

UCLA

UCLA Previously Published Works

Title

Development of a Preference-Based Index From the National Eye Institute Visual Function Questionnaire—25

Permalink

<https://escholarship.org/uc/item/5nn4q873>

Journal

JAMA Ophthalmology, 132(3)

ISSN

2168-6165

Authors

Rentz, Anne M

Kowalski, Jonathan W

Walt, John G

et al.

Publication Date

2014-03-01

DOI

10.1001/jamaophthalmol.2013.7639

Peer reviewed

## Original Investigation | SOCIOECONOMICS AND HEALTH SERVICES

# Development of a Preference-Based Index From the National Eye Institute Visual Function Questionnaire–25

Anne M. Rentz, MSPH; Jonathan W. Kowalski, PharmD, MS; John G. Walt, MBA; Ron D. Hays, PhD; John E. Brazier, PhD; Ren Yu, MA; Paul Lee, MD; Neil Bressler, MD; Dennis A. Revicki, PhD

**IMPORTANCE** Understanding how individuals value health states is central to patient-centered care and to health policy decision making. Generic preference-based measures of health may not effectively capture the impact of ocular diseases. Recently, 6 items from the National Eye Institute Visual Function Questionnaire–25 were used to develop the Visual Function Questionnaire–Utility Index health state classification, which defines visual function health states.

**OBJECTIVE** To describe elicitation of preferences for health states generated from the Visual Function Questionnaire–Utility Index health state classification and development of an algorithm to estimate health preference scores for any health state.

**DESIGN, SETTING, AND PARTICIPANTS** Nonintervention, cross-sectional study of the general community in 4 countries (Australia, Canada, United Kingdom, and United States). A total of 607 adult participants were recruited from local newspaper advertisements. In the United Kingdom, an existing database of participants from previous studies was used for recruitment.


**INTERVENTIONS** Eight of 15 625 possible health states from the Visual Function Questionnaire–Utility Index were valued using time trade-off technique.

**MAIN OUTCOMES AND MEASURES** A  $\theta$  severity score was calculated for Visual Function Questionnaire–Utility Index–defined health states using item response theory analysis. Regression models were then used to develop an algorithm to assign health state preference values for all potential health states defined by the Visual Function Questionnaire–Utility Index.

**RESULTS** Health state preference values for the 8 states ranged from a mean (SD) of 0.343 (0.395) to 0.956 (0.124). As expected, preference values declined with worsening visual function. Results indicate that the Visual Function Questionnaire–Utility Index describes states that participants view as spanning most of the continuum from full health to dead.

**CONCLUSIONS AND RELEVANCE** Visual Function Questionnaire–Utility Index health state classification produces health preference scores that can be estimated in vision-related studies that include the National Eye Institute Visual Function Questionnaire–25. These preference scores may be of value for estimating utilities in economic and health policy analyses.

*JAMA Ophthalmol.* doi:10.1001/jamaophthalmol.2013.7639  
Published online January 16, 2014.

 Supplemental content at  
jamaophthalmology.com

**Author Affiliations:** Evidera, Bethesda, Maryland (Rentz, Yu, Revicki); Allergan Inc, Irvine, California (Kowalski, Walt); Department of Medicine, University of California, Los Angeles (Hays); School of Health and Related Research, Sheffield University, Sheffield, England (Brazier); Duke University, Durham, North Carolina (Lee); now with Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor (Lee); Retina Division, Wilmer Eye Institute, Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, Maryland (Bressler).

**Corresponding Author:** Anne M. Rentz, MSPH, Evidera, 7101 Wisconsin Ave, Ste 600, Bethesda, MD 20814 (anne.rentz@evidera.com).

**Section Editor:** Paul P. Lee, MD.

Use of preference-based health-related quality-of-life (HRQOL) measures has increased owing to the increased use of economic evaluation in creating health policy.<sup>1,2</sup> The US Public Health Service Panel on Cost-effectiveness in Health and Medicine issued recommendations supporting the use of preference-based measures to calculate quality-adjusted life-years (QALYs) for economic evaluations.<sup>3</sup> These QALYs are used to quantify HRQOL outcomes for economic evaluations.<sup>4-7</sup> The QALYs represent the product of HRQOL and survival, allowing effectiveness to be quantified in terms of the degree that the intervention changes both. The UK National Institute for Health and Clinical Guidance issued a guidance expressing preference for economic evaluations using generic preference-based utility measures, specifically the EuroQol EQ-5D.<sup>8</sup> This guidance stated that in the absence of EQ-5D data, empirical mapping to the EQ-5D from other HRQOL instruments or the valuation of health states based on other instruments may be used as an alternative.

Because there are many conditions for which utilities are not available, have inconsistent results across studies, or are inadequately represented by available generic preference instruments, interest in expressing preference for different health technologies using disease-targeted measures is growing.<sup>6,7,9,10</sup> Disease-targeted measures are viewed as more sensitive to treatment changes and more relevant to the impact on HRQOL, especially for chronic medical conditions.<sup>6,7,10-16</sup> The EQ-5D, for example, has been shown to have limited sensitivity to visual function in patients with age-related macular degeneration.<sup>17</sup> Recently, innovative methods have been used to estimate utilities from existing instruments such as the 36-Item Short Form Health Survey, the King's Health Questionnaire, and the Cambridge Pulmonary Hypertension Outcome Review.<sup>6,10,18</sup> Using a disease-targeted HRQOL measure for preference measurement has the potential advantage of using a more sensitive descriptive system to classify people into health states.

The most widely used ocular disease-targeted HRQOL measure is the National Eye Institute Visual Function Questionnaire-25 (NEI VFQ-25).<sup>19,20</sup> Six items (ie, items 6, 11, 14, 18, 20, and 25) representing 6 of the NEI VFQ-25 subscales were recently converted into a health state classification system called the *Visual Function Questionnaire-Utility Index* (VFQ-UI).<sup>6,21,22</sup> This article provides an overview of the development of the VFQ-UI health state classification system, the preference elicitation for selected vision-related health states from the general public, and development of the VFQ-UI scoring algorithm.

