AMERICAN ACADEMY OF OPHTHALMOLOGY®

# Measuring Quality of Life in Diabetic Retinal Disease: A Narrative Review of Available Patient-Reported Outcome Measures

*Stela Vujosevic, MD, PhD,*[1,2] *Emily Chew, MD,*[3] *Leanne Labriola, DO, MBA,*[4,5] *Sobha Sivaprasad, MD,*[6] *Ecosse Lamoureux, PhD*[7,8]

***Topic:*** Several patient-reported outcome measures (PROMs) are available to measure health-related quality of life (HRQoL) in patients with late-stage clinical diabetic retinal diseases (DRDs). However, an understanding of the psychometric properties of PROMs is needed to assess how they could relate to severity levels of a revised DRD grading system. This narrative review assessed the available generic-, vision-, and DRD-related PROMs used in DRD research and highlights areas for improvement.

***Clinical Relevance:*** Diabetic retinal disease is a common complication of diabetes and can lead to sight-threatening complications with a devastating effect on HRQoL.

***Methods:*** The Quality of Life working group is one of 6 working groups organized for the DRD Staging System Update Effort, a project of the Juvenile Diabetes Research Foundation Mary Tyler Moore Vision Initiative. PubMed, Cochrane Library, Embase, and Google Scholar databases were searched using core keywords to retrieve ophthalmology-related review articles, randomized clinical trials, and prospective, observational, and cross-sectional studies in the English language. A detailed review of 12 PROMs (4 QoL questionnaires and 8 utilities) that met a minimum level of evidence (LOE) was conducted. The relevance of each PROM to DRD disease stage and Biomarker Qualification guidelines (*Biomarkers, EndpointS, and other Tools*) categories was also defined.

***Results:*** The National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25), Impact of vision impairment-computerized adaptive testing, and Diabetic Retinopathy and Macular Edema Computerized Adaptive Testing System had a LOE of II in detecting change due to late-stage DRD (diabetic macular edema), although several areas for improvement (e.g., psychometrics and generalizability) were identified. Other PROMs, particularly the utilities, had a LOE of III due to cross-sectional evidence in late-stage clinical DRD. Although the NEI VFQ-25 has been the most widely used PROM in late-stage DRD, more work is required to improve its multidimensional structure and other psychometric limitations. No PROM was deemed relevant for subclinical or early/mid-DRD.

***Conclusion:*** This narrative review found that the most commonly used PROM is NEI VFQ-25, but none meets the ideal psychometric, responsiveness, and clinical setting digital administration requirements that could be included in an updated DRD staging system for diagnosis and monitoring of DRD progression.

***Financial Disclosure(s):*** Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2024;4:100378* 

*Supplemental material available at www.ophthalmologyscience.org.*

Diabetic retinal disease (DRD) is a major public health burden, with clinical DRD defined as visible vascular lesions of diabetic retinopathy (DR), affecting nearly one-third of people with diabetes.[1] Classification of DRD is currently based on severity levels of clinical DR. Approximately 10% of people with DR have vision-threatening disease, specifically diabetic macular edema (DME) or proliferative diabetic retinopathy (PDR) that comprises late-stage clinical DRD. However, the current DRD classification system does not take into account the patient's quality of life (QoL).[2,3]

Patient-reported outcome measures (PROMs) are commonly used to measure health-related QoL (HRQoL). They can be generic (i.e., general health), symptom specific (e.g., vision), or disease-specific (e.g., DRD) and may be delivered as questionnaires or utility instruments (Table S1, available at www.ophthalmologyscience.org.).[4] *Questionnaires* comprise a set of items (questions) whose scores can be summed or averaged to provide measurement of a latent construct (e.g., HRQoL). The overall score is typically used as a study end point or outcome measure. *Utilities* reflect the health status of a patient and the



https://doi.org/10.1016/j.xops.2023.100378
ISSN 2666-9145/23

associated value of that health status through the patient's preferences. Preferences can be derived directly from patient ratings of health-related hypothetical scenarios or indirectly by patients rating their health status from a multi-attribute classification system (similar to a questionnaire). Utilities can be used to generate the quality-adjusted life-years (QALYs) that combines both the quantitative (length of life) and the qualitative (QoL) into a single index using a utility score, as in the equation, (years of life × utility value = QoL) years.[5,6]

The inclusion of PROMs in clinical trials is now mandated by regulatory authorities, such as the Food and Drug Administration (FDA) and the European Medicine Agency.[7] Moreover, PROMs can guide clinical care, support patient−provider decision making, provide data for comparative effectiveness research for practice improvement or assessment of performance, and assess metrics for value-based payments.[8] Because of the importance of these parameters in health care, a better understanding of the context of use and psychometric properties of current and emerging DRD-associated PROMs is required to guide public health decisions and policies[9] and clinical trial end point measurements.[10] The aims of this paper were as follows: (1) to conduct a narrative review of the context of use, appropriateness, and psychometric properties of current and emerging HRQoL PROMs used in DRD research; and (2) to highlight gaps in knowledge that may serve as potential guidance for future work to develop and validate DRD-specific PROMs.

## Methods

This narrative review was undertaken by the Quality of Life working group, one of 6 working groups organized for the DRD Staging System Update Effort, a project of the Juvenile Diabetes Research Foundation Mary Tyler Moore Vision Initiative supported by The Mary Tyler Moore and S. Robert Levine, MD Charitable Foundation (see Appendix X for a complete listing of Initiative leadership and participants, available at www.ophthalmologyscience.org). The researchers in the Quality of Life working group were led by Stela Vujosevic (Italy) and other members included Emily Chew (United States), Leanne Labriola (United States), Ecosse Lamoureux (Singapore), and Sobha Sivaprasad (United Kingdom).

### Literature Search

Four databases (PubMed, Cochrane Library, Embase, and Google Scholar) were searched for QoL PROMs that have been used in DRD research using the following keywords: *"Diabetic retinopathy" AND "Diabetic Macular Edema/Oedema" AND "Quality of life" AND "Functioning"; "Utility," "Time Trade Off," "Standard Gamble," "Direct Elicitation," "Health Related Quality of Life," "ED-BOLT ON," AND "RetCAT"*. The search was completed in April 2022, and ophthalmology-related review articles, randomized clinical trials (RCTs), and prospective, observational, and cross-sectional studies were included. The review followed all 6 items according to the Scale for the Assessment of Narrative Review Articles (SANRA), an accepted critical appraisal tool to assess the quality of nonsystematic review articles.[11] A detailed review was completed for 12 PROMs that met the key criteria of at least level III evidence (LOE) (Table 2). These were: European Quality of Life Questionnaire (EQ-5D) and EQ-5D/visual analog scale (VAS); National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25); Impact of Vision Impairment (IVI) questionnaire; Diabetic Retinopathy Dependent Quality of Life (RetDQoL); Diabetic Retinopathy and Macular Edema Computerized Adaptive Testing System (RetCAT); Short form-6D (SF-6D); the Health Utilities Index (HUI3); Vision and Quality of Life Index (VisQoL); Time trade-off (TTO); Standard Gamble (SG); and Diabetic Retinopathy Utility instrument (DR-U) (Table 2).

These PROMs' parameter characteristics, scientific understanding of their relationship to DRD, performance expectations in regard to DRD, evidence level, statistical data, gap analysis, and additional information, such as availability and applicability (detailed in Table S3, available at www.ophthalmologyscience.org) were reviewed in a standardized manner according to the FDA Biomarker Qualification guidelines (*Biomarkers, EndpointS, and other Tools,* FDA 2020).[12]

This multistep review process started with an initial review by a single researcher, then proceeded with a collaborative review by the working group that met monthly between November 2020 and June 2021 and thereafter 4 more times (12 times in total) and also communicated by email until July 2022. Thereafter, the working group lead synthesized the results and formulated conclusions, which were sent to the Steering Committee.

