the literature, the regression models were developed starting from either the total average score (Model 1), the average scores for each domain (Model 2) or by treating each question item as independent variable (Model 3), treated as continuous values, as well as from individual items recognized as ordinal values

(Model 4) (Tables [S1 and S2](#)). Each model was tested twice: first by including only the questions of section A, common to all age groups (Version A); and then by using a distinct model for each age group, including all age-appropriate questions (Version B). The age groups were also entered in all regression models as a dummy variable to account for the confounding effect.

Seven different regression methods were used:

- **Ordinary least squares (OLS)** minimizes the sum of squares between predicted and observed utility scores. This model assumes homoscedasticity, errors are normally distributed with mean zero and constant variance, which is not always possible.¹⁷
- **Tobit** is used in the presence of outcome variables censored at lower or upper endpoints. This model assumes that there is a latent utility score that could be greater than 1, instead of modeling a maximum score of 1.¹⁷
- **Censored least absolute deviations (CLAD)** is similar to Tobit, but uses the median of the latent utility instead of its mean, which is more robust to distributional assumption and heteroscedasticity.¹⁸ In this model, the sum of absolute deviations is minimized instead of the sum of squares of the differences between observed and predicted utility scores.¹⁹
- **Robust MM-estimator (MM)** creates an M-estimator, which minimizes a function of the residuals, and an S-estimator, which minimizes a measure of dispersion of the residuals, which increases robustness in presence of outliers or heteroscedasticity.²⁰
- **Generalized linear model (GLM)** allows non-normal distributions to be specified for a certain a priori distribution. Family functions (Gaussian, inverse Gaussian, and gamma) and link functions (identity, log) were investigated with AIC and BIC criteria.²¹
- **Multivariable fractional polynomials (MFP)** is used when the dependent and the independent variables have a non-linear relationship. We only used the GLM method for the MFP model.²²
- **Categorical regression with optimal scaling (CATREG)** approach regresses the responses of categorical predictor variables by using optimal scaling to assign numerical quantifications to the categories $Q_{i1}, \dots, Q_{i7}$ of each variable (Q_i). The optimal quantification for each scaled variable is achieved with alternating least squares. One coefficient (β) is estimated for each categorical variable and numerical quantifications are estimated for each of these categories $\{Q_{i1}, \dots, Q_{i7}\}$. Ordinal variables were discretized by ranking and the dependent variable was discretized by the multiplying method.²³

Models with continuous variables (1–3) included squared variables to model a non-linear relationship for regression techniques (OLS, Tobit, CLAD, MM).^{21, 24} Squared variables were omitted from GLM and MFP since they already attempt to fit a non-linear model. All models used backward elimination method ($p < .05$) to choose the significant predictors to be included in the final models, except for CLAD and MM, which were performed with the predictors selected by OLS.²⁰ Multi collinearity was assessed in the selected model with variance inflation factor. Sensitivity analyses included a backward methods (a) using a cut-off value of $p < .10$, and (b) excluding items that were inversely related to disutility.

2.5 Assessing goodness of fit

Goodness of fit was assessed in both the index and the external validation cohort using four criteria: (a) the mean absolute error ($MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$) of the absolute difference between predicted and observed utility values; (b) the root mean square error ($RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$) of the squared differences between the predicted and observed utilities; (c) the proportion of predicted values deviating from observed values by absolute error < 0.05 ; and (d) the concordance coefficient correlation (CCC) which is a measure of intra-class correlation.²⁵ The best model was determined based on an appreciation of goodness of fit and consistency between the two cohorts.

2.6 Additional validation

Utility estimates derived from the best model were further compared to utility estimates calculated using a previously published FAQLQ-PF mapping algorithm based on the AQOL-6D generic questionnaire in cohorts #1 and #2.²⁶ This comparison included only children aged 4 and above, since that algorithm includes questions not answered by parents of younger children.

The difference in overall mean utility detected by the SF-6Dv2 and the best FAQLQ-PF mapping algorithm model in cohort #3 were compared with a *t*-test. The difference in distribution was compared using the Kolmogorov–Smirnov test.

All statistical analyses were conducted using STATA software version 17.0 except for CATREG Optimal Scaling regression, which was performed on SPSS software version 26.0.

3 RESULTS

3.1 Study sample characteristics

Of the 1257 respondents in the index cohort (#1), a total of 257 were excluded: 151 were aged between 13 and 17 years; 93 did not answer the SF-6Dv2; and 13 had more than 20% missing values in the FAQLQ-PF. The median age of the remaining 1000 patients was 5.6 (IQR 3.2–8.6) with 34.6% female (Table 2). They had a median of three food allergies (IQR 1–5), the most frequent being peanuts, nuts, and eggs (76.1%, 45.3%, and 35.3%, respectively). The mean utility score calculated from the SF-6Dv2 was 0.8189 (SD 0.1329), ranging from 0.1620 to 1.0000. The Shapiro-Wilk test rejected the hypothesis that the SF-6Dv2 utilities were normally distributed ($p < .001$) (Figure S1). The SF-6Dv2 utility score was associated with the number of food allergies, allergy to ubiquitous foods, a previous history of anaphylaxis, concomitant asthma, or atopic dermatitis diagnoses (Table S3) as well as younger patient age (Table S4). The average FAQLQ-PF total score on a scale of 0–6 was 2.95 (SD 1.33) or 3.01 (SD 1.35), when considering only section A or all age-applicable sections, respectively.

TABLE 2. Characteristics of patients in study cohorts

Characteristics	Cohort #1 (N = 1000)	Cohort #2 (N = 248)	Cohort #3 (N = 141)
Age, Median (IQR)	5.6 (3.2–8.6)	5.0 (2.5–9.0)	4.7 (2.9–7.5)
Female, n (%)	346 (34.6)	-	49 (34.8)
Number of food allergies, Median (IQR)	3.0 (1.0–5.0)	3.0 (1.0–5.0)	2.0 (1.0–4.0)
Accidental reactions since birth, Median (IQR)	3.0 (1.0–6.0)	3.0 (2.0–5.0)	2.0 (1.0–5.0)
Food allergy, n (%)			
Peanuts	761 (76.1)	181 (73.0)	92 (65.2)
Nuts	453 (45.3)	116 (46.8)	56 (39.7)
Eggs	353 (35.3)	100 (40.3)	51 (36.2)
Milk	194 (19.4)	61 (24.6)	36 (25.5)
Sesame	196 (19.6)	47 (19.0)	28 (19.9)
Shellfish/mollusc	82 (8.2)	15 (6.0)	7 (5.0)
Legumes	82 (8.2)	8 (3.2)	4 (2.8)
Fish	76 (7.6)	25 (10.1)	10 (7.1)
Mustard	66 (6.6)	19 (7.7)	11 (7.8)

Abbreviations: FAQLQ-PF, Food Allergy Quality of Life Questionnaire—Parent Form; IQR, Interquartile range; SD, Standard deviation.