The objectives of this study were to obtain health preferences for VFQ-UI-related health states in more than 600 members of the general population in Australia, Canada, the United Kingdom, and the United States and to develop an algorithm for estimating utility scores. The utility scores derived using the VFQ-UI algorithm can then be used in estimating QALYs for economic evaluations comparing treatment interventions for ocular diseases and for health technology assessments and policy decisions. Utility scores derived from the disease-targeted NEI VFQ-25 may prove to provide more sensitive preference-based estimates of utilities for patients with vary-

ing levels of vision loss compared with generic preference measures such as the EQ-5D or SF-6D.

## Methods

The aim of this study was to generate a preference-based scoring algorithm for estimating utility scores based on the VFQ-UI health state classification. The overall approach for developing the VFQ-UI health state classification system and VFQ-UI scores comprised 3 stages: (1) developing a health state classification system; (2) conducting a valuation study of the health states; and (3) developing the scoring algorithm using multivariate regression analyses (Box).<sup>23</sup>

### Development of the VFQ-UI Health State Classification

A health state classification system is a multidimensional framework that can be used to define health states. Such classifications define a set of health states by selecting 1 level from each dimension in the system. A secondary analysis was performed using NEI VFQ-25 data collected in several studies of patients with either central (n = 968)<sup>24-28</sup> or peripheral (n = 2451)<sup>29,30</sup> vision loss (eAppendix 1 and eTable 1 in Supplement) to identify the best subset of NEI VFQ-25 items that capture the overall content of the instrument.<sup>23</sup> The general health item was removed because it is not specific to visual function. The driving subscale was removed because of high missing data rates and to be more generalizable to countries other than the United States, while the ocular pain subscale was excluded owing to mis-fit within both samples.<sup>23</sup> Next, we used Rasch analysis<sup>31,32</sup> to reduce the NEI VFQ-25 to a simpler descriptive health state classification by identifying the severity level captured by the items.

Rasch analysis was used to transfer categorical item responses to points on a latent scale using a logit model, where the underlying scale is treated as continuous.<sup>22</sup> Central and peripheral vision loss data sets that included NEI VFQ-25 data

#### Box. Developing the Visual Function Questionnaire-Utility Index (VFQ-UI) Health State Classification System and Scoring Methods

- Stage 1: Developing a health state classification system
- Dimensionality analyses
  - Selecting subset of relevant National Eye Institute Visual Function Questionnaire-25 items
  - Defining the VFQ-UI health state classification system
  - Identifying representative health states
- Stage 2: Conducting a valuation study of the health states
- Selecting a general population sample
  - Preference valuation for selected representative VFQ-UI health states
  - Time trade-off techniques used for valuation exercise
- Stage 3: Developing the VFQ-UI scoring algorithm
- Mapping relationship between preference valuations and item levels for VFQ-UI health state classification system
  - Multivariate regression analysis used for mapping relationship
  - Develop VFQ-UI scoring algorithm based on regression model results

were examined independently to identify items fitting each vision loss type and any differences in item performance. Rasch models were used to evaluate the unidimensionality of the VFQ-25 domains, item level ordering and fit, and differential item functioning (eAppendix 1 in Supplement). After that review, central and peripheral vision loss data sets were pooled to allow for selection of items relevant across ocular indications. Similar analyses were performed on these pooled data. These 3 sets of analyses were reviewed by clinical and psychometric experts, and decisions were made regarding the final composition of the health state classification.<sup>23</sup>

Based on these analyses and clinical review, 1 item from each of 6 NEI VFQ-25 subscales (near vision activities, distance vision activities, vision-specific social functioning, role difficulties, dependency, and mental health) was selected for the reduced health state classification (ie, VFQ-UI health state classification).

### Design and Conduct of the Valuation Study

#### Development of Health States and Pilot Testing

Based on the VFQ-UI health state classification, 8 vision-related health states were developed for the preference valuation study (of a total possible 15 625).<sup>23</sup> These health states were selected to describe best vision function states (111111), worst vision function states (555555), and intermediate health states reflecting vision-related functioning and well-being. We pilot tested the health states in interviews with 22 patients with age-related macular degeneration, glaucoma, or macular edema to explore the content validity of the health states. Patient interviews confirmed the relevance and face validity of the health states and that overall they were understandable and accurate (eAppendix 2 in Supplement).

#### General Population Valuation Sample

The health state valuation survey was conducted with participants from the general public from Australia, Canada, the United Kingdom, and the United States. Trained personnel conducted one-on-one interviews. By interviewing the general public as opposed to those with specified eye conditions, we were able to elicit findings more generalizable to the preferences of overall society in the participating countries.

Participants were recruited from local newspaper advertisements. In the United Kingdom, an existing database of participants from previous studies was also used for recruitment. Eligible participants must have been aged 18 years or older; a current resident of the interview country; able to understand and complete the survey (as judged by the interviewer); and willing and able to give written informed consent. The study protocol and consent form were submitted to a central institutional review board, approved, and met Health Insurance Portability and Accountability Act of 1996 requirements.

#### Interview for Preference Elicitation

Participants took part in a health state ranking exercise and a time trade-off (TTO) interview and were asked to complete the EQ-5D. In the ranking task, participants were asked to take the 8 health states along with the anchor states (full health and

dead) and put them in order from best to worst. This task allowed participants to familiarize themselves with the health states and to understand their differences.