The PROMs were classified as "Ready" (for current use or use within 2 years), "Promising" (unmet, but defined research needs that can be accomplished within 2−5 years), and "Potential" (unmet research needs that will need > 5 years to accomplish) based on time frame of anticipated validation for clinical or research use (Table 4) based on subject matter expertise and consensus agreement among all 5 members of the working group. Moreover, the relevance of the 12 PROMs to disease stage (preclinical, early, mid, and late-stage DRD) (Table 4) was also determined. Late-stage DRD was defined as PDR and/or DME. The sensitivity of each of the 12 PROMs to identify progression of DRD, responsiveness of DME and/or PDR to treatment, and visual acuity change were also considered.

Finally, the potential relevance of the 12 PROMs was classified according to the FDA Biomarker Qualification guidelines (*Biomarkers, EndpointS, and other Tools*, FDA 2020)[12] (Table 5).

Conclusions about the LOE rating, psychometric performance, responsiveness, appropriateness, and generalizability of each PROM to detect change in DRD severity levels and treatment response informed areas for future research that are required for an ideal PROM.

## Results

### Questionnaires (Vision-Specific)

**NEI VFQ-25.** *Parameter Characteristics.* The NEI VFQ-25 elicits patient perceptions of their visual impairment (VI) and its relation to HRQoL.[13] Content for the questionnaire was developed from patients with DR, age-related macular degeneration, cataract, primary open-angle glaucoma, cytomegalovirus retinitis, and low vision.[13] The NEI VFQ-25 includes 1 general health item and 11 vision-specific domains, including general vision, ocular pain, near vision, distance vision, social function, mental health, role limitations, dependency, driving, color vision, and peripheral vision. Composite and subscale scores range from 0 to 100, where higher scores indicate better

Table 2. Currently Available Level of Evidence for the 12 Reviewed PROMs

| I | II | III | IV-V |
|---|---|---|---|
| NEI VFQ-25 | Substantial evidence in RCTs and population-based studies for late-stage clinical DRD | | |
| IVI | Evidence in clinical and population-based samples | | |
| RetDQoL | | CS with mostly clinical samples | |
| DR-U | | CS in a large clinical sample | |
| RetCAT | CS in a large clinical sample | | |
| EQ-5D/VAS | | CS with mostly clinical samples | |
| SF-6D | | CS with mostly clinical samples | |
| HUI3 | | CS with mostly clinical samples | |
| VisQoL | | CS with mostly clinical samples | |
| TTO/SG | | CS with mostly clinical samples | |

CS = cross-sectional; DRD = diabetic retinal disease; DR-U = Diabetic Retinopathy Utility instrument; EQ-5D = European Quality of Life Questionnaire; HUI-3 = Health Utilities Index; IVI = Impact of Vision Impairment; NEI VFQ-25 = National Eye Institute 25-item Visual Function Questionnaire; PROM = patient-reported outcome measure; RCT = randomized controlled trial; RetCAT = Diabetic Retinopathy and Macular Edema Computerized Adaptive Test; RetDQoL = Retinopathy Dependent Quality of Life questionnaire; SF-6D = Short Form 6-D; SG = standard gamble; TTO = time trade-off; VAS = visual analog scale; VisQoL = Vision and Quality of Life index.

functioning or well-being.[14] Although minimally important differences that correlate with severity level of clinical DR have not yet been established for the NEI VFQ-25, a $\geq$ 6-point change in composite or subscale scores has been reported to reflect a clinically meaningful change in DME.[15] The reliability and validity of the NEI VFQ-25 has been demonstrated in a variety of eye conditions, including DRD (LOE II). For more information on NEI VFQ-25, refer to Supplementary Materials.

*Scientific Understanding and Performance Expectations in DRD.* The NEI VFQ-25 has shown consistent correlation with visual acuity.[16] For example, in the Wisconsin Epidemiology study of Type 1 Diabetes, NEI VFQ-25 scores were strongly associated with vision, independent of severity of DR and other diabetes complications.[17] The NEI VFQ-25 is also associated with severity of DR, with scores decreasing with increasing severity of DR in people with no diabetes-related comorbidities.

*Type of Data and Availability for Evidential Evaluation.* The NEI VFQ-25 has been used in population-based studies, case series, and RCTs, particularly with novel treatments for DME. Although the NEI VFQ-25 scores are appropriate for monitoring of vision-related QoL (VRQoL) because they are associated with vision and DR severity, their sensitivity has not been evaluated for earlier stages of DRD. It can be either self- or interviewer-administered.

*Statistical Considerations.* The NEI VFQ-25 has undergone substantial psychometric testing using classical test theory (e.g., construct validity, criterion validity, dimensionality via factor analysis, etc.).[13] However, Rasch analysis of the NEI VFQ-25 has demonstrated improved psychometric performance and interpretability of the VRQoL scale using a 2-domain structure (visual functioning and socioemotional well-being).[18–20] A more recent evaluation of the Rasch analysis on combined clinical trials and observational data calibrated item measures and rating category thresholds for the NEI VFQ-25 as well as for the 2 domain-specific versions, the NEI VFQ that included the visual function time and the socioemotional items. This has greatly improved the usefulness of the NEI VFQ by enabling the estimation of measures on an invariant scale and comparisons between patients and across studies.[21] A shorter version of NEI VFQ-25 is also being developed.

*Gap Analysis.* Because the NEI VFQ-25 has very few items in each subscale, the results should be interpreted with caution. Moreover, it does not evaluate other aspects of VRQoL, such as functioning under low luminance. It has to be extended to more focus groups (with various severity levels of DR). It also should be considered for change into the electronic format due to the high costs in trials. The NEI VFQ-25 is not currently approved by regulatory bodies as an end point for trials. The NEI VFQ is strongly correlated with visual acuity, but it does not correlate with DRD or systemic disease. The psychometric properties of NEI VFQ-25 using modern psychometric theory has been explored. As such, revision of the NEI VFQ-25 to include more granularity or spread of the scale for each item and also to expand the number of work is still required to optimize its multidimentionality, especially in patients with DRD.[18–20] As such, revision of domains may be required so it can be multidimensional and achieve the standards set out by regulatory bodies, such as the FDA and European Medicine Agency.

*Miscellaneous Questions.* The NEI VFQ-25 is readily available online at no cost, has been translated into multiple languages, and is quick to administer. With strong correlation with visual acuity, it merits further research and development to incorporate into the DRD staging process.

**IVI.** *Parameter Characteristics.* The 28-item IVI measures the impact of VI (from any eye condition) on VRQoL and comprises 3 domains; reading and accessing information, mobility and independence, and emotional well-being.[22–24] Higher scores indicate better VRQoL outcomes and vice versa. The IVI has been extensively validated in clinical and population-based studies (LOE II),[25–28] including patients with DRD, and has excellent psychometric properties (LOE II).

Table 4. Relevance of PROMs to Disease Stage

| | Ready* | Promising† | Potential‡ |
|---|---|---|---|
| Subclinical DRD | | | |
| Early-stage clinical DRD | | | |
| Midstage clinical DRD | | | |
| Late-stage clinical DRD | NEI VFQ-25 | | RetCAT; RetDQoL; EQ-5D/VAS; DR-U |

DRD = diabetic retinal disease; DR-U = Diabetic Retinopathy Utility instrument; EQ-5D = European Quality of Life questionnaire; NEI VFQ-25 = National Eye Institute 25-item Visual Function Questionnaire; PROM = patient-reported outcome measure; RetCAT = Diabetic Retinopathy and Macular Edema Computerized Adaptive Test; RetDQoL = Retinopathy Dependent Quality of Life questionnaire; VAS = visual analog scale.
*For current use or use in < 2 years.
†Unmet, but defined research needs that can be accomplished within 2–5 years.
‡Unmet research needs that will need > 5 years to accomplish.