^a Prior to food oral immunotherapy.

Of the 295 parents participating to the external validation cohort (online survey), 43 were aged between 13 and 17 years, and 4 did not complete the SF-6Dv2. The median age of the remaining 248 participants was 5 years (IQR 2.5–9.0). The rate of ubiquitous food allergies (53.2%) was higher than for cohort #1 (34.6%). Patients had a median of three food allergies (IQR 1.0–5.0), the most frequent being peanuts, nuts, and eggs (73.0%, 46.8%, and 40.3%, respectively). The mean utility score calculated from the SF-6Dv2 was 0.7956 (SD 0.1374), ranging from 0.1890 to 1.0000. The average FAQLQ-PF total score on a scale of 0–6 were the same when calculated based on Section A (2.90, SD 1.386) and all age-applicable sections (2.90, SD 1.393).

3.2 Regression models

A negative moderate correlation was observed between the FAQLQ-PF total scores and SF-6Dv2 scores (Spearman correlation coefficient (r) = $-.4127$, $p < .001$ based on section A and (r) = $-.3986$, $p < .001$ based on all applicable sections). The items that were selected in the various regressions are presented in the supplemental material (Table [S5](#)). Question 6 in the FAQLQ-PF (physical distress) was the only item selected by all regression models and was systematically given the highest β value.

Table [3](#) summarizes the key goodness-of-fit statistics of the regression models. The best performing model was 4B with CATREG, which had the lowest MAE (0.0791) and RMSE (0.1020) as well as the highest CCC (0.5820). CATREG also performed best when including only models based on FAQLQ-PF section A, with a MAE of 0.0811, a RMSE of 0.1066, and a CCC of 0.5254 (Figure [S2](#)). The sensitivity analysis using a cut-off of $p < .1$ in the backward elimination method selected the same items and variance inflation factors were all below 3 (Table [S6](#)).

TABLE 3. Goodness of fit statistics of regression models in the index cohort ($N = 1000$)

Models	Version A—Section A items							Version B – Age-applicable		
	Mean (SD)	Min	Max	MAE	RMSE	A.D. <0.05	CCC	Mean (SD)	Min	Max
Observed	0.8189 (0.1329)	0.1620	1.0000							
Model 1										
OLS/GLM	0.8189 (0.0705)	0.5923	0.9652	0.0873	0.1126	35.2	0.4388	0.8189 (0.0689)	0.5790	0.94
TOBIT	0.8237 (0.0746)	0.5836	0.9795	0.0871	0.1128	35.8	0.4522	0.8236 (0.0725)	0.5767	0.95
CLAD	0.8367 (0.0619)	0.6374	0.9454	0.0867	0.1146	38.1	0.3972	0.8323 (0.0660)	0.5730	0.93
MM	0.8343 (0.0626)	0.6326	0.9622	0.0865	0.1140	38.8	0.4043	0.8354 (0.0620)	0.5976	0.93
MFP	0.8189 (0.0710)	0.5610	0.9381	0.0872	0.1123	36.0	0.4439	0.8189 (0.0691)	0.5681	0.91
Model 2										

Note

Gaussian family function with identity link function was selected for all the models that had a GLM regression, which is equivalent to the OLS regression method. Therefore, results for OLS and GLM are presented together. Dependent variable: SF-6Dv2. Bold font indicates the model for best each for (MAE, RMSE, A.D. < 0.05, CCC).

Abbreviations: A.D. < 0.05%, proportion of predicted utilities whose absolute values deviate from the mean of the observed utility values by less than 0.05; CATREG, Categorical regression with optimal scaling; CCC, Concordance correlation coefficient; CLAD, Censored least absolute deviations; GLM, General linear model; MAE, Mean Absolute Error; MFP, Multivariable fractional polynomials; MM, Robust MM-estimator; OLS, Ordinary least square; RMSE, Root Square Mean Error; TOBIT, Censored regression.

When the two CATREG models were compared in the external validation cohort, Model 4A was found to be more consistent than Model 4B, (MAE of 0.0898 and 0.0956, respectively, Table 4). The β values and level quantifications for each item in model 4A are presented in Table S7. This equation uses the FAQLQ-PF item results and patient age group, to calculate a Z-score relative to the distribution of SF-6Dv2-derived utility values in the index cohort ($Z_{FAQLQ-PF}$). It therefore needs to be converted back to utility value using the mean and standard deviation of utility in the index cohort (0.8189, SD 0.1329) as described below, where Q_{ij} is the quantification at level j of item i and age_j is the quantification for the patient's age group.

- $$Z_{FAQLQ-PF} = \beta_1 Q_{1j} + \beta_3 Q_{3j} + \beta_4 Q_{4j} + \beta_6 Q_{6j} + \beta_7 Q_{7j} + \beta_8 Q_{8j} + \beta_{10} Q_{10j} + \beta_{12} Q_{12j} + \beta_{14} Q_{14j} + \beta_{age} age_j$$
- Utility = $Z_{FAQLQ-PF} * 0.1329 + 0.8189$

TABLE 4. Goodness of fit statistics from validation analysis ($N = 248$)

Model	A – Section A items only			B – Age-applicable sections		
	Mean utility	MAE	RMSE	Mean utility	MAE	RMSE
Observed	0.7956			0.7956		
Model 1						
OLS/GLM	0.8185	0.0944	0.1256	0.8185	0.0940	0.1257
TOBIT	0.8234	0.0952	0.1269	0.8234	0.0946	0.1270
CLAD	0.8308	0.0955	0.1292	0.8308	0.0955	0.1304
MM	0.8337	0.0952	0.1290	0.8337	0.0945	0.1291
MFP	0.8178	0.0953	0.1266	0.8178	0.0946	0.1261

Model	A – Section A items only			B – Age-applicable sections		
	Mean utility	MAE	RMSE	Mean utility	MAE	RMSE
Model 2						
OLS/GLM	0.8219	0.0931	0.1246	0.8219	0.0925	0.1245
TOBIT	0.8270	0.0940	0.1259	0.8270	0.0931	0.1258
CLAD	0.8342	0.0946	0.1280	0.8342	0.0926	0.1261
MM	0.8368	0.0944	0.1288	0.8368	0.0937	0.1287

Note

Dependent variable: SF-6Dv2. Bold indicates the best model for each statistic (MAE, RMSE, A.D. < 0.05, CCC).