For the TTO interview, we used a TTO board<sup>33</sup> containing time lines for the health state comparisons. The TTO asks participants to imagine they will be in a given vision-related health state for 10 years and then asks them to compare this vision-related health state with a number of shorter periods in full health (x) after which time the future is uncertain. The valuation of the targeted state is given as  $x/10$  when the participant reaches a situation where he or she is indifferent or unable to select between the health state and full health alternatives. The TTOs used a 10-year time horizon given the long-term chronic nature of loss of visual functioning for many ocular diseases (eAppendix 1 in Supplement).

Target vision-related health states were rated against full health and pits state (worst visual functioning health state defined by the classification) and then pits state against full health and dead.

The pits or worst health state as defined by the VFQ-UI was used as the lower anchor because using dead as the lower anchor during the interviews could be insensitive to the effect of loss of visual acuity (VA). Results need to be presented on a scale of full health to dead, so the rating for the pits state on the scale of full health to dead was also separately collected during interviews and then used to calibrate previous ratings to the scale of full health to dead. The equation for this adjustment is  $TTOADJ = [TTO + (1 - P)] \times P$ , where  $TTOADJ$  is the TTO-adjusted score and  $P$  is the pits state.<sup>21</sup>

Participants also completed the EQ-5D,<sup>34</sup> a generic preference-based health status measure.

### Regression Analyses for Modeling Health State Preferences

Our aim was to develop a scoring algorithm that could assign values for all states defined by the VFQ-UI health state classification. We used an approach summarized by Young et al<sup>35</sup> for developing preference-based, disease-specific measures when the items in the instrument are not independent (ie, are correlated). Additional details on the analyses are included in eAppendix 1 in the Supplement.

As described later, we applied item response theory (IRT) analyses to obtain an indicator of severity for each health state defined by the VFQ-UI classification system and then mapped the severity indicator onto the utilities of targeted study health states.<sup>22,36</sup> We combined information from 2 different sets of data: (1) VFQ-UI data used in developing the health state classification from patients with central and peripheral vision loss; and (2) health preference valuation interview study data. Patient data used to develop the health state classification are described in detail elsewhere.<sup>23</sup>

First, with the data set from the VFQ-UI data used in developing the health state classification from patients with central and peripheral vision loss, we estimated severity ( $\theta$ ) scores from the patient-level responses to the 6 NEI VFQ-25 items that compose the VFQ-UI using a graded response model.<sup>37</sup> The  $\theta$  value represents the location of the health states in terms of vision-related function, where higher scores indicate better functioning.<sup>38,39</sup> Regression models were then run to map the

Table 1. Time Trade-Off Interview Participant Demographic Characteristics by Country

Characteristic	No. (%)			
	Australia (n = 150)	Canada (n = 150)	United Kingdom (n = 152)	United States (n = 155)
Age, mean (SD), y	35.6 (14.5)	41.2 (15.8)	41.8 (13.7)	51.6 (16.3)
Sex				
Male	57 (38.0)	65 (43.3)	55 (36.2)	69 (44.5)
Female	93 (62.0)	85 (56.7)	97 (63.8)	86 (55.5)
Employment status				
Employed full time	55 (36.7)	74 (49.3)	69 (45.4)	59 (38.1)
Employed part time	39 (26.0)	23 (15.3)	32 (21.1)	26 (16.8)
Homemaker	2 (1.3)	6 (4.0)	2 (1.3)	3 (1.9)
Student	24 (16.0)	14 (9.3)	18 (11.8)	7 (4.5)
Unemployed	9 (6.0)	10 (6.7)	10 (6.6)	12 (7.7)
Retired	10 (6.7)	18 (12.0)	11 (7.2)	38 (24.5)
Disabled or unable to work	3 (2.0)	4 (2.7)	4 (2.6)	9 (5.8)
Other	2 (1.3)	0	6 (3.9)	0
Missing	6 (4.0)	1 (0.7)	0	1 (0.6)
Highest education				
No formal qualification	2 (1.3)	NA	4 (2.6)	NA
Some high school	NA	4 (2.7)	NA	4 (2.6)
High school	27 (18.0)	30 (20.0)	NA	26 (16.8)
General Certificate of Secondary Education, O levels, or equivalent	NA	NA	21 (13.8)	NA
A levels or equivalent	NA	NA	21 (13.8)	NA
Vocation or work-based qualification	38 (25.3)	NA	14 (9.2)	NA
University degree or college	74 (49.3)	84 (56.0)	75 (49.3)	71 (45.8)
Graduate school	NA	24 (16.0)	NA	49 (31.6)
Other	7 (4.7)	7 (4.7)	14 (9.2)	4 (2.6)
Missing	2 (1.3)	1 (0.7)	3 (2.0)	1 (0.6)
AMD				
No	145 (96.7)	148 (98.7)	148 (97.4)	150 (96.8)
Yes	3 (2.0)	1 (0.7)	3 (2.0)	4 (2.6)
Missing	2 (1.3)	1 (0.7)	1 (0.7)	1 (0.6)
Know someone with an eye disease				
No	131 (87.3)	114 (76.0)	130 (85.5)	95 (61.3)
Yes				
AMD	11 (7.3)	18 (12.0)	12 (7.9)	35 (22.6)
Glaucoma	3 (2.0)	10 (6.7)	3 (2.0)	12 (7.7)
Other	4 (2.7)	7 (4.7)	6 (3.9)	12 (7.8)
Missing	1 (0.7)	1 (0.7)	1 (0.7)	1 (0.6)

Abbreviations: AMD, age-related macular degeneration; NA, not applicable.

relationship between TTO preference scores and selected demographic variables and VFQ-UI  $\theta$  values. Different regression models were explored to determine whether linear or non-linear regressions represented a better fit in estimating TTO scores. The regression models included age, sex, and education. All modeling was done using SAS version 9.1 statistical software (SAS Institute, Inc).  $P < .05$  was considered statistically significant.