*Scientific Understanding and Performance Expectations in DRD.* The IVI is correlated with visual acuity, which is reflected in its association with severity of DRD,[25–28] response to treatment,[29] and low-vision rehabilitation.[30]

*Type of Data and Availability for Evidence Evaluation.* In addition to cross-sectional and population-based studies, the IVI has been used in 2 small intervention studies in Australia, including phase II RCTs in DR[29,30] in which IVI scores correlated with improvement in vision after treatment for DME.

*Statistical Considerations.* The IVI provides an overall score of VRQoL in addition to 3 domain scores. Scores can be calculated using simple summary scoring; however, conducting Rasch analysis on the raw data is recommended to convert the summed scores into those approximating interval-level measurement, ready for use in parametric testing such as regression analyses. Using Rasch-scaled IVI scores rather than summed IVI scores may improve measurement precision and increase sensitivity when assessing associations between variables and changes over time.[30] A recent study by Goldstein et al[31] has calibrated IVI item measures from various studies, enabling researchers and clinicians to more easily compare VRQoL in patients with VI.

*Gap Analysis.* Although the IVI has been used in 2 phase II DRD trials to date, its responsiveness to other DRD treatment regimens has not yet been evaluated, and further work in this area is needed. It is nonspecific to DRD and provides measurement of only 3 QoL domains.

*Miscellaneous Questions.* The IVI is available commercially and can be administered in paper–pencil format or on a digital platform; as such, it can be easily implemented in both high and low-resource settings. It can be either self- or interviewer-administered. The IVI has relatively low-patient burden, taking approximately 15 minutes to administer. Shortened versions of the IVI are also available, including the Brief IVI,[32] currently endorsed for use in assessing VRQoL in patients with age-related macular degeneration by the International Consortium for Health Outcomes Measurement and the IVI-CAT,[33] both of which cut administration time by 50% to 80%.

Table 5. Potential Relevance of PROMs to BEST Categories

| | Diagnostic | Monitoring | Predictive | Prognostic | Pharmacodynamic/Response | Safety | Susceptibility/Risk |
|---|---|---|---|---|---|---|---|
| NEI VFQ-25 | | | | | X | | |
| DR-U | | | | | X | | |
| RetDQoL | | | | | | | |
| IVI | | | | | X | | |
| RetCAT | | | | | X | | |
| EQ-5D | | | | | | | |
| EQ-5D/VAS | | | | | | | |
| SF-6D | | | | | | | |
| HUI3 | | | | | | | |
| VisQoL | | | | | | | |
| TTO | | | | | | | |
| SG | | | | | | | |

BEST = *Biomarkers, EndpointS, and other Tools*; DR-U = Diabetic Retinopathy Utility instrument; EQ-5D = European Quality of Life Questionnaire; HUI3 = Health Utilities Index; IVI = Impact of Vision Impairment; NEI VFQ-25 = National Eye Institute 25-item Visual Function Questionnaire; PROM = patient-reported outcome measure; RetCAT = Diabetic Retinopathy and Macular Edema Computerized Adaptive Testing System; RetDQoL = Diabetic Retinopathy Dependent Quality of Life; SF-6D = Short Form-6D; SG = standard gamble; TTO = time trade-off; VAS = visual analog scale; VisQoL = Vision and Quality of Life index.

## Questionnaires (DRD-Specific)

**RetDQoL.** *Parameter Characteristics.* The RetDQoL specifically evaluates the QoL of patients diagnosed with DR and is modeled on the Audit of Diabetes-Dependent QoL.[34] The questionnaire has 2 broad questions related to present QoL and patient perception of QoL if they did not have a diabetes-related eye problem. The remaining 24 items refer to specific aspects of QoL, in which patients are asked to both evaluate the impact of the domain on the QoL and how important each domain is to their life. The RetDQoL has demonstrated high internal consistency and good construct validity (LOE III).

*Scientific Understanding and Performance Expectations in DRD.* The RetDQoL has been shown to be sensitive to different levels of DRD-induced VI and late-stage DRD as the expected relationships were found for both the overview items and the average weighted index.[35]

*Type of Data and Availability for Evidential Evaluation.* The reported use of RetDQoL in clinical practice, population-based studies, or clinical trials to assess its correlation to DR severity level or treatment outcome has been limited.[36,37]

*Statistical Considerations.* The multiplicative scoring system employed by the RetDQoL is problematic,[38] as shown in previous work utilizing Rasch analysis in similar instruments (the Macular Disease Quality of Life questionnaire[39]); the majority of response categories were underutilized and category thresholds were disordered.

*Gap Analysis.* The psychometric properties of the RetDQoL have not been explored using modern psychometric theory, and its responsiveness to DRD treatment regimens has not yet been evaluated.

*Miscellaneous Questions.* The RetDQoL is available upon request from the authors and it can be easily implemented in both high and low-resource settings. It can be either self- or interviewer-administered and is suitable for administration both face to face or via telephone (although it is recommended not to mix the mode of interview in the same study). With only 24 items, the RetDQoL is relatively quick to administer.

**RetCAT.** *Parameter Characteristics.* Diabetic Retinopathy and Macular Edema Computerized Adaptive Testing System is a digital and computerized adaptive testing (CAT) PROM that measures the impact of DRD, associated VI, and effectiveness of related treatments on QoL. Diabetic Retinopathy and Macular Edema CAT System contains 279 items within 10 QoL item banks, namely visual symptoms, activity limitation, mobility, emotional, health concerns, convenience, driving, lighting, social, and economic (Table S6, available at www.ophthalmologyscience.org). Diabetic Retinopathy and Macular Edema CAT System provides a continuous score in logits (transformed to a 1−99 percentile for better interpretability) for each QoL domain, with higher scores indicating "better" QoL outcomes. Items are administered from the item banks by the CAT system using an algorithm based on the item calibrations and patients' responses to all previous items. Diabetic Retinopathy and Macular Edema CAT System has undergone a rigorous, multistage development and validation process (LOE II) (see Supplementary Materials).[40−44]

*Scientific Understanding and Performance Expectations in DRD.* With patient-involved content development, it is unsurprising that RetCAT scores are related to the severity of DRD and VI.[43,44] For instance, 4 RetCAT tests (activity limitation, health concerns, lighting, and visual symptoms) have demonstrated reductions in test scores as DRD severity increases and, for binocular VI, RetCAT scores consistently decreased as the severity of VI worsened for all domains except convenience.

*Type of Data and Availability for Evidential Evaluation.* To date, RetCAT has been applied in clinical, cross-sectional studies in White and Asian populations as part of its developmental processes.[40,43,44]

*Statistical Considerations.* Unlike paper−pencil questionnaires, there is no need to treat RetCAT data with modern psychometric methods because the scores have been generated from a Rasch-calibrated item bank.

*Gap Analysis.* Responsiveness to treatment options and psychometric performance in DRD progression need to be further explored by its use in RCTs, clinical and population-based studies, and health care systems.

*Miscellaneous Questions.* Diabetic Retinopathy and Macular Edema CAT System is ready for use and commercially available. It provides fast measurement of QoL (~2 minutes per QoL domain) while maintaining high measurement precision.[43] With 10 domains available, it provides more comprehensive measurement than 2 to 3 domain paper−pencil questionnaires. Because RetCAT requires internet access and computer hardware (e.g., laptop and tablet) for administration, it may be less accessible for low-resource settings. However, for high resource settings, the RetCAT system reduces data storage, entry, and analysis needs, and enables real-time feedback via integration with e-health record systems. Diabetic Retinopathy and Macular Edema CAT System can be administered in clinic or at home via a secure URL on patients' own smart devices, making it attractive for health care in the coronavirus disease 2019 era where patients are spending minimal time physically attending clinical appointments. Qualification from the FDA is currently being sought for RetCAT.