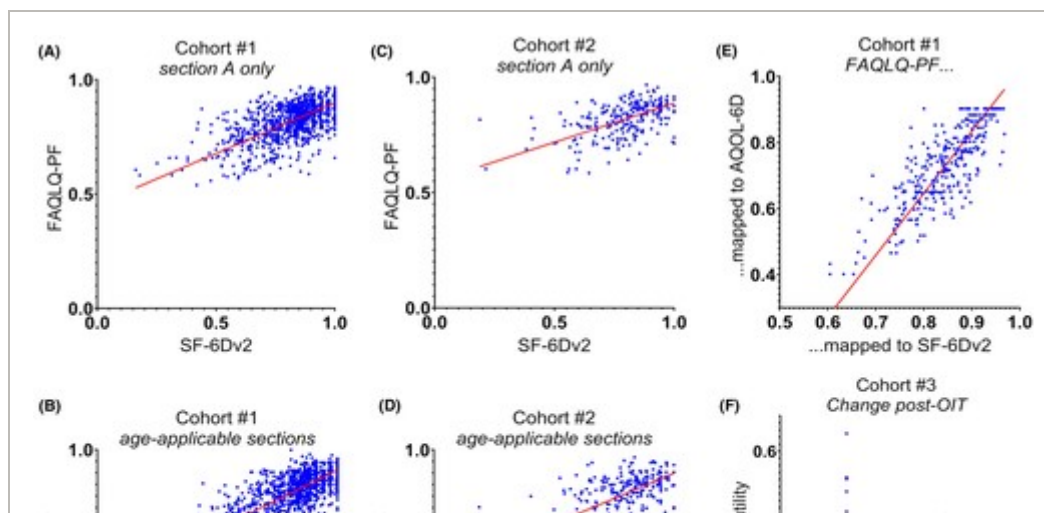
Abbreviations: A.D. < 0.05, proportion of predicted utilities whose absolute values deviate from the mean of the observed utility values by less than 0.05; CATREG, Categorical regression with optimal scaling; CLAD, Censored least absolute deviations; GLM, General linear model; MAE, Mean Absolute Error; MFP, Multivariable fractional polynomials; MM, Robust MM-estimator; OLS, Ordinary least square; RMSE, Root Square Mean Error; TOBIT, Censored regression.

This equation is further simplified in [Figure 2](#).

The sensitivity analysis using backward elimination excluding items inversely related to disutility led to the same model but without question 3. It had higher MAE both in the index and validation cohorts.

3.3 Additional validation

Utility estimates calculated using the final model were also found to correlate with those calculated with the prior mapping algorithm for adolescents based on the AQOL-6D²⁶ in both cohorts ($r = .8048$ and $r = .7624$) ([Figure 1E](#)).



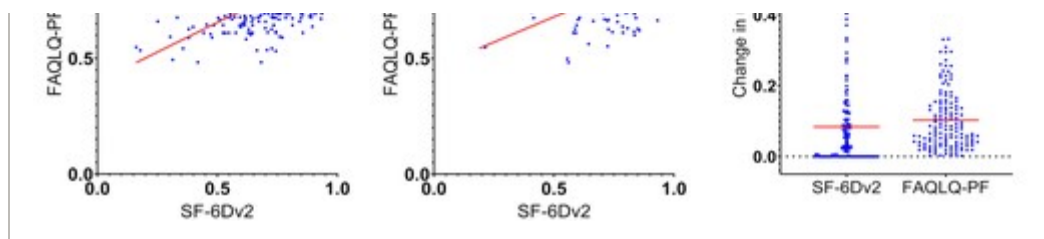


FIGURE 1

[Open in figure viewer](#) | [PowerPoint](#)

Model performances: These scatter plots compare the observed SF-6Dv2-derived utilities and the utilities predicted using the model generated with categorical regression with optimal scaling based on only section A or all age-applicable sections of the FAQLQ-PF in the index (Panels A and B) and validation cohorts (Panels C and D). Panel E compares utility estimates derived from the FAQLQ-PF using the best-fitting model vs the mapping algorithm from Chen et al. based on the AQOL-6D. Panel F compares the mean change in utility estimate following oral immunotherapy calculated with the SF-6Dv2 or with the current algorithm. The lines denote the mean difference. FAQLQ-PF, Food Allergy Quality of Life Questionnaire—Parent Form; SF-6Dv2, Short-Form Six-Dimension version 2

When comparing the change in utility following treatment, the overall mean change in SF-6Dv2- and FAQLQ-PF-derived utility were found to be comparable ($+0.084 \pm 0.124$ vs. $+0.103 \pm 0.085$; $p = .14$) (Figure 1F). There was however a significant difference in the distribution of these estimates (Kolmogorov–Smirnov $D = 0.34$; $p < .0001$), with the SF-6Dv2 failing to detect any improvement in 41 patients (35%).

4 DISCUSSION

In this study, two large independent food allergy cohorts were used to develop a mapping algorithm to calculate food allergy-related utility values directly from the FAQLQ-PF questionnaire. Categorical regression with optimal scaling produced the most consistent model for the prediction of SF-6Dv2-derived utility value. With a MAE of 0.0811, the final model is similar to those from SF-6D mapping studies for comparable diseases.^{17, 27, 28}

There is no universal approach to choosing the right regression technique to predict utility.³ Here, seven different regression techniques were used to cope with various potential issues relating to the dependent variable such as outliers, heteroscedasticity, skewness, and ceiling effect. To account for the different ways that the FAQLQ-PF score could be calculated, models were developed based on FAQLQ-PF average total score, domain scores, or item scores, treating values as numerical or ordinal; based on either all age-relevant items or only items answered by all age categories (section A). Multiple goodness of fit parameters were used to choose the overall best-fitting model. While RMSE is widely used in questionnaire mapping studies, it is more likely to be affected by outliers compared to MAE, hence a benefit of using both methods.²⁹

The performance of the different models provides insight as to the validity of FAQLQ-PF

reporting methods. The inferiority of models 1 and 2 for all indicators (MAE, RMSE, CCC, absolute differences) is evidence that the various items in FAQLQ-PF have different weights. The fact that model 4 performed slightly better than model 3 for MAE, RMSE, and CCC supports the idea of non-equidistant intervals between item levels. For example, the difference between “1-barely” and “2-slightly” may not be the same as the difference between “5-very” and “6-extremely,” which implies that the values should not be treated as continuous numerical values. The MM and CLAD regressions resulted in higher proportion of values with low absolute differences, which likely reflects their robustness compared to other regression techniques. In the index cohort, the CATREG model based on all age-applicable sections of the questionnaire performed better than the one based on only section A. This is expected since each age subgroup is given its best-fitting scenario. However, when tested in the external validation cohort, this model was found to have lower consistency compared to the one based only on section A. The small size of each age subgroup likely affected the robustness of the model based on all age-applicable sections. This is reflected by a lower MAE and RMSE scores in the external validation cohort.

A practical advantage of selecting a model based on section A only is that it greatly simplifies calculation of utility values. The exclusion of questions 15–30 (sections B and C) means that some relevant information may be lost. However, the backward elimination method tended to exclude these questions anyway for the older children. Question 6 was the only item to be systematically included in all models, which suggests that physical distress from food allergy is associated with the greatest deficit in HRQoL. Question 3 (frustrated by dietary restriction) was inversely related to disutility, which may appear counter-intuitive. The variance inflation factor indicates that this was not an artifact of multi-collinearity (Tables [S6](#) and [S8](#)). Removing the term reduced the model's robustness and consistency, which suggests it is important. A possible explanation is that the item describes a less meaningful aspect of food allergy quality of life, which may serve as barometer to control for leniency bias.