To evaluate the validity of the VFQ-UI scores, we compared baseline mean VFQ-UI scores from a clinical trial of patients with uveitis by best-corrected VA (BCVA) groups using analysis of covariance, adjusting for sex. The BCVA groups were defined as 20/40 or better, worse than 20/40 to better than 20/200, and 20/200 or worse. The relationships between base-

line VFQ-UI and NEI VFQ-25 composite and subscale scores were evaluated using Spearman rank-order correlations.

## Results

### Preference Survey Respondents

In total, 607 participants in Australia, Canada, the United Kingdom, and the United States took part in the valuation survey. The mean age within each country ranged from 36 to 52 years, with the Australian sample being the youngest (Table 1). All countries had a larger proportion of female participants. A small proportion (11 participants [2%]) had age-related macular degeneration.

**Table 2. EQ-5D Index and Visual Analog Scale Scores by Country**

EQ-5D Scale	Australia (n = 150)	Canada (n = 150)	United Kingdom (n = 152)	United States (n = 155)
Index score				
Mean (SD)	0.92 (0.12)	0.93 (0.11)	0.93 (0.12)	0.92 (0.13)
Range	0.41-1.00	0.31-1.00	0.52-1.00	0.45-1.00
IQR	0.85-1.00	0.83-1.00	0.85-1.00	0.83-1.00
Visual analog scale score				
Mean (SD)	81.6 (12.8)	84.5 (11.1)	84.4 (14.6)	84.5 (15.3)
Range	25.0-100.0	40.0-100.0	4.0-100.0	15.0-100.0
IQR	84.0-100.0	90.0-100.0	85.0-100.0	85.0-100.0

Abbreviation: IQR, interquartile range.

**Table 3. Summary of Visual Function Questionnaire–Utility Index Health State Time Trade-Off Utility Scores**

Health State <sup>a</sup>	Participants, No.	Mean (SD)				
		Overall	Australia	Canada	United Kingdom	United States
E: 111111 <sup>b</sup>	606	0.956 (0.124)	0.954 (0.155)	0.989 (0.028)	0.916 (0.168)	0.964 (0.079)
G: 211111	607	0.906 (0.147)	0.901 (0.148)	0.955 (0.100)	0.851 (0.189)	0.917 (0.117)
I: 312211	604	0.828 (0.173)	0.829 (0.183)	0.856 (0.152)	0.795 (0.185)	0.832 (0.166)
K: 323322	607	0.750 (0.212)	0.760 (0.212)	0.761 (0.194)	0.717 (0.231)	0.762 (0.209)
M: 333322	606	0.717 (0.215)	0.721 (0.198)	0.721 (0.218)	0.687 (0.220)	0.737 (0.220)
O: 433354	606	0.550 (0.283)	0.590 (0.255)	0.508 (0.260)	0.534 (0.302)	0.567 (0.307)
R: 444354	607	0.402 (0.312)	0.446 (0.261)	0.365 (0.290)	0.378 (0.335)	0.418 (0.350)
T: 555555 <sup>c</sup>	607	0.343 (0.395)	0.318 (0.364)	0.377 (0.360)	0.264 (0.401)	0.413 (0.436)

<sup>a</sup> See eAppendix 1 in Supplement for full health state descriptions.

<sup>b</sup> Best health state.

<sup>c</sup> Worst health state.

Mean EQ-5D single index scores were similar for all 4 countries; however, there were differences in ranges, with the United Kingdom showing the smallest range (0.52-1.0) (Table 2). Mean EQ-5D visual analog scale scores were comparable across the 4 countries and, as expected, were lower than the index scores.<sup>40</sup> All rank-order scores of the 8 health states completed by each participant were in the expected order of no difficulty (111111) to dead (data not shown).

**Health State Values**

In all, 4850 health state valuations were elicited (8 health states for each of the 607 participants), with only 6 valuations that either were not completed or could not achieve the indifference point. Descriptive statistics on values by health states after adjustment onto the scale of full health to dead are shown in Table 3. Significant country-specific differences were observed for 2 of the 8 health states (Table 3). Differences in mean utilities were seen for health state E and health state G (both  $P < .001$ ). For both health states, UK participants reported the lowest scores and Canadian participants provided the highest scores—the maximum differences were 0.05 and 0.15, respectively. Mean (SD) adjusted health state utilities ranged from 0.343 (0.395) to 0.956 (0.124). The United Kingdom had the widest range of mean health state utilities but the lowest valuation for the best health state (111111) (range, 0.264-0.916). Canadian participant valuations resulted in a utility score of 0.377 for the worst health state (555555) but offered the highest score for the best health state 111111 (0.989). Australian utilities ranged

from 0.318 to 0.954, and the US utilities ranged from 0.413 to 0.964. These preferences showed a linear decline in utility values as the health states increased in severity. Most participants were willing to trade time for health, and the health state utilities were well distributed with little skewness. Results indicate that the 8 evaluated states from the VFQ-UI describe states that participants viewed as spanning a large proportion of the continuum from full health to dead.

**Regression Modeling**

Regression models were used to generate algorithms to estimate preference (ie, utility) scores for the range of potential health states defined by the VFQ-UI health state classification. Items 18, 20, and 25 of the NEI VFQ-25 were recoded so their scoring would be in the same direction as items 6, 11, and 14. The patient data (data set from the VFQ-UI data used in developing the health state classification from patients with central and peripheral vision loss) were analyzed using an IRT graded response model.<sup>37</sup> All items fit the graded response model (Table 4), except item 25 (I worry about doing things that may embarrass myself or others) (data not shown). However, given the large sample size (approximately 3000), it is not surprising that 1 item showed some evidence of mis-fit.<sup>41</sup>

We then produced  $\theta$  scores for each level of each dimension based on this IRT analysis. Level 5, the worst response category (stopped doing because of eyesight, all of the time, or definitely true), was used as the reference category. The  $\theta$  scores are consistent with the ordinality of the VFQ-UI items.