## Utilities (Generic)

**EQ-5D and EQ-5D/VAS.** *Parameter Characteristics.* The EQ-5D is a multiattribute utility instrument that measures generic HRQoL.[45] It has 5 health dimensions (mobility, self-care, usual activities, pain/discomfort, and depression/anxiety).[46] Each dimension has either 3 levels of severity (EQ-5D-3L, describing 243 health states) or 5 levels of severity (EQ-5D-5L, describing 3125 health states). Health status can also be valued directly using the EQ-5D/VAS ranging from 0 to 100, with 0 representing the worst imaginable health state and 100 the best.[47]

*Scientific Understanding and Performance Expectations in DRD.* In health technology assessment, a difference of 0.06 to 0.07 points is considered a clinically meaningful change in DME in the better seeing eye. Change in EQ-5D does not correlate with changes in visual acuity

outcomes in clinical trials evaluating interventions for DME.[6,48–50] Similarly, the EQ-5D has been unable to elicit a difference between DR severity levels classified according to the better eye, worse eye, or VI in the worse eye in cross-sectional studies.[51,52] The disutility of blindness in one eye ranged from −0.074 to −0.054.[53,54] Only extreme visual acuity categories demonstrate a statistically significant difference (LOE III).

*Type of Data and Availability for Evidential Evaluation.* The EQ-5D has been used in cross-sectional, clinical, population-based, and RCT studies in patients with DR.[51,55,56]

*Statistical Considerations.* The EQ-5D-3L index ranges from −0.039 to 1,[4–6] with scores derived from responses from 3395 people randomly selected from the general population in the United Kingdom.[47] Value sets from a range of different countries are available for the EQ-5D-3L and EQ-5D-5L, with cross-walk value sets also available to map 3L scores on 5L scores.

*Gap Analysis.* Overall, studies investigating the validity and reliability of EQ-5D in various severity levels of DR are limited, and there is a lack of longitudinal validity to measure a change in health-related utility based on DR over time.[51,52] Changes in EQ-5D may also reflect the feelings of vulnerability to vision loss in patients with mild DR, which confound the values.[57] Similarly, the EQ-5D may capture changes related to side effects of potential treatments of DR in the context of general health dimensions of an individual.

To address the lack of sensitivity of the EQ-5D to vision loss, a set of vision questions have been added as a supplemental bolt-on,[58] although this has not been validated against severity levels of DR.[59] Moreover, because visual acuity is unaffected even in high-risk PDR without complications, the EQ-5D bolt-on is also unlikely to capture changes that correlate with changes in DR severity levels.

*Miscellaneous Questions.* Despite the gaps reported above, the EQ-5D is the most used generic preference-based measure of health in ophthalmology and is quick to use and freely available.

Although the EQ-5D is useful to define the impact of ocular disease compared with other health parameters, it has limited usefulness as a biomarker of disease-specific vision deterioration. Additional modifications to the scale of EQ-5D that can account for responsiveness to change in the DRD staging system may be worthwhile because it may serve as a comparator when evaluating other ranking or discrete choice methods because these methods may also correlate with other systemic comorbidities in people with diabetes and help decipher the findings observed in other methods.

**SF-6D.** *Parameter Characteristics.* The SF-6D is derived from the SF-36 and SF-12 health questionnaires and has 6 dimensions (physical functioning, role limitation, social functioning, bodily pain, mental health, and vitality), and each dimension has 4 to 6 severity levels.[47,60] This results in 18 000 possible unique health states, which is higher than the EQ-5D and therefore may mean the SF-6D can detect more subtle changes in health state than the EQ-5D. A tariff of values for each health state is also available for several countries.

*Scientific Understanding and Performance Expectations in DRD.* In a cohort study on people with type 2 diabetes with 90% having ≥ 2 comorbidities over 1 year, both the EQ-5D and SF-6D performed equally well in people with well controlled diabetes but differed in people with multiple systemic comorbidities (including 38% with an eye disease).[61] There are no data on the performance of the SF-6D based on severity levels of DR or its correlation with changes in DR severity level (LOE III).[62]

*Type of Data and Availability for Evidence Evaluation.* The SF-6D has been compared with EQ-5D and -15D utilities in a cross-sectional study on type 2 diabetes with and without comorbidities and complications, including DR.[56] The SF-6D utilities were not sensitive to differentiate people with and without DR.

*Statistical Considerations.* The theoretical range of SF-6D preference-based utility scores based on the SG technique ranges from 1 for full health to 0.345 for the worst possible health state. Unlike the EQ-5D, the SF-6D includes role limitation and social functioning, meaning it provides a more holistic measure of HRQoL. Any patient who completes the SF-36 or the SF-12 can be uniquely classified according to the SF-6D. However, while the EQ-5D suffers from a ceiling effect, the SF-6D has a floor effect,[55] meaning that it lacks capability to differentiate HRQoL in those at the lower end of the "ability" spectrum.

*Gap Analysis.* There is no evidence that the SF-6D can accurately correlate with DRD.[55]

*Miscellaneous Questions.* The SF-6D is quick to use, freely available, and can be easily calculated from its parent PROMs, SF-12 or SF-36, making it a flexible PROM option.

## Utilities (Vision-Specific)

**HUI3.** *Parameter Characteristics.* The HUI3 has 9 dimensions (vision, hearing, speech, ambulation, dexterity, emotion, cognition, self-care, and pain), and each dimension has 5 or 6 severity levels, resulting in 972 000 possible unique health states. Compared with other generic preference-based measures, it also has a vision dimension.[48,49]

*Scientific Understanding and Performance Expectations in DRD.* Using TTO to generate utilities from the HUI3, mean values of disutility in no DR, background DR, PDR, and legal blindness were 0.94, 0.87, 0.83, and 0.81, respectively, with a standard deviation of 0.14.[63] When considering DR severity in the worse eye, the mean QALY weights for no DR, background DR, PDR, DME, and legal blindness were 0.88, 0.78, 0.81, 0.79, and 0.39, respectively. The HUI3 showed a high correlation with visual acuity and with the NEI VFQ-25.[55] A decrement of the HUI3 of 0.057 has also been reported in those with DR and severe visual loss.[64] However, the influence of severity levels of DR was not assessed (LOE III).

*Type of Data and Availability for Evidence Evaluation.* The HUI3 has been used within the ACCORD clinical trial as a PROM context to measure QALYs and

cardiovascular event-free years gained.[55] It relates to VI better than EQ-5D, but it was not sensitive enough to identify severity levels of DR unless there was associated VI.

*Statistical Considerations.* The value of the HUI3 multiattribute utility score ranges from −0.36 to 1, where 1 represents perfect health, 0 represents death, and a negative index score indicates a health state considered to be worse than death. These values were elicited from a random sample of 504 people in a Canadian general population using the VAS in combination with the SG method.[47,61,62]

*Gap Analysis.* Overall, studies investigating the association of HUI3 utilities to severity levels of DR are limited, and there is a lack of longitudinal data to measure a change in utility based on DR progression over time. Future research is required to understand whether it is worthwhile to be incorporated in the DRD staging system.

*Miscellaneous Questions.* The HUI3 requires a license to use, with an associated cost. With multiple domains and items, it is relatively lengthy to administer, and the scoring algorithm to convert raw scores to utilities is relatively complex.

**VisQoL.** *Parameter Characteristics.* The VisQoL is a descriptive system derived from the Assessment of QoL-7D that covers 6 dimensions of self-reported VRQoL: physical well-being, independence, social well-being, self-actualization, planning, and organization.[65] The health states defined by the VisQoL responses are translated into VisQoL utilities using an available value set that was derived from participant surveys of the general population using the TTO method.