In the past, food allergy cost-utility analyses had to rely on utility estimates from generic questionnaires or direct elicitation approaches that had not been validated in the food allergic population to estimate utility values for their models and the health state utility value for desensitized patients was somewhat arbitrarily defined.^{8, 30–34} The new conversion algorithm will allow to directly assess the individual utility gain from interventions in clinical trial. The FAQLQ-PF is already widely used in the food allergy literature to measure changes in HRQoL, which can now be translated to utility. Beyond its use for cost-effective analyses, FAQLQ-PF-derived utility provides a more valid alternative to the FAQLQ-PF average total score to measure food allergy-related quality of life in a quantitative manner.

Recently, Chen et al. mapped the FAQLQ-PF to the AQoL-6D in parents of children aged 7–17 years.²⁶ The SF-6Dv2 and AQoL-6D are similar yet distinctive generic questionnaires that can both be used to derive health utility. Although the AQoL-6D-based algorithm was created for a different age group, derived estimates strongly correlated with those

calculated with the current algorithm, adding yet another element of validity. When comparing the estimates, the current algorithm suffered from less discretization and ceiling effect, likely reflecting the fact that it was generated from a larger sample which allowed more terms to be included in the equation. Lesser ceiling effect is also a characteristic of the SF-6Dv2, compared to other generic questionnaires.^{35, 36}

This study has several strengths. First, multiple regression models were considered to account for the different ways that FAQLQ-PF results can be expressed and control for various issues relating to data distribution. These included the CATREG regression technique, which allows to respect the ordinal nature of FAQLQ-PF item scales. The utility values used as dependent variable were derived from a SF-6Dv2 preference set that had been specifically validated in the Canadian food allergic population, thus providing a more valid estimate than other generic questionnaires.¹⁰ This had been done using a unitary approach that allowed to validate the SF-6Dv2 in a parent population and thus to map its results to a disease-specific parent questionnaire. The algorithm was developed based on a very large index cohort and not only was it directly validated in an independent external sample, but it was also further validated in a pre-post treatment sample assessing change over time, thus exceeding the standards set forth by the National Institute of Clinical Excellence.^{14, 15}

This study has limitations. First, the FAQLQ-PF can only be used in patients aged from 0 to 12 years old. This can limit projections in health economic models with long time horizons. Similar mapping algorithms will need to be developed for the adolescent and adult populations. Second, because the FAQLQ-PF is a proxy questionnaire, it does not include the child's perspective. Future studies mapping FAQLQ child form to utility derived from generic health-related questionnaires, ideally validated in the food allergic population, would be mandated. Validation in other countries would also be warranted. Third, the third cohort was only meant to validate the model. The observed improvement in utility may not be representative of the general food-allergic population or attributable solely to the intervention. Finally, even if the FAQLQ-PF is highly specific to food allergy, the associated utility values remain indirect utility measures. It would be interesting to compare performance with estimates derived from direct elicitation techniques.

In clinical studies, it is generally considered preferable to calculate utilities using the same generic questionnaire that was used to elicit the preference value set. However, in a disease state such as food allergy, where the generic questionnaire fails to capture important subtleties of the disease, an argument could be made that despite a potential loss of trueness due to error propagation when inferring utility from FAQLQ-PF, the net increase in precision would still lead to more accurate measure overall. In fact, the low sensitivity of the SF-6Dv2 for a change in food-related utility in our third cohort (Figure 2) strongly supports this argument. This said, FAQLQ-PF-derived utility can only reflect changes in food allergy-related disutility and does not allow to reliably estimate the absolute global health utility value.

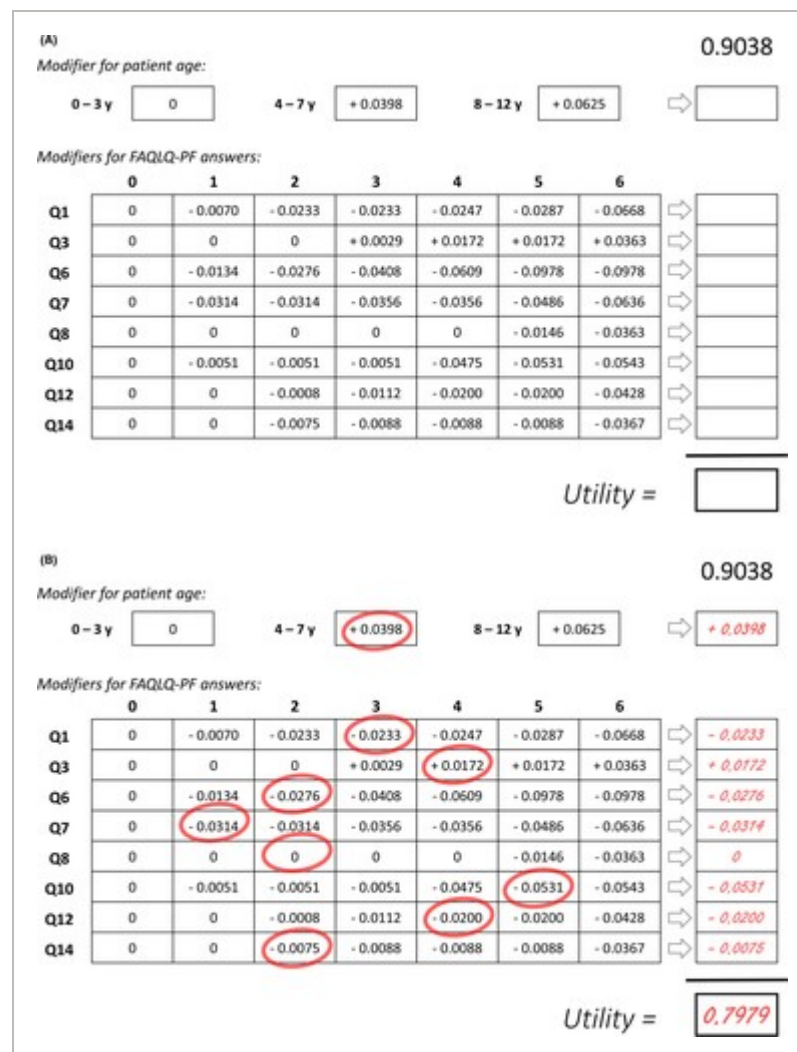


FIGURE 2

[Open in figure viewer](#) | [PowerPoint](#)

Final algorithm allowing to derive health-related utility directly from the FAQLQ-PF. Panel A presents a simplified version of the final mapping algorithm starting from the default utility of a patient answering 0 for each question. Panel B presents the example of utility estimation for a 4-year-old child whose parents answered the FAQLQ-PF with Q1 = 3, Q3 = 4, Q6 = 2, Q7 = 1, Q8 = 2, Q10 = 5, Q12 = 4, Q14 = 2

In conclusion, categorical regression with optimal scaling was used to produce a conversion algorithm mapping SF-6Dv2-derived utility values based on a preference set validated in the food allergic population onto the FAQLQ-PF questionnaire. The indirect elicitation of health utility values through the FAQLQ-PF will allow to measure the intangible benefits associated with food allergy interventions in a quantitative manner for future cost-utility studies or patient-oriented outcomes research.