**Table 4. Item Parameters and Fit Statistics for 6-Item Visual Function Questionnaire–Utility Index for Item Response Theory Analysis of Combined Data**

Item No.	Slope, a <sup>a</sup>	Category Threshold <sup>a</sup>				Item Fit Statistics <sup>b</sup>		df
		b1	b2	b3	b4	Likelihood Ratio G <sup>2</sup> (P Value)	Pearson $\chi^2$ (P Value)	
6 <sup>c</sup>	2.51	-0.36	0.67	1.54	2.31	84.15 (.002)	78.59 (.006)	50
11 <sup>d</sup>	2.86	0.71	1.35	2.02	2.90	62.93 (.06)	63.33 (.06)	47
14 <sup>e</sup>	3.46	0.52	1.09	1.55	2.08	80.59 (.003)	80.34 (.003)	49
18 <sup>f</sup>	2.31	0.29	0.94	1.75	2.37	80.81 (.01)	80.36 (.02)	55
20 <sup>g</sup>	3.12	0.92	1.43	1.57	2.18	60.66 (.12)	58.36 (.17)	49
25 <sup>h</sup>	2.79	0.88	1.44	1.61	2.16	102.4 (<.001)	110.3 (<.001)	52

<sup>a</sup> Using item response theory software MULTILOG.<sup>38</sup>

<sup>b</sup> Using SAS macro IRTFIT.<sup>39</sup>

<sup>c</sup> See well up close.

<sup>d</sup> See people's reaction to things I say.

<sup>e</sup> Going out for films, sports event.

<sup>f</sup> Limited work time owing to vision.

<sup>g</sup> Stay at home because of vision.

<sup>h</sup> Worry about doing things that may embarrass because of vision.

**Table 5. Regression Models for Estimating Health Preference Scores**

	Model <sup>a</sup>			
	1	2	3	4
Intercept	0.8336 <sup>b</sup>	0.8740 <sup>b</sup>	0.8656 <sup>b</sup>	0.8841 <sup>b</sup>
Age	0.0009 <sup>b</sup>	0.0009 <sup>b</sup>	0.0009 <sup>b</sup>	0.0009 <sup>b</sup>
$\theta$	-0.1667 <sup>b</sup>	-0.1062 <sup>b</sup>	-0.0966 <sup>b</sup>	-0.0867 <sup>b</sup>
$\theta^2$	-0.00456 <sup>c</sup>	-0.1122 <sup>b</sup>	-0.0505 <sup>b</sup>	-0.0853 <sup>b</sup>
$\theta^3$	NA	0.0278 <sup>b</sup>	NA	0.0114 <sup>d</sup>
Participants, No.	4842	4842	4842	4842
Adjusted R <sup>2</sup>	0.4065	0.4175	0.3992	0.4002

Abbreviation: NA, not applicable.

<sup>a</sup> Model 1 uses age, predicted  $\theta$  based on dimensions (continuous variable), and  $\theta^2$ ; model 2 uses age, predicted  $\theta$  based on dimensions (continuous variable),  $\theta^2$ , and  $\theta^3$ ; model 3 uses age, predicted  $\theta$  based on dimensions (class variable), and  $\theta^2$ ; and model 4 uses age, predicted  $\theta$  based on dimensions

(class variable),  $\theta^2$ , and  $\theta^3$ .

<sup>b</sup>  $P < .001$ .

<sup>c</sup>  $P < .05$ .

<sup>d</sup>  $P < .01$ .

Regression analyses mapping the relationship between the IRT severity score ( $\theta$ ) and TTO preference scores for the 8 health states are presented in **Table 5**. Education and sex were dropped because they were not significant in any regression model. The simplest regression model, model 1, included age,  $\theta$  as a continuous variable, and  $\theta^2$ . The  $\theta^2$  term was not significant, so model 2 was run with a  $\theta^3$  term. All of the variables were significant in the model with a  $\theta^3$  term. These 2 models were re-run with the predicted  $\theta$  term included as a class variable. For both models, all terms were significant. The adjusted R<sup>2</sup> values were all similar, ranging from 0.3992 to 0.4175, and other model fit statistics showed that these regression models had a good fit to the observed data.

**Validity of VFQ-UI Scores**

Assessment of the validity of the VFQ-UI outside the development sample was performed using data from a recent interventional study that administered the NEI VFQ-25 to patients with uveitis. A detailed description of the clinical trial and key inclusion and exclusion criteria have been reported elsewhere.<sup>42</sup> Briefly, patients included in the trial were at least 18 years old, were diagnosed as having intermediate or posterior uveitis, and had decreased VA attributable to uveitis, with

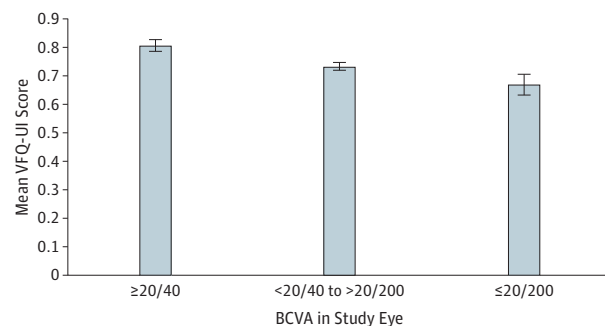
a vitreous haze score at least +1.5 and BCVA between 10 and 75 letters in the study eye measured by the Early Treatment Diabetic Retinopathy Study method at baseline. Patients were excluded if they had uncontrolled systemic disease, any ocular condition in the study eye that would prevent improvement in VA, or any condition or treatment that would otherwise confound the results of the study. The mean age of the sample was 44.6 years, 65% were female, and most patients (61%) were white. The median VA in the study eye was 62.0 letters, and 84% received treatment in their worse-seeing eye.