*Scientific Understanding and Performance Expectations in DRD.* Vision and Quality of Life Index changes correlated with very severe visual loss in DME and were independently associated with VisQoL utilities (LOE III).[66]

*Types of Data and Availability for Evidence Evaluation.* The VisQoL has been utilized only in cross-sectional studies related to DR/DME.

*Statistical Considerations.* Item utilities are combined using a multiplicative model and the scale of the utility index ranges from −0.25 to 1, where 0.0 represents death, 1.0 represents full health, and −0.25 represents worse than death.[67] The VisQoL utilities are categorized into 5 groups using 15th, 30th, 45th, 60th, and 75th percentiles as cut points ranging from poorest to highest VRQoL. Regression coefficients for each group are provided and represent the change in VisQoL utility score per unit change of each covariate for that particular level.

It should be noted that the predicted utility score may be less accurate for application at the individual level. However, the highly accurate group-level predictions ensure the usefulness of mapping algorithms in the cost utility analysis.

*Gap Analysis.* The VisQoL has been evaluated only in cross-sectional studies, and criterion and convergent validity have not been established.

*Miscellaneous Questions.* The VisQoL is a freely available, noninvasive test that takes < 10 minutes to complete. It can be either self- or interviewer-administered.

## Utilities (DRD-Specific)

**TTO and SG.** *Parameter Characteristics.* Time trade-off and SG are both direct elicitation methods that measure the utility scale and are a currently accepted model for QoL measurements in ophthalmology (Table S7, available at www.ophthalmologyscience.org). A utility value of 1.0 is equivalent to perfect health. A utility value of 0.0 is equal to death. Negative values represent a state worse than death.[68] In SG, a subject is given the scenario that a treatment, when successful, will provide a perfect health state, but an unsuccessful treatment will result in immediate death. The subject is then asked what percentage chance of treatment failure they would accept before agreeing to the treatment.[68]

*Scientific Understanding and Performance Expectations in DRD.* Values have been established for different levels of visual acuity using TTO and SG, although values vary considerably (Supplementary Materials). For example, for 20/40 vision, utility scores of 0.80 (TTO) and 0.89 (SG) have been established, whereas for 20/400 vision, scores are 0.54 (TTO) and 0.59 (SG). For no-light perception, utility scores are 0.35 (TTO) and 0.49 (SG). Because of their numeric scale, utility scores can be compared with other health disease states[48] (e.g., major stroke [0.61][69] and rheumatoid arthritis [0.77][70]) that have also been scored on the TTO scale. This comparative capability is advantageous to other methods of utility measurements,[52] as being able to compare disease states across different conditions is a critical requirement for informing public health decisions.

Useful cut points with defined values have been proposed for visual acuity for both TTO and SG and for TTO for severity of DR (LOE III).[6]

*Types of Data and Availability for Evidence Evaluation.* Brown et al[68] established the landmark publication for TTO values that were assessed in 325 subjects (0.77, standard deviation = 0.23, 95% confidence interval = 0.02), and these values are most often used today. Since then, other groups in the United States, Canada,[71] Iran,[72] Thailand,[73] Sweden, Taiwan,[74] and China[75] have also utilized TTO questions to establish utility scores for patients with DR.[6]

*Statistical Considerations.* Covariables sex and age have been previously shown to have an effect on HRQoL[52] and therefore should be adjusted for when analyzing utility scores. Duration of disease may also impact TTO utility ratings scores, and utility scores have been shown to shift over time, suggesting that patients with vision loss adjust to living with the disability.[49,63,76] The reliability of TTO/SG is affected by variability in the process (style of questioning and choice of words[77]) as well as in the patients (cultural variations,[75] comfort with discussing death,[75] degree of cognitive loss,[78] duration of disease,[79] current comorbilities,[80] perception of their own life compared with others, opinion of ideal health,[81] and past experience with health care system[81,82]). Other critical limitations include mathematical inconsistences that compromise the validity of the equation and participants' rating of worse than death.[81]

From a statistical standpoint, utility scales are interval scales (where 0 is an arbitrary value for death). This number is then entered a ratio-scale with a true zero. This violates the mathematical principles of use of interval numbers and compromises the validity of the equation.

*Gap Analysis.* The pivotal TTO studies for DR had small sample sizes and lacked diversity, which limits their applicablilty in larger scale studies. Additionally, although one study showed good retest reliability[79] with TTO in patients with ocular disease, several other studies highlight reliability as a limitation of this method.[71,76,83,84] The sensitivity of TTO/SG utility scores to detect progression of DRD is not known,[72] and additional information on the ability of this tool to measure change after intervention is needed.

*Miscellaneous Questions.* Overall, evaluation of TTO/SG utility scores is highly useful in comparing DRD to other disease states. However, the lack of testing standardization and the variable cultural acceptance of this form of questioning creates significant challenges for use of this measure as a robust biomarker for QoL. Other methods of eliciting utilities, such as ranking methods and discrete choice experiments (DCEs) have been shown to have higher reproducibility and reliability.

**DR-U.** *Parameter Characteristics.* The DR-U is a preference-based utility instrument developed using a DCE in a clinical sample of 220 patients with DR/DME across the spectrum of disease severity.[85] Utilities are estimated from the DR-U via a descriptive system comprising 5 dimensions of DR-specific QoL, including visual symptoms, lighting and glare, activity limitation and mobility, socio-emotional well-being, and inconvenience, with each dimension rated using 3 severity levels ranging from "no difficulty" to "a lot of difficulty."

*Scientific Understanding and Performance Expectations in DRD.* The DR-U is sensitive to DR severity and DR-related VI, with utilities decreasing as DR severity (nonvision-threatening DR = 0.87; vision-threatening DR = 0.80; $P = 0.021$) and better eye VI (none = 0.84; mild = 0.78; moderate/severe = 0.72; $P = 0.012$) increased. The DR-U also demonstrated divergent and convergent validity, with low ($r = 0.39$) and moderate ($r = 0.58$) correlations with EQ-5D and VisQoL utilities, respectively (LOE III).[7]

*Type of Data and Availability for Evidence Evaluation.* The use of the DR-U beyond the development population (cross-sectional) is limited.

*Statistical Considerations.* The value of the DR-U utility ranges from 0.58 to 1, where 1 and 0.58 represent the best and worst possible health states, respectively. The bottom anchor was rescaled from 0 to 0.58 based on the mean TTO value for DR-related blindness in 7 reported studies.[52] The raw scores collected using the DR-U descriptive system are converted to utilities using the utility weights determined by the DCE. The scale of the DR-U differs from the "full health" and "being dead" QALY scale, meaning that comparisons across different diseases and treatments on a common metric and generation of QALYs are difficult. Studies that intend to inform decision making on a general population level may also need to incorporate a generic utility instrument.

*Gap Analysis.* The DR-U is not approved by regulatory bodies. Discrete choice experiments use a flexible methodology to estimate which attributes are important in decision making.[86] The DR-U's responsiveness to change over time due to intervention and its association with all levels of DRD remain unclear as does the applicability and reproducibility of this measure across a diverse population.

*Miscellaneous Questions.* Because the DR-U is commercially available in both paper—pencil and digital format, it can be easily implemented in both high and low-resource settings. It can be either self- or interviewer-administered. The DR-U takes approximately 5 to 10 minutes to administer. Qualification from the FDA is currently being sought for the DR-U.

## Summary of Key Findings

Of the 12 reviewed PROMs, the NEI VFQ-25 had a LOE II and a wide usage in cross-sectional, longitudinal, and RCT studies (Table 2). The IVI and RetCAT were given a LOE of II due to their large-scale application in both clinical and population-based studies and robust development and validation protocol, respectively. Most PROMs, particularly the utilities, had a LOE of III due to most of the evidence being cross-sectional in nature. The NEI VFQ-25 is ready to be included in late-stage DRD, mainly in DME (Table 4), whereas EQ-5D/VAS, RetCAT, RetDQoL, and DR-U are considered potential for late-stage clinical DRD. No PROMs were deemed currently relevant to subclinical DRD or early stages of clinical DRD.