ACKNOWLEDGMENT

The authors thank the patients and their family for their participation in this research project as well as acknowledge Allergies Québec, Bye Bye Allergies, Déjouer les allergies and Food Allergy Canada for their invaluable help in recruiting participants. This work was

supported by a research grant from the Canadian Allergy, Asthma and Immunology Foundation. TGP is member of the FRQS-funded Centre de recherche de l'IUSMM. PB is supported by a FRQS clinical research scholar grant.

CONFLICT OF INTEREST

The authors declare no conflict of interest related to the study. KS reports personal fees from Novartis outside the submitted work. ADR reports grants from Merck and ALK outside the submitted work. PB reports personal fees from Novartis, Pfizer, Sanofi, ALK, and Aralez Pharmaceuticals, as well as grants from DBV technologies, Regeneron, and Sanofi outside the submitted work.

Supporting Information



Filename	Description
all15190-sup-0001-SupInfo.docx Word 2007 document , 195 KB	Supplementary Material

Please note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.

REFERENCES



- 1 Warren CM, Jiang J, Gupta RS. Epidemiology and burden of food allergy. *Curr Allergy Asthma Rep.* 2020; **20**(2): 6.
[Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)
- 2 Shaker MS, Schwartz J, Ferguson M. An update on the impact of food allergy on anxiety and quality of life. *Curr Opin Pediatr.* 2017; **29**(4): 497- 502.
[Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)
- 3 Rigbi NE, Goldberg MR, Levy MB, Nachshon L, Golobov K, Elizur A. Changes in patient quality of life during oral immunotherapy for food allergy. *Allergy.* 2017; **72**(12): 1883- 1890.
[Wiley Online Library](#) | [CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)
- 4 DunnGalvin A, Fleischer DM, Campbell DE, et al. Improvements in quality of life in children following epicutaneous immunotherapy (EPIT) for peanut allergy in the PEPITES and PEOPLE studies. *J Allergy Clin Immunol Pract.* 2021; **9**(1): 216- 24 e1.
[Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

5 DunnGalvin A, de BlokFlokstra BM, Burks AW, Dubois AE, Hourihane JO. Food allergy QoL questionnaire for children aged 0–12 years: content, construct, and cross-cultural validity. *Clin Exp Allergy*. 2008; **38**(6): 977- 986.

[Wiley Online Library](#) | [CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

6 Liddell TM, Kruschke JK. Analyzing ordinal data with metric models: what could possibly go wrong? *J Exp Soc Psychol*. 2018; **79**: 328- 348.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

7 Azzano P, Dufresne E, Poder T, Begin P. Economic considerations on the usage of biologics in the allergy clinic. *Allergy*. 2021; **76**(1): 191- 209.

[Wiley Online Library](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

8 Shaker M, Greenhawt M. A primer on cost-effectiveness in the allergy clinic. *Ann Allergy Asthma Immunol*. 2019; **123**(2): 120- 8 e1.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

9 Fauteux VPT. Overview of QALY elicitation methods. *Int J Health Pref Research*. 2017; **1**: 2- 14.

[Google Scholar](#)

10 Dufresne E, Poder TG, Samaan K, et al. SF-6Dv2 preference value set for health utility in food allergy. *Allergy*. 2021; **76**(1): 326- 338.

[Wiley Online Library](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

11 Heniford BT, Walters AL, Lincourt AE, Novitsky YW, Hope WW, Kercher KW. Comparison of generic versus specific quality-of-life scales for mesh hernia repairs. *J Am Coll Surg*. 2008; **206**(4): 638- 644.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

12 Maly M, Vondra V. Generic versus disease-specific instruments in quality-of-life assessment of chronic obstructive pulmonary disease. *Methods Inf Med*. 2006; **45**(2): 211- 215.

[Crossref](#) | [CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

13 Chen G, Garcia-Gordillo MA, Collado-Mateo D, et al. Converting Parkinson-specific scores into health state utilities to assess cost-utility analysis. *Patient*. 2018; **11**(6): 665- 675.

[Crossref](#) | [PubMed](#) | [Google Scholar](#)

14 Longworth L, Rowen D. *NICE DSU Technical Support Document 10: The Use of Mapping Methods to Estimate Health State Utility Values*. NICE Decision Support Unit Technical Support Documents; 2011.

[Google Scholar](#)

15 Wailoo AJ, Hernandez-Alava M, Manca A, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. *Value Health*. 2017; **20**(1): 18- 27.

[Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

16 Le Blanc V, Samaan K, Paradis L, et al. Treatment expectations in food-allergic patients referred for oral immunotherapy. *J Allergy Clin Immunol Pract*. 2021; **9**(5): 2087- 2089.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

17 Abdin E, Chong SA, Seow E, Verma S, Tan KB, Subramaniam M. Mapping the Positive and Negative Syndrome Scale scores to EQ-5D-5L and SF-6D utility scores in patients with schizophrenia. *Qual Life Res*. 2019; **28**(1): 177- 186.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

18 Lamu AN, Olsen JA. Testing alternative regression models to predict utilities: mapping the QLQ-C30 onto the EQ-5D-5L and the SF-6D. *Qual Life Res*. 2018; **27**(11): 2823- 2839.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

19 Pullenayegum EM, Tarride JE, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? *Value Health*. 2010; **13**(4): 487- 494.

[Wiley Online Library](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

20 Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 multiattribute utility instruments. *Med Decis Making*. 2016; **36**(2): 160- 175.

[Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

21 Kularatna S, Senanayake S, Chen G, Parsonage W. Mapping the Minnesota living with heart failure questionnaire (MLHFQ) to EQ-5D-5L in patients with heart failure. *Health Qual Life Outcomes*. 2020; **18**(1): 115.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

22 Royston P, Sauerbrei W. Multivariable modeling with cubic regression splines: a principled approach. *Stata J*. 2007; **7**(1): 45- 70.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

23 Panchagnula S, Sun X, Montejo JD, et al. Validating the transformation of PROMIS-GH to EQ-5D in adult spine patients. *J Clin Med*. 2019; **8**(10): 1506.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

24 Kaambwa B, Chen G, Ratcliffe J, Iezzi A, Maxwell A, Richardson J. Mapping between the Sydney asthma quality of life questionnaire (AQLQ-S) and five multi-attribute utility instruments (MAUIs). *Pharmacoeconomics*. 2017; **35**(1): 111- 124.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

25 Lin L, Hedayat AS, Sinha B, Yang M. Statistical methods in assessing agreement: models, issues, and tools. *J Am Stat Assoc*. 2002; **97**(457): 257- 270.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

26 Chen G, DunnGalvin A, Greenhawt M, Shaker M, Campbell DE. Deriving health utility indices from a food allergy quality-of-life questionnaire. *Pediatr Allergy Immunol*. 2021; **32**(8): 1773- 1780. [Wiley Online Library](#) | [Web of Science®](#) | [Google Scholar](#)

27 Monroy M, Ruiz MA, Rejas J, Soto J. Mapping of the Gastrointestinal Short Form Questionnaire (GSF-Q) into EQ-5D-3L and SF-6D in patients with gastroesophageal reflux disease. *Health Qual Life Outcomes*. 2018; **16**(1): 177. [Crossref](#) | [Google Scholar](#)

28 Chen G, Iezzi A, McKie J, Khan MA, Richardson J. Diabetes and quality of life: comparing results from utility instruments and Diabetes-39. *Diabetes Res Clin Pract*. 2015; **109**(2): 326- 333. [Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

29 Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast*. 2006; **22**(4): 679- 688. [Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

30 Carroll AE, Downs SM. Improving decision analyses: parent preferences (utility values) for pediatric health outcomes. *J Pediatr*. 2009; **155**(1): 21- 25, e1-e5. [Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

31 Mittmann N, Trakas K, Risebrough N, Liu BA. Utility scores for chronic conditions in a community-dwelling population. *Pharmacoeconomics*. 1999; **15**(4): 369- 376. [Crossref](#) | [CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

32 Protudjer JL, Jansson SA, Ostblom E, et al. Health-related quality of life in children with objectively diagnosed staple food allergy assessed with a disease-specific questionnaire. *Acta Paediatr*. 2015; **104**(10): 1047- 1054. [Wiley Online Library](#) | [CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

33 Shaker M, Chan ES, Protudjer JLP, Soller L, Abrams EM, Greenhawt M. The cost-effectiveness of preschool peanut oral immunotherapy in the real-world setting. *J Allergy Clin Immunol Pract*. 2021; **9**(7): 2876- 2884.e4. [Crossref](#) | [Google Scholar](#)

34 Tice JA, Guzauskas GF, Hansen RN, et al. The effectiveness and value of oral immunotherapy and Viaskin peanut for peanut allergy. *J Manag Care Spec Pharm*. 2020; **26**(5): 620- 623. [Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

35 Poder TG, Fauteux V, He J, Brazier JE. Consistency between three different ways of administering the short form 6 dimension version 2. *Value Health*. 2019; **22**(7): 837- 842. [Crossref](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

36 Poder TG, Carrier N, Kouakou CRC. Quebec health-related quality-of-life population norms using the EQ-5D-5L: decomposition by sociodemographic data and health problems. *Value Health*.

2020; **23**(2): 251- 259.

[Crossref](#) | [Web of Science®](#) | [Google Scholar](#)

[Download PDF](#)

About Wiley Online Library

Privacy Policy

Terms of Use

Cookies

Accessibility

Publishing Policies

Help & Support

Contact Us

Training and Support

DMCA & Reporting Piracy

Opportunities

Subscription Agents

Advertisers & Corporate Partners

Connect with Wiley

The Wiley Network

Wiley Press Room

Copyright © 1999-2021 John Wiley & Sons, Inc. All rights reserved

SUPPLEMENTARY – TABLES AND FIGURES

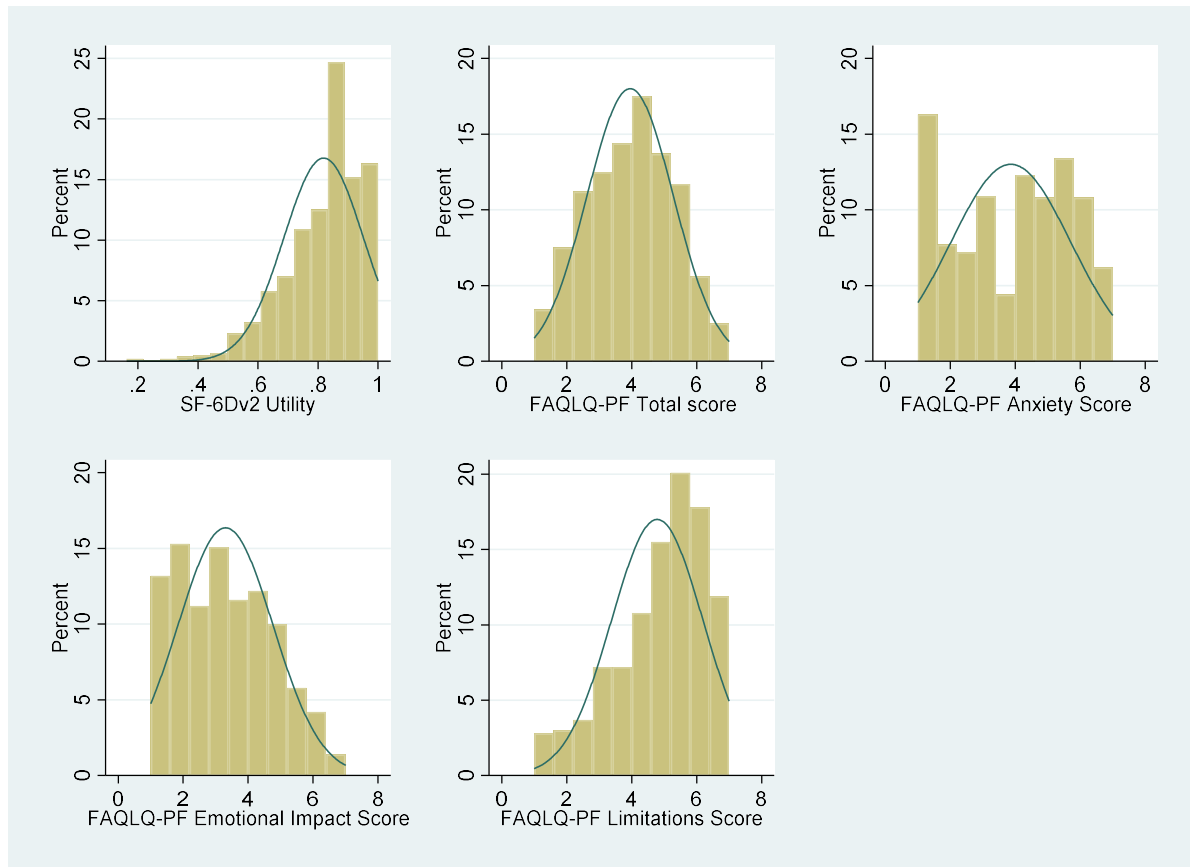


Figure S1 – Frequency distribution of SF-6Dv2 total score and FAQLQ-PF total and domain scores. Histograms bars present the distribution of the SF-6Dv2 total score, the FAQLQ-PF total average score and the average score for its anxiety, emotional impact, and limitations domains in the index cohort. The over-layed teal line represents the best fit for a normal distribution. FAQLQ-PF: Food Allergy Quality of Life Questionnaire – Parent Form.