Mean VFQ-UI scores significantly discriminated among study-eye BCVA groups ( $\geq 20/40$ ,  $< 20/40$  to  $> 20/200$ , and  $\leq 20/200$ ) at baseline ( $P < .05$ ) (**Figure**).<sup>43</sup> The VFQ-UI scores significantly correlated with NEI VFQ-25 composite scores ( $r = 0.92$ ) and subscale scores (ranging from a low of  $r = 0.42$  for general health to  $r = 0.85$  for social functioning).

**Discussion**

We developed a VFQ-UI algorithm based on the VFQ-UI health state classification that will enable estimation of preference scores using responses to 6 items of the NEI VFQ-25. As a re-

**Figure. Visual Function Questionnaire–Utility Index Validity of Known Groups From a Recent Interventional Uveitis Study<sup>42,a</sup>**



BCVA indicates best-corrected visual acuity; VFQ-UI, Visual Function Questionnaire–Utility Index; and error bars, standard error of the mean.

<sup>a</sup>The overall nonstudy eye mean (SD) Early Treatment Diabetic Retinopathy Study BCVA at baseline across all subjects was 76.9 (14.74).

sult, preference scores can be estimated when these 6 items from the NEI VFQ-25 have been completed. The final regression models took into account the patient data used in development of the VFQ-UI health state classification.<sup>23</sup> Using patient-level data and valuation study data, we mapped the relationship between  $\theta$  scores and preference scores for VFQ-UI health states. For future research applications using the VFQ-UI, we recommend using the model 2 regression equation because it is parsimonious and yielded the largest adjusted  $R^2$  value (eAppendix 3, Appendix 4, eTable 2, and eTable 3 in Supplement).

This study was conducted to obtain health preferences for VFQ-UI health states in members of the general population in Australia, Canada, the United Kingdom, and the United States. Because the resources allocated for different health technologies come from the public in Australia, Canada, and the United Kingdom, it is argued that the preferences should come from the public.<sup>3-5,7</sup> In addition, societal preferences are important because generic health preference measures are based on societal data, and it is important to have comparable sources of data for these disease-targeted states. Furthermore, several health technology assessment agencies require preferences based on the general population.<sup>8</sup>

The analytical method used in this study is different from the method used by Brazier and colleagues in developing health state classifications based on the 36-Item Short Form Health Survey, Asthma Quality of Life Questionnaire, and other disease-targeted measures.<sup>6,10,18,22</sup> Our data could not be analyzed using their approach owing to the lack of independence of the selected vision-related dimensions. The approach estimates the  $\theta$  scores for the target health states using IRT analysis and then models the relationship between these  $\theta$  scores and the health preference scores. This is comparable to methods used by Brazier et al<sup>9</sup> and Young et al<sup>35</sup> to develop a flushing-specific preference score.

Using data from a separate clinical trial of patients with noninfectious uveitis, we found preliminary support for the validity of the VFQ-UI. The VFQ-UI scores highly correlated with NEI VFQ-25 composite scores and moderately to highly

correlated with subscale scores. The VFQ-UI scores varied significantly by BCVA groups, with patients with more impaired VA reporting greater impairment in vision-specific preference scores.

One limitation of the valuation survey is that participants were required to attend an interview and therefore were likely healthier than the general population. The survey participants were healthier and more educated than the general population in these countries. Standard deviations for VFQ-UI scores were smaller overall and by country for better visual functioning health states. Standard deviations increased with worsening health states, suggesting that the general population has a broader range of preferences and willingness to trade off years as visual functioning decreases. The results indicating a high valuation are consistent with public opinion polls in the United States that show a high fear of going blind and a high value for vision itself.<sup>44</sup> Additionally, although this was a large sample from the general public, collecting utilities from patients who experience vision-related conditions could also be a useful comparison. Further research is needed to compare patient and general public preferences for different vision-related states. Previous research has demonstrated that patients and clinicians have different preferences for vision-related impairment states.<sup>45,46</sup>

Second, we observed some differences between countries on 2 of the TTO utilities. However, the rank ordering of the utilities within countries was identical, and the differences may only seem to reflect more optimistic (or pessimistic) valuations across countries. These findings may be attributable to cultural differences among the countries. Individuals were all from Western countries, and there may be differences in valuing vision-related health states among Asian, African, North American, and European countries that may affect VFQ-UI scores. Additional research is needed to examine differences in preferences in Asian and other cultures.

Areas for further research include assessing the reliability and validity, including responsiveness to change of utility scores from the VFQ-UI in both central and peripheral vision loss populations as well as for various vision-defined health states used in economic analyses. As the NEI VFQ-25 and the VFQ-UI are completed in terms of vision in both eyes, the VFQ-UI captures utility associated with bilateral vision. Future research can assess utility considering bilateral or monocular conditions and whether a study eye is the better- or worse-seeing eye. In addition, research is needed to assess VFQ-UI estimated health state utility scores and measurement properties compared with generic preference measures such as the EQ-5D, Health Utilities Index Mark 2, Health Utilities Index Mark 3, SF-6D, and/or the Quality of Well-being Scale.<sup>47</sup> A recent study demonstrated that the VFQ-UI was more responsive than generic health preference measures in patients undergoing cataract surgery.<sup>47</sup>

## Conclusions

We developed an algorithm for converting VFQ-UI scores into health preferences for use in economic evaluations. These vision-related preference scores are expected to be more respon-



sive to differences among the effects of ophthalmologic interventions than generic health preference measures. The VFQ-UI represents the patient's perspective on the impact of ocular conditions on functioning and well-being, and VFQ-UI scores

allow for comparisons across eye disorders. The VFQ-UI health preference scores for different vision-related treatments may be of value to estimate QALYs for economic evaluations and health policy decisions.

#### ARTICLE INFORMATION

**Submitted for Publication:** April 18, 2013; final revision received July 12, 2013; accepted July 26, 2013.

**Published Online:** January 16, 2014.

doi:10.1001/jamaophthalmol.2013.7639.