## Discussion

In our critical review of the available generic-, vision-, and DRD-related questionnaires and utility instruments, we showed that NEI VFQ-25, EQ-5D/VAS, RetCAT, RetDQoL, and DR-U have the potential to be included in an updated DRD staging system, provided further satisfactory modifications and validation of these tools meet the required LOE. Further areas of focus for these measures include the ability to elicit a difference between different severity levels of clinical DRD and demonstrate longitudinal validity in terms of responsiveness to DRD interventions. A minimal identifiable difference or a clinically meaningful change as a response to interventions needs to be identified, especially if the worse eye is being treated. Moreover, these PROMs will need further development and validation if new DRD staging is developed that includes preclinical or early structural or functional changes that precede visual acuity changes.[87]

A key advantage of the NEI VFQ-25 is that it has been widely used in DRD research, and therefore there is much comparative data available. However, its psychometric limitations and optimal functionality must be taken into consideration if it is to be further used in future DRD staging. Similarly, the lack of cut-off scores limits its

usefulness in DRD staging and points to an area for future research.

Importantly, there is also an unmet need to move from paper−pencil questionnaires to computerized systems in the digital and post-coronavirus disease 2019 era. One potential solution is item banking and CAT. Computerized adaptive testing is a digital method of administering items (questions) from calibrated item banks. Each item bank measures a latent construct (e.g., "activity limitation") and usually contains more items than normally found in a paper−pencil questionnaire. Using an algorithm based on artificial intelligence, CAT iteratively selects an item from the bank that most closely targets the person's level of the construct *at that point in the test*. This "smart-technology" means that scores are generated quickly using only a subset of the available items. To our knowledge, the RetCAT, reviewed in this paper, is the first CAT in DRD available for clinical, research, and health care needs.

Our review of generic multiattribute utility instruments, such as SF-6D, SF-12D, and HUI3, suggests that they require significantly more research to understand whether any of them can be incorporated into the DRD staging system because they lack sensitivity to DRD and related clinical outcomes. Similarly, the TTO and SG have several psychometric limitations, including flawed assumptions of the model, methodological and cultural issues, and response shift.[81]

An alternative approach would be to use DCE and ranking methods.[82] Such methods have shown promise for health state valuation as they offer a flexible methodology to estimate which attributes are important in decision making. However, these methods are also affected by participants drawing on past experiences and taking short cuts in their decision making.[88] The DR-U is the only currently available DRD-specific utility measure that has been developed using the DCE method. Future studies to determine its responsiveness to intervention/change over time, the detailed relationship with DRD levels, and to establish valid cut points and assess its cross-cultural applicability would strengthen the robustness of evidence of the DR-U's measurement properties.

Strengths of this study include the use of validated and comparative frameworks (i.e., LOE and *Biomarkers, EndpointS, and other Tools*) to assess the 12 chosen PROMs, which added robustness and objectivity of the review process. The main limitation of the review process is the potential for bias in terms of which PROMs were selected for review, given that this was not a formal, systematic review nor a full scoping review process but simply a narrative review based on the perspectives of a limited group of PROM and clinical experts. However, tiered evaluation by this group of experts should have mitigated the likelihood for inconsistencies or controversial assessments.

Current DRD-related PROMs face a major challenge in that early and midstages of clinical DRD may not affect visual acuity, which primarily impacts patient's HRQoL. Unless other visual function parameters correlate better than visual acuity in detecting change in early DRD, it is challenging to meet the standards of an ideal PROM. Consequently, the ultimate PROM for DRD, potentially using a CAT system, needs to be further validated using both classical and modern psychometric methods in a large sample of patients across the spectrum of DRD, with a proven capacity to detect disease progression through all clinical DRD stages and respond to interventions at any DRD severity level. Digital administration options of this ideal DRD PROM are needed, especially via tablets and URLs sent to patients' own smart devices. However, digital exclusions and disparities need to be considered. An ability for time-effective and streamlined PROMs reports to be available in real-time, and even routed to health care electronic medical record systems to facilitate doctor−patient DRD management discussions, is ideally sought.

In conclusion, this narrative review has determined that no current PROM has all necessary parameters to be included in the updated DRD staging system. Collaborative efforts involving clinicians, trialists, modelers, economists, patient focus groups, and regulatory and funding agencies (e.g., FDA, Patient-Centered Outcomes Research Institute, NEI, European Medicine Agency, and National Institute for Health and Care Excellence) are being planned to develop and validate PROMs. Novel digital solutions (i.e., virtual reality), changes to existing validated PROM questionnaires, and application of utilities to patient-centered outcomes are some of the areas of key interest.

## Footnotes and Disclosures

Originally received: May 2, 2023.
Final revision: July 21, 2023.
Accepted: August 2, 2023.
Available online: August 9, 2023. Manuscript no. XOPS-D-23-00087R1.

[1] Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy.

[2] Eye Clinic, IRCCS MultiMedica, Milan, Italy.

[3] Division of Epidemiology and Clinical Applications, Clinical Trials Branch, National Eye Institute, National Institutes of Health, Bethesda, Maryland.

[4] Ophthalmology Department, Carle Foundation Hospital, Urbana, Illinois.

[5] Surgery Department, University of Illinois College of Medicine, Urbana, Illinois.

[6] Moorfields Biomedical Research Centre, Moorfields Eye Hospital, London, United Kingdom.

[7] Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

[8] Duke-NUS Medical School, National University of Singapore, Singapore.

Dr Emily Chew, the Editor-in-Chief of this journal, was recused from the peer-review process of this article and had no access to information regarding its peer-review.

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures: S.V.: Consultant − AbbVie, Apellis, Bayer, Boehringer & Ingelheim, Novartis, Roche; Honoraria − Zeiss.

E.C.: Advisory board − DSMC member, Advisor for Genentech.

S.S.: Research support − AbbVie, Bayer, Boehringer Ingleheim, Roche; Consultant − Apellis, Bayer, Biogen, Boehringer Ingleheim, Eyebiotech, Janssen, Novartis, Ocular Therapeutix, OcuTerra, Roche; Honoraria − AbbVie, Bayer, Boehringer Ingleheim, Roche; Travel expenses − Bayer, Boehringer Ingleheim, Roche; Advisory board − Bayer, Nova Nordisk; Medical writing − Boehringer Ingleheim.

The other authors have no proprietary or commercial interest in any materials discussed in this article. This paper is a part of the DRD Staging System Update Effort, a project of the "Mary Taylor Moore Vision Initiative" supported by the Mary Taylor Moore and S. Robert Levine, MD Charitable Foundation and JDRF.

HUMAN SUBJECTS: Human subjects were not included in this study.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Vujosevic, Chew, Labriola, Sivaprasad, Lamourex

Data collection: Vujosevic, Chew, Labriola, Sivaprasad, Lamourex

Analysis and interpretation: Vujosevic, Chew, Labriola, Sivaprasad, Lamourex

Obtained funding: N/A

Overall responsibility: Vujosevic, Chew, Labriola, Sivaprasad, Lamourex

Abbreviations and Acronyms:
**CAT** = computerized adaptive testing; **DCE** = discrete choice experiment; **DME** = diabetic macular edema; **DR** = diabetic retinopathy; **DRD** = diabetic retinal disease; **DR-U** = Diabetic Retinopathy Utility instrument; **EQ-5D** = European Quality of Life Questionnaire; **FDA** = Food and Drug Administration; **HRQoL** = health-related quality of life; **HUI3** = Health Utilities Index; **IVI** = Impact of Vision Impairment; **LOE** = level of evidence; **NEI VFQ-25** = National Eye Institute 25-item Visual Function Questionnaire; **PDR** = proliferative diabetic retinopathy; **PROM** = patient-reported outcome measure; **QALY** = quality-adjusted life-year; **QoL** = quality of life; **RCT** = randomized clinical trial; **RetCAT** = Diabetic Retinopathy and Macular Edema Computerized Adaptive Testing System; **RetDQoL** = Diabetic Retinopathy Dependent Quality of Life; **SF-6D** = Short Form-6D; **SG** = standard gamble; **TTO** = time trade-off; **VAS** = visual analog scale; **VI** = visual impairment; **VisQoL** = Vision and Quality of Life index; **VRQoL** = vision-related quality of life.