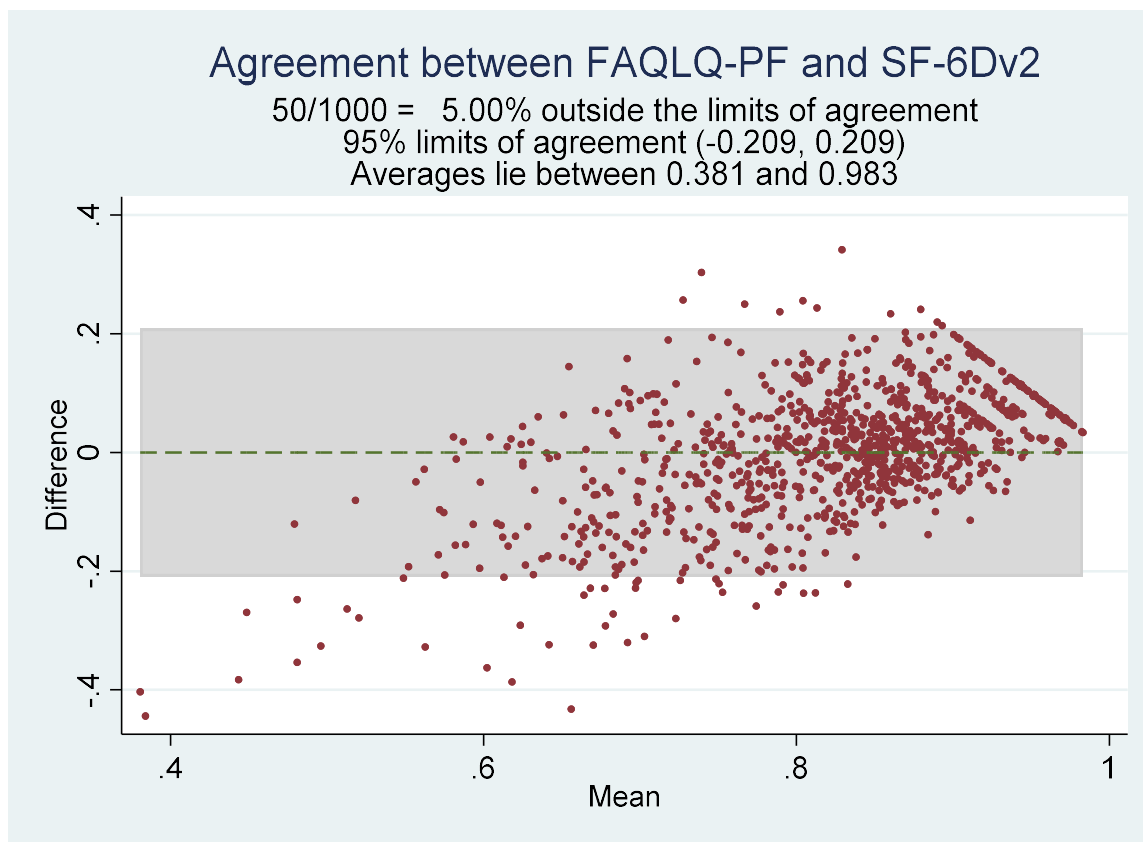


Figure S2 – Bland-Altman plot for the differences between the observed and the predicted SF-6Dv2 utility scores.

Model 1	$u_{SF6Dv2} = \beta_0 + \beta_1 \times FAQLQ - PF \text{ total score} + \beta_2 \times FAQLQ - PF \text{ total score}^2 + \beta_3 \times age \text{ group}$
Model 2	$u_{SF6Dv2} = \beta_0 + \beta_1 \times Emotional \text{ impact domain} + \beta_2 \times Emotional \text{ impact domain}^2 + \beta_3 \times Anxiety \text{ domain} + \beta_4 \times Anxiety \text{ domain}^2 + \beta_5 \times Limitations \text{ domain} + \beta_6 \times Limitations \text{ domain}^2 + \beta_7 \times age \text{ group}$
Model 3	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{14} \beta_i \times FAQLQ - PF \text{ item scores} + \sum_{i=1}^{14} \beta_i \times FAQLQ - PF \text{ item scores}^2 + \beta_2 \times age \text{ group}$
Model 4	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{14} \beta_i \times \{Q_{1_1}, \dots, Q_{1_7}\} + \beta_2 \times age \text{ group}$

Table S1 – Regression models (FAQLQ-PF section A). Four models adjusted with FAQLQ-PF items from section A (ranging from 1 to 14). The total score and domain scores were calculated using only section A items. Model 4 is the CATREG regression method. Dependent variable: SF-6Dv2 (u_{SF6Dv2}). FAQLQ-PF: Food Allergy Quality of Life Questionnaire – Parent Form.

Age group (0 – 3y)	
Model 1	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{FAQLQ total score} + \beta_2 \times \text{FAQLQ total score}^2$
Model 2	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{Emotional impact domain} + \beta_2 \times \text{Emotional impact domain}^2 + \beta_3 \times \text{Anxiety domain} + \beta_4 \times \text{Anxiety domain}^2 + \beta_5 \times \text{Limitations domain} + \beta_6 \times \text{Limitations domain}^2$
Model 3	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{14} \beta_i \times \text{FAQLQ item scores} + \sum_{i=1}^{14} \beta_i \times \text{FAQLQ item scores}^2$
Model 4	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{14} \beta_i \times \{Q_{1_1}, \dots, Q_{1_7}\}$
Age group (4 – 7y)	
Model 1	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{FAQLQ total score} + \beta_2 \times \text{FAQLQ total score}^2$
Model 2	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{Emotional impact domain} + \beta_2 \times \text{Emotional impact domain}^2 + \beta_3 \times \text{Anxiety domain} + \beta_4 \times \text{Anxiety domain}^2 + \beta_5 \times \text{Limitations domain} + \beta_6 \times \text{Limitations domain}^2$
Model 3	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{26} \beta_i \times \text{FAQLQ item scores} + \sum_{i=1}^{14} \beta_i \times \text{FAQLQ item scores}^2$
Model 4	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{26} \beta_i \times \{Q_{1_1}, \dots, Q_{1_7}\}$
Age group (8 – 12y)	
Model 1	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{FAQLQ total score} + \beta_2 \times \text{FAQLQ total score}^2$
Model 2	$u_{SF6Dv2} = \beta_0 + \beta_1 \times \text{Emotional impact domain} + \beta_2 \times \text{Emotional impact domain}^2 + \beta_3 \times \text{Anxiety domain} + \beta_4 \times \text{Anxiety domain}^2 + \beta_5 \times \text{Limitations domain} + \beta_6 \times \text{Limitations domain}^2$
Model 3	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{30} \beta_i \times \text{FAQLQ item scores} + \sum_{i=1}^{30} \beta_i \times \text{FAQLQ item scores}^2$
Model 4	$u_{SF6Dv2} = \beta_0 + \sum_{i=1}^{30} \beta_i \times \{Q_{1_1}, \dots, Q_{1_7}\}$