**Author Contributions:** Mss Rentz and Yu had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Rentz, Kowalski, Walt, Hays, Lee, Bressler, Revicki.

**Acquisition of data:** Rentz, Kowalski.

**Analysis and interpretation of data:** All authors.

**Drafting of the manuscript:** Rentz, Kowalski, Hays, Brazier, Revicki.

**Critical revision of the manuscript for important**

**intellectual content:** Kowalski, Walt, Hays, Yu, Lee, Bressler, Revicki.

**Statistical analysis:** Rentz, Kowalski, Hays, Brazier, Yu, Revicki.

**Obtained funding:** Rentz, Kowalski, Walt, Lee, Revicki.

**Administrative, technical, and material support:** Kowalski.

**Study supervision:** Walt, Revicki.

**Conflict of Interest Disclosures:** Mss Rentz and Yu and Dr Revicki are employed by Evidera (United BioSource Corp at the time of the study), which provides consulting and other research services to pharmaceutical, device, government, and nongovernment organizations; as Evidera employees, they work with a variety of companies and organizations and are expressly prohibited from receiving any payment or honoraria directly from these organizations for services rendered. Dr Kowalski and Mr Walt are employees of Allergan, Inc. Dr Hays received financial support for this study from Allergan, Inc. Dr Brazier had received payment from Allergan, Inc, as a consultant for this study. Dr Lee is currently a consultant for Pfizer, Genentech, Quorum Consulting, and Novartis and was a consultant for Allergan, Inc, from March 2010 to March 2011; has in the last 3 years received research grants from Pfizer; and has a proprietary interest in Merck, GlaxoSmithKline, and Vitaspring Health Technologies. Dr Bressler's employer, the Johns Hopkins University (but not Dr Bressler), received funding from Allergan, Inc, for this project; the terms of this project are negotiated and administered by Johns Hopkins University's Office of Research Administration. Under Johns Hopkins University's policy, support for the costs of research, administered by the institution, does not constitute a financial conflict of interest. No other disclosures were reported.

**Funding/Support:** This work was supported by Allergan, Inc. Dr Hays was supported in part by grants P30AG021684 from the National Institute on Aging to the University of California, Los Angeles (UCLA) Resource Center for Minority Aging Research Center for Health Improvement in Minority Elderly, 2P20MD000182 from the National Center on Minority Health and Health Disparities to the UCLA/Drew Project EXPORT, and

P30AG028748 from the National Institute on Aging to the UCLA Older Americans Independence Center.

**Role of the Sponsor:** Allergan, Inc, was involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The other funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Disclaimer:** Dr Bressler is the *JAMA Ophthalmology* Editor but was not involved in the review process or the acceptance of the manuscript.

**Additional Contributions:** We thank the following researchers, their staff, and patients for their support and participation in piloting and cognitive debriefing of the 8 health states prior to the country-level valuation surveys: David S. Boyer, MD, Tammy Gasparian, Lissette Pleitez, Retina-Vitreous Associates Medical Group, Beverly Hills, California; Jeffrey C. Hong, MD, Marie Rodriguez, Huntington Eye Medical Group, Pasadena, California; and Steven Simmons, MD, Martin Kaback, MD, and Ralph Sanchez, MD, Glaucoma Consultants of the Capital Region, Slingerlands, New York. Beenish Nafees, MSc, Sean O'Quinn, MPH, Anne Brooks, BS, Scott Doyle, DPhil, PhD, Sarah Hearn, MSc, Pamela Murray, MSc, Alise Nacson, MPH, Sherilyn Notte, BS, Paul O'Donohoe, BSc, Megan Stafford, BSc, and Randall Winnette, MSc, provided assistance with this study. Julie Meilak and Aria Gray helped with proofreading and formatting the manuscript.

#### REFERENCES

1. Commonwealth Department of Health, Housing, and Community Services. *Guidelines for the Pharmaceutical Industry on the Submission to the Pharmaceutical Benefits Advisory Committee*. Canberra, Australia: Australian Government Publishing Service; 1992.
2. National Institute of Health and Clinical Excellence. Guide to the methods of technology appraisal, 2004. <http://www.nice.org.uk>. Accessed January 9, 2009.
3. Gold M, Franks P, Erickson P. Assessing the health of the nation: the predictive validity of a preference-based measure and self-rated health. *Med Care*. 1996;34(2):163-177.
4. Lipscomb J, Drummond M, Fryback D, Gold M, Revicki D. Retaining, and enhancing, the QALY. *Value Health*. 2009;12(suppl 1):S18-S26.
5. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value Health*. 2009;12(suppl 1):S5-S9.
6. Yang Y, Brazier JE, Tsuchiya A, Young TA. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the Asthma Quality of Life Questionnaire. *Med Decis Making*. 2011;31(2):281-291.

7. Petrillo J, Cairns J. Converting condition-specific measures into preference-based outcomes for use in economic evaluation. *Expert Rev Pharmacoecon Outcomes Res*. 2008;8(5):453-461.

8. National Institute of Health and Clinical Excellence. Guide to the methods of technology appraisal, 2008. <http://www.nice.org.uk>. Accessed January 28, 2009.

9. 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.

10. Rowen D, Brazier J, Young T, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health*. 2011;14(5):721-731.

11. Bansback N, Marra C, Tsuchiya A, et al. Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Rheum*. 2007;57(6):963-971.

12. Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. *Value Health*. 2007;10(3):214-220.

13. Lenert LA, Sherbourne CD, Sugar C, Wells KB. Estimation of utilities for the effects of depression from the SF-12. *Med Care*. 2000;38(7):763-770.

14. Lo PS, Tong MC, Revicki DA, et al. Rhinitis Symptom Utility Index (RSUI) in Chinese subjects: a multiattribute patient-preference approach. *Qual Life Res*. 2006;15(5):877-887.