Keywords:
Diabetic retinal disease, Patient-reported outcomes, Quality of life, Systematic review.

Correspondence:
Stela Vujosevic, MD, PhD, Via San Vittore 12, 20123 Milan, Italy. E-mail: stela.vujosevic@unimi.it.

## References

1. Vujosevic S, Aldington SJ, Silva P, et al. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol*. 2020;8:337−347.
2. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie House classification. ETDRS report number 10. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology*. 1991;98:786−806.
3. Wilkinson CP, Ferris FL, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677−1682.
4. What is health related quality of life?. https://www.isoqol.org. Accessed June 30, 2022.
5. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res*. 1972;7:118−133.
6. Brown MM, Brown GC, Sharma S, Shah G. Utility values and diabetic retinopathy. *Am J Ophthalmol*. 1999;128:324−330.
7. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.
8. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, et al. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q*. 2014;92:754−775.
9. Fenwick EK, Pesudovs K, Rees G, et al. The impact of diabetic retinopathy: understanding the patient's perspective. *Br J Ophthalmol*. 2011;95:774−782.
10. Sun JK, Aiello LP, Abràmoff MD, et al. Updating the staging system for diabetic retinal disease. *Ophthalmology*. 2021;128:490−493.
11. Baethge C, Goldbeck-Wood S, Mertens S. SANRA—a scale for the quality assessment of narrative review articles. *Res Integr Peer Rev*. 2019;4:5.
12. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009;101:1446−1452.
13. Visual Function Questionnaire (VFQ-25). https://www.rand.org/health-care/surveys_tools/vfq.html. Accessed June 30, 2022.
14. Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol*. 2001;119:1050−1058.
15. Burr JM, Cooper D, Ramsay CR, et al. Interpretation of change scores for the National Eye Institute Visual Function Questionnaire-25: the minimally important difference. *Br J Ophthalmol*. 2022;106:1514−1519.
16. Varma R, Paz SH, Azen SP, et al. The Los Angeles Latino Eye Study: design, methods, and baseline data. *Ophthalmology*. 2004;111:1121−1131.
17. Klein R, Moss SE, Klein BE, et al. The NEI-VFQ-25 in people with long-term type 1 diabetes mellitus: the Wisconsin Epidemiologic Study of Diabetic Retinopathy. *Arch Ophthalmol*. 2001;119:733−740.
18. Petrillo J, Bressler NM, Lamoureux E, et al. Development of a new Rasch-based scoring algorithm for the National Eye Institute Visual Functioning Questionnaire to improve its interpretability. *Health Qual Life Outcomes*. 2017;15:157.
19. Marella M, Pesudovs K, Keeffe JE, et al. The psychometric validity of the NEI VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci*. 2010;51:2878−2884.
20. Pesudovs K, Gothwal VK, Wright T, Lamoureux EL. Remediating serious flaws in the National Eye Institute Visual

Function Questionnaire. *J Cataract Refract Surg*. 2010;36:718–732.
21. Goldstein JE, Bradley C, Gross AL, et al. The NEI VFQ-25C: calibrating items in the National Eye Institute Visual Function Questionnaire-25 to enable comparison of outcome measures. *Transl Vis Sci Technol*. 2022;11:10.
22. Weih LM, Hassell JB, Keeffe J. Assessment of the impact of vision impairment. *Invest Ophthalmol Vis Sci*. 2002;43:927–935.
23. Lamoureux EL, Pallant JF, Pesudovs K, et al. The Impact of Vision Impairment Questionnaire: an evaluation of its measurement properties using Rasch analysis. *Invest Ophthalmol Vis Sci*. 2006;47:4732–4741.
24. Lamoureux EL, Pallant JF, Pesudovs K, et al. The impact of vision impairment questionnaire: an assessment of its domain structure using confirmatory factor analysis and Rasch analysis. *Invest Ophthalmol Vis Sci*. 2007;48:1001–1006.
25. Fenwick EK, Man REK, Gan ATL, et al. Beyond vision loss: the independent impact of diabetic retinopathy on vision-related quality of life in a Chinese Singaporean population. *Br J Ophthalmol*. 2019;103:1314–1319.
26. Man RE, Fenwick EK, Sabanayagam C, et al. Differential impact of unilateral and bilateral classifications of diabetic retinopathy and diabetic macular edema on vision-related quality of life. *Invest Ophthalmol Vis Sci*. 2016;57:4655–4660.
27. Lamoureux EL, Hassell JB, Keeffe JE. The impact of diabetic retinopathy on participation in daily living. *Arch Ophthalmol*. 2004;122:84–88.
28. Lamoureux EL, Tai ES, Thumboo J, et al. Impact of diabetic retinopathy on vision-specific function. *Ophthalmology*. 2010;117:757–765.
29. Aroney C, Fraser-Bell S, Lamoureux EL, et al. Vision-related quality of life outcomes in the BEVORDEX study: a clinical trial comparing ozurdex sustained release dexamethasone intravitreal implant and bevacizumab treatment for diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2016;57:5541–5546.
30. Gillies MC, Lim LL, Campain A, et al. A randomized clinical trial of intravitreal bevacizumab versus intravitreal dexamethasone for diabetic macular edema: the BEVORDEX study. *Ophthalmology*. 2014;121:2473–2481.
31. Goldstein JE, Fenwick E, Finger RP, et al. Calibrating the Impact of Vision Impairment (IVI): creation of a sample-independent visual function measure for patient-centered outcomes research. *Transl Vis Sci Technol*. 2018;7:38.
32. Fenwick EK, Man RE, Rees G, et al. Reducing respondent burden: validation of the Brief Impact of Vision Impairment questionnaire. *Qual Life Res*. 2017;26:479–488.
33. Fenwick EK, Loe BS, Khadka J, et al. Optimizing measurement of vision-related quality of life: a computerized adaptive test for the impact of vision impairment questionnaire (IVI-CAT). *Qual Life Res*. 2020;29:765–774.
34. Bradley C, Todd C, Gorton T, et al. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. *Qual Life Res*. 1999;8:79–91.
35. Brose LS, Bradley C. Psychometric development of the individualized Retinopathy-Dependent Quality of Life Questionnaire (RetDQoL). *Value Health*. 2010;13:119–127.
36. Deswal J, Narang S, Gupta N, et al. To study the impact of diabetic retinopathy on quality of life in Indian diabetic patients. *Indian J Ophthalmol*. 2020;68:848–853.
37. Ramu J, Chatziralli I, Yang Y, et al. Health-related quality of life, visual function and treatment satisfaction following intravitreal dexamethasone implant for diabetic macular edema. *Patient Prefer Adherence*. 2017;11:579–586.
38. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Multiplicative rating scales do not enable measurement of vision-related quality of life. *Clin Exp Optom*. 2011;94:52–62.
39. Finger RP, Fenwick E, Pesudovs K, et al. Rasch analysis reveals problems with multiplicative scoring in the macular disease quality of life questionnaire. *Ophthalmology*. 2012;119:2351–2357.
40. Fenwick EK, Pesudovs K, Khadka J, et al. The impact of diabetic retinopathy on quality of life: qualitative findings from an item bank development project. *Qual Life Res*. 2012;21:1771–1782.
41. Willis G. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage Publications; 2005.
42. Fenwick EK, Pesudovs K, Khadka J, et al. Evaluation of item candidates for a diabetic retinopathy quality of life item bank. *Qual Life Res*. 2013;22:1851–1858.
43. Fenwick EK, Khadka J, Pesudovs K, et al. Diabetic retinopathy and macular edema quality-of-life item banks: development and initial evaluation using computerized adaptive testing. *Invest Ophthalmol Vis Sci*. 2017;58:6379–6387.
44. Fenwick EK, Barnard J, Gan A, et al. Computerized adaptive tests: efficient and precise assessment of the patient-centered impact of diabetic retinopathy. *Transl Vis Sci Technol*. 2020;9:3.
45. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37:53–72.
46. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20:1727–1736.
47. Busija L, Pausenberger E, Haines TP, et al. Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care Res (Hoboken)*. 2011;63:S383–S412.
48. Brown MM, Brown GC, Sharma S, Busbee B. Quality of life associated with visual loss: a time tradeoff utility analysis comparison with medical health states. *Ophthalmology*. 2003;110:1076–1081.
49. Sharma S, Oliver-Fernandez A, Bakal J, et al. Utilities associated with diabetic retinopathy: results from a Canadian sample. *Br J Ophthalmol*. 2003;87:259–261.
50. Smith DH, Johnson ES, Russell A, et al. Lower visual acuity predicts worse utility values among patients with type 2 diabetes. *Qual Life Res*. 2008;17:1277–1284.
51. Heintz E, Wiréhn AB, Peebo BB, et al. QALY weights for diabetic retinopathy-a comparison of health state valuations with HUI-3, EQ-5D, EQ-VAS, and TTO. *Value Health*. 2012;15:475–484.
52. Bagust A, Beale S. Modelling EuroQol health-related utility values for diabetic complications from CODE-2 data. *Health Econ*. 2005;14:217–230.
53. Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Med Decis Making*. 2002;22:340–349.
54. Clarke PM, Simon J, Cull CA, Holman RR. Assessing the impact of visual acuity on quality of life in individuals with