Table S2 – Regression models (FAQLQ-PF according to age group (Version B)). Four models adjusted with FAQLQ-PF items for age-applicable sections (a model per age group). The total score and domain scores were calculated for each age group. Model 4 is the CATREG

regression method. Dependent variable: SF-6Dv2 (u_{SF6Dv}). FAQLQ-PF: Food Allergy Quality of Life Questionnaire – Parent Form.

Clinical background	Group		Mean difference (No – Yes)	p-value
	Yes (n) Mean $\pm$ SD	No (n) Mean $\pm$ SD		
Anaphylaxis	(252) 0.794 $\pm$ 0.135	(748) 0.828 $\pm$ 0.131	0.034	< 0.001
Atopic dermatitis	(156) 0.793 $\pm$ 0.139	(844) 0.824 $\pm$ 0.131	0.031	0.009
Asthma	(382) 0.804 $\pm$ 0.135	(601) 0.829 $\pm$ 0.131	0.025	0.005
Multiple food allergies (>1)	(712) 0.810 $\pm$ 0.135	(288) 0.842 $\pm$ 0.125	0.031	< 0.001
High Risk food allergies (>1)	(433) 0.791 $\pm$ 0.146	(567) 0.841 $\pm$ 0.118	0.050	< 0.001

Table S3 – T-test comparisons (unequal variance) SF-6Dv2 for clinical background.

Parameter	Group		Mean difference (I – J)	p-value*
	I (n) Mean ± SD	J (n) Mean ± SD		
Age	0-3 (307) 0.789 ± 0.141	4-7 (376) 0.822 ± 0.134	0.033	0.003
	0-3 (307) 0.789 ± 0.141	8-12 (317) 0.844 ± 0.118	0.055	< 0.001
	4-7 (376) 0.822 ± 0.134	8-12 (317) 0.844 ± 0.118	0.022	0.091

Table S4 – ANOVA for SF-6Dv2 utility comparisons by patient's age. * Multiplicity of tests controlled by Bonferroni's correction.

Models	Version A – Section A items	Version B – Age-applicable sections
Model 3		
OLS/GLM		• Section A : Q1, Q6, Q14² • Section B : Q6, Q8², Q10², Q19² • Section C : Q2, Q29, Q3², Q6², Q10², Q12², Q23², Q29²
CLAD	Q7, Q8, Q1 ² , Q6 ² , Q8 ² , Q10 ² , Q11 ² , Q12 ² , Q14 ² , A2, A3	
MM		
TOBIT	Q6, Q7, Q8, Q1 ² , Q3 ² , Q8 ² , Q10 ² , Q12 ² , Q14 ² , A2, A3	• Section A : Q1, Q6, Q14² • Section B : Q6, Q4², Q8², Q10², Q19² • Section C : Q25, Q29, Q2², Q3², Q6², Q10², Q12², Q23², Q29²
MFP	Q1, Q3, Q6, Q7, Q8, Q10, Q12, Q14	• Section A : Q1, Q6, Q14 • Section B : Q6, Q8, Q10, Q23, Q25 • Section C : Q2, Q3, Q6, Q10, Q12, Q23
Model 4		
CATREG	Q1, Q3, Q6, Q7, Q8, Q10, Q12, Q14	• Section A : Q1, Q6, Q14 • Section B : Q6, Q7, Q8, Q10, Q11, Q13, Q23 • Section C : Q1, Q2, Q3, Q5, Q6, Q10, Q15, Q19, Q25

Table S5 – Selected questions for items in models 3 and 4 for both versions with backwards elimination method. OLS regression method selects the questions for CLAD and MM. GLM chose the family gaussian with identity link function, meaning OLS and GLM have the same results. Version A is only Section A items and Version B is all age-applicable sections with one model by section. Dependent variable – SF-6Dv2. OLS – Ordinary least square, GLM – General linear model, TOBIT – Censored regression, CLAD – Censored least absolute deviations, MM – Robust MM-estimator, MFP – Multivariable fractional polynomials, CATREG – Categorical regression with optimal scaling.

Items	VIF
Q1	1.91
Q3	1.88
Q6	1.50
Q7	2.48
Q8	1.47
Q10	1.76
Q12	1.88
Q14	2.02

Table S6 – Variance inflation factor for items included in the CATREG A model.

Items	β (SE)	<i>Item Levels (j)</i>						
		1	2	3	4	5	6	7
<i>Q1</i>	-0.133 (0.038)	-1.237	-0.839	0.084	0.084	0.160	0.385	2.543
<i>Q3</i>	0.091 (0.036)	-0.979	-0.979	-0.979	-0.740	0.443	0.443	2.021
<i>Q6</i>	-0.220 (0.035)	-0.706	-0.247	0.238	0.688	1.376	2.640	2.640
<i>Q7</i>	-0.141 (0.037)	-1.503	0.173	0.173	0.395	0.395	1.089	1.890
<i>Q8</i>	-0.086 (0.041)	-0.468	-0.468	-0.468	-0.468	-0.468	0.809	2.709
<i>Q10</i>	-0.155 (0.035)	-0.690	-0.442	-0.442	-0.442	1.618	1.887	1.944
<i>Q12</i>	-0.117 (0.033)	-1.704	-1.704	-1.651	-0.986	-0.416	-0.416	1.049
<i>Q14</i>	-0.100 (0.041)	-1.114	-1.114	-0.551	-0.454	-0.454	-0.454	1.644
β (SE)		<i>Age (years)</i>						
		0 – 3		4 – 7		8 – 12		
0.188 (0.031)		-1.392		0.202		1.108		

Table S7 – Coefficients and quantifications for the CATREG A regression model.

Items	Q1	Q3	Q6	Q7	Q8	Q10	Q12	Q14
Q1	1.000							
Q3	0.600	1.000						
Q6	0.206	0.204	1.000					
Q7	0.597	0.567	0.497	1.000				
Q8	0.295	0.286	0.432	0.410	1.000			
Q10	0.475	0.480	0.338	0.589	0.386	1.000		
Q12	0.241	0.350	0.241	0.359	0.402	0.359	1.000	
Q14	0.346	0.421	0.262	0.423	0.368	0.454	0.657	1.000

Table S8 – Correlations (Pearson) between selected items in CATREG A.