15. Revicki DA, Leidy NK, Brennan-Diemer F, Sorensen S, Toggias A. Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest*. 1998;114(4):998-1007.

16. Revicki DA, Leidy NK, Brennan-Diemer F, Thompson C, Toggias A. Development and preliminary validation of the multiattribute Rhinitis Symptom Utility Index. *Qual Life Res*. 1998;7(8):693-702.

17. Espallargues M, Czoski-Murray CJ, Bansback NJ, et al. The impact of age-related macular degeneration on health status utility values. *Invest Ophthalmol Vis Sci*. 2005;46(11):4016-4023.

18. 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.

19. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD; NEI-VFQ Field Test Investigators. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). *Arch Ophthalmol*. 1998;116(11):1496-1504.

20. Mangione CM. NEI VFQ-25 scoring algorithm. [http://www.nei.nih.gov/resources/visionfunction/manual\\_cm2000.pdf](http://www.nei.nih.gov/resources/visionfunction/manual_cm2000.pdf). Accessed December 5, 2012.

21. Yang Y, Brazier J, Tsuchiya A, Coyne K. Estimating a preference-based single index from the Overactive Bladder Questionnaire. *Value Health*. 2009;12(1):159-166.
22. 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.
23. Kowalski JW, Rentz AM, Walt JG, et al. Rasch analysis in the development of a simplified version of the National Eye Institute Visual-Function Questionnaire-25 for utility estimation. *Qual Life Res*. 2012;21(2):323-334.
24. Haller JA, Bandello F, Belfort R Jr, et al; Ozurdex GENEVA Study Group. Dexamethasone intravitreal implant in patients with macular edema related to branch or central retinal vein occlusion twelve-month study results. *Ophthalmology*. 2011;118(12):2453-2460.
25. Callanan DG, Gupta S, Boyer DS, et al; Ozurdex PLACID Study Group. Dexamethasone intravitreal implant in combination with laser photocoagulation for the treatment of diffuse diabetic macular edema. *Ophthalmology*. 2013;120(9):1843-1851.
26. Dong LM, Childs AL, Mangione CM, et al; Submacular Surgery Trials Research Group. Health- and vision-related quality of life among patients with choroidal neovascularization secondary to age-related macular degeneration at enrollment in randomized trials of submacular surgery: SST report No. 4. *Am J Ophthalmol*. 2004;138(1):91-108.
27. Schiffman RM, Walt JG, Jacobsen G, Doyle JJ, Lebovics G, Sumner W. Utility assessment among patients with dry eye disease. *Ophthalmology*. 2003;110(7):1412-1419.
28. Globe DR, Wu J, Azen SP, Varma R; Los Angeles Latino Eye Study Group. The impact of visual impairment on self-reported visual functioning in Latinos: The Los Angeles Latino Eye Study. *Ophthalmology*. 2004;111(6):1141-1149.
29. Vasudevan SK, Gupta V, Crowston JG. Neuroprotection in glaucoma. *Indian J Ophthalmol*. 2011;59(suppl):S102-S113.
30. Muir KW, Santiago-Turla C, Stinnett SS, et al. Health literacy and adherence to glaucoma therapy. *Am J Ophthalmol*. 2006;142(2):223-226.
31. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. London, England: Psychology Press; 2007.
32. Wright BD, Masters GN. *Rating Scale Analysis: Rasch Measurement*. Chicago, IL: MESA Press; 1982.
33. Furlong W, Feeny D, Torrance GW, Barr R, Horsman J. *Guide to Design and Development of Health-State Utility Instrumentation*. Hamilton, ON: Centre for Health Economics & Policy Analysis, McMaster University; 1990.
34. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
35. Young TA, Rowen D, Norquist J, Brazier JE. Developing preference-based health measures: using Rasch analysis to generate health state values. *Qual Life Res*. 2010;19(6):907-917.
36. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ*. 2010;11(2):215-225.
37. Samejima F. Graded response model. In: van der Linden WJ, Hambleton RK, eds. *Handbook of Modern Item Response Theory*. New York, NY: Springer; 1997.
38. Thissen D, Chen WH, Bock RD. *MULTILOG* [computer program]. Version 7. Lincolnwood, IL: Scientific Software International; 2003.
39. Bjorner JB, Smith KJ, Orlando M, Stone C, Thissen D, Sun X. *IRTFIT: A Macro for Item Fit and Local Dependence Tests Under IRT Models*. Lincoln, RI: QualityMetric; 2006.
40. Brooks R, Rabin R, de Charro F. *The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective*. Dordrecht, the Netherlands: Kluwer; 2003.
41. Hambleton RK, Swaminathan H, Rogers HJ. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage; 1991.
42. Lowder C, Belfort R Jr, Lightman S, et al; Ozurdex HURON Study Group. Dexamethasone intravitreal implant for noninfectious intermediate or posterior uveitis. *Arch Ophthalmol*. 2011;129(5):545-553.
43. Naik RK, Rentz AM, Foster CS, et al. Normative comparison of patient-reported outcomes in patients with noninfectious uveitis. *JAMA Ophthalmol*. 2013;131(2):219-225.
44. American Foundation for the Blind. Attitudes about blindness/severe vision loss. <http://www.afb.org>. Accessed December 10, 2010.
45. Kymes SM, Lee BS. Preference-based quality of life measures in people with visual impairment. *Optom Vis Sci*. 2007;84(8):809-816.
46. Stein JD, Brown MM, Brown GC, Hollands H, Sharma S. Quality of life with macular degeneration: perceptions of patients, clinicians, and community members. *Br J Ophthalmol*. 2003;87(1):8-12.
47. Feeny D, Spritzer K, Hays RD, et al. Agreement about identifying patients who change over time: cautionary results in cataract and heart failure patients. *Med Decis Making*. 2012;32(2):273-286.