type 2 diabetes using the Short Form-36. *Diabetes Care*. 2006;29:1506–1511.
55. Lloyd A, Nafees B, Gavriel S, et al. Health utility values associated with diabetic retinopathy. *Diabet Med*. 2008;25:618–624.
56. Kontodimopoulos N, Pappa E, Chadjiapostolou Z, et al. Comparing the sensitivity of EQ-5D, SF-6D and 15D utilities to the specific effect of diabetic complications. *Eur J Health Econ*. 2012;13:111–120.
57. Coyne KS, Margolis MK, Kennedy-Martin T, et al. The impact of diabetic retinopathy: perspectives from patient focus groups. *Fam Pract*. 2004;21:447–453.
58. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18:1–224.
59. Finch AP, Brazier JE, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: examining their contribution to health-related quality of life. *Value Health*. 2019;22:50–61.
60. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21:271–292.
61. The Health Utilities Index. http://healthutilities.com/. Accessed June 30, 2022.
62. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcome*. 2003;1:54.
63. Tung TH, Chen SJ, Lee FL, et al. A community-based study for the utility values associated with diabetic retinopathy among type 2 diabetics in Kinmen, Taiwan. *Diabetes Res Clin Pract*. 2005;68:265–273.
64. Shao H, Yang S, Fonseca V, et al. Estimating quality of life decrements due to diabetes complications in the United States: the Health Utility Index (HUI) diabetes complication equation. *Pharmacoeconomics*. 2019;37:921–929.
65. Peacock S, Misajon R, Iezzi A, et al. Vision and quality of life: development of methods for the VisQoL vision-related utility instrument. *Ophthalmic Epidemiol*. 2008;15:218–223.
66. Fenwick EK, Xie J, Pesudovs K, et al. Assessing disutility associated with diabetic retinopathy, diabetic macular oedema and associated visual impairment using the Vision and Quality of Life Index. *Clin Exp Optom*. 2012;95:362–370.
67. Misajon R, Hawthorne G, Richardson J, et al. Vision and quality of life: the development of a utility measure. *Invest Ophthalmol Vis Sci*. 2005;46:4007–4015.
68. Brown GC. Vision and quality-of-life. *Trans Am Ophthalmol Soc*. 1999;97:473–511.
69. Hallan S, Asberg A, Indredavik B, Widerøe TE. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med*. 1999;246:309–316.
70. Tijhuis GJ, Jansen SJ, Stiggelbout AM, et al. Value of the time trade off method for measuring utilities in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2000;59:892–897.
71. Sharma S, Brown GC, Brown MM, et al. Validity of the time trade-off and standard gamble methods of utility assessment in retinal patients. *Br J Ophthalmol*. 2002;86:493–496.
72. Alinia C, Mohammadi SF, Lashay A, Rashidian A. Impact of diabetic retinopathy on health-related quality of life in Iranian diabetics. *Iran J Public Health*. 2017;46:55–65.
73. Au Eong KG, Chan EW, Luo N, et al. Validity of EuroQOL-5D, time trade-off, and standard gamble for age-related macular degeneration in the Singapore population. *Eye (Lond)*. 2012;26:379–388.
74. Socialstyrelsen, Nationella riktlinjer för diabetesvården 2010 (***National Guidelines*** for diabetes care 2010). Stockholm: Socialstyrelsen; 2011 [cited 2012 29th of February]; Available at: http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/17824/Bilaga4Metod.pdf; Accessed February 29, 2012.
75. Wu B, Li J, Wu H. Strategies to screen for diabetic retinopathy in Chinese patients with newly diagnosed type 2 diabetes: a cost-effectiveness analysis. *Medicine (Baltimore)*. 2015;94:e1989.
76. Hollands H, Lam M, Pater J, et al. Reliability of the time trade-off technique of utility assessment in patients with retinal disease. *Can J Ophthalmol*. 2001;36:202–209.
77. Shen J, Hill S, Mott D, et al. Conducting a time trade-off study alongside a clinical trial: a case study and recommendations. *Pharmacoecon Open*. 2019;3:5–20.
78. Ekman M, Berg J, Wimo A, et al. Health utilities in mild cognitive impairment and dementia: a population study in Sweden. *Int J Geriatr Psychiatry*. 2007;22:649–655.
79. Jansen SJ, Stiggelbout AM, Wakker PP, et al. Unstable preferences: a shift in valuation or an effect of the elicitation procedure? *Med Decis Making*. 2000;20:62–71.
80. Lugnér AK, Krabbe PFM. An overview of the time trade-off method: concept, foundation, and the evaluation of distorting factors in putting a value on health. *Expert Rev Pharmacoecon Outcomes Res*. 2020;20:331–342.
81. Arnesen T, Trommald M. Roughly right or precisely wrong? Systematic review of quality-of-life weights elicited with the time trade-off method. *J Health Serv Res Policy*. 2004;9:43–50.
82. Ali S, Ronaldson S. Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. *Br Med Bull*. 2012;103:21–44.
83. Zhu X, Sun Q, Zou H, et al. Disparities between ophthalmologists and patients in estimating quality of life associated with diabetic retinopathy. *PLOS ONE*. 2015;10:e0143678.
84. Attema AE, Edelaar-Peeters Y, Versteegh MM, Stolk EA. Time trade-off: one methodology, different methods. *Eur J Health Econ*. 2013;14(Suppl 1):S53–S64.
85. Fenwick EK, Bansback N, Gan ATL, et al. Validation of a novel diabetic retinopathy utility index using discrete choice experiments. *Br J Ophthalmol*. 2020;104:188–193.
86. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. *J Health Econ*. 2012;31:306–318.
87. Khoo K, Man REK, Rees G, et al. The relationship between diabetic retinopathy and psychosocial functioning: a systematic review. *Qual Life Res*. 2019;28:2017–2039.
88. Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? *Health Econ*. 2003;12:393–402.