Estimating Minimally Important Differences for Two Vision-Specific Quality of Life Measures

Brenda W. Gillespie,[1] David C. Musch,[2,3] Leslie M. Niziol,[2] and Nancy K. Janz[4]

[1]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, United States
[2]Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, Michigan, United States
[3]Department of Epidemiology, University of Michigan, Ann Arbor, Michigan, United States
[4]Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, Michigan, United States

PURPOSE. To estimate minimally important differences (MIDs) for the Visual Activities Questionnaire (VAQ) and the National Eye Institute-Visual Function Questionnaire (NEI-VFQ).

METHODS. A total of 607 subjects with newly-diagnosed open-angle glaucoma (OAG) was enrolled in the Collaborative Initial Glaucoma Treatment Study (CIGTS) and randomized to initial treatment with medications or surgery. Subjects underwent an ophthalmic examination and telephone-administered quality of life (QOL) interview before randomization and every six months thereafter. The VAQ and NEI-VFQ were used to assess participants' perceptions of their visual function. Clinical measures included the mean deviation (MD) from Humphrey 24-2 full threshold visual field (VF) testing, and best-corrected visual acuity (VA) measured using the Early Treatment Diabetic Retinopathy Study (ETDRS) protocol. Anchor-based (using MD and VA) and distribution-based methods were used to estimate MIDs.

RESULTS. Anchor-based cross-sectional analyses at 66 months follow-up found a 10-letter increment in better eye VA corresponded to MIDs of 5.2 units for VAQ and 3.8 units for NEI-VFQ total scores. A 3-dB increment in the better eye MD yielded MIDs of 2.6 and 2.3 units for the same two questionnaires. In longitudinal analyses, MIDs for the VAQ were 3.2 units for a 10-letter change of VA and 3.4 units for a 3-dB change in the MD. Distribution-based MIDs were larger.

CONCLUSIONS. A range of MIDs for the VAQ (2.6–6.5 units) and NEI-VFQ (2.3–3.8 units) was found. Although similar in magnitude, MIDs were sensitive to the MID estimation method, the anchor chosen, and differences between questionnaires. (ClinicalTrials.gov number, NCT00000149.)

Keywords: MID, glaucoma, CIGTS, NEI-VFQ, VAQ

Patient-reported outcomes from health-related quality of life (HR-QOL) instruments represent important measures of the effect of a condition and its treatment on symptoms and functioning.[1] The HR-QOL questionnaire scores should be responsive to clinically evident disabilities, such that a certain change in score from HR-QOL questionnaires should reflect corresponding changes in the clinical condition.

The clinical utility of HR-QOL scores is enhanced by the estimation of a minimally important difference (MID), sometimes called a minimal clinically important difference (MCID),[2] which is the smallest difference in the measure that patients perceive as important, either in terms of benefit or harm, and which would lead a care provider to consider changing the patient's management.[3,4] Estimating the MID of a HR-QOL measure is considered essential to the interpretation of HR-QOL results when assessing within-patient or between-group differences.[1]

Methods for MID estimation fall into three general categories: distribution-based, anchor-based cross-sectional, and anchor-based longitudinal. Distribution-based MIDs are crude, being estimated from the variability of the HR-QOL distribution without reference to clinical anchors, such as visual field (VF) or visual acuity (VA) measures. Both of the anchor-based methods rely on regression analyses of HR-QOL scores on a clinical anchor. With cross-sectional data, the estimates are based on between-person differences in HR-QOL and anchor values. Using longitudinal data as change scores, the estimates reflect within-person differences, and are considered the most relevant of the MID methods.

A widely-used vision-specific HR-QOL measure is the National Eye Institute Visual Function Questionnaire (NEI-VFQ), with initial MID estimates from the Los Angeles Latino Eye Study (LALES). Using mean deviation (MD) as the anchor, McKean-Cowdin et al.[5] estimated MIDs in a cross-sectional sample of subjects with VF loss due to any cause. Those with a worse MD by 4 to 5 dB in the better eye reported a 4- to 5-point lower score on the NEI-VFQ. In participants with glaucoma,[6] they reported a 2.5-point lower score in the NEI-VFQ for similar loss in MD. In longitudinal analyses with 4 years of follow-up, the MID was reported to depend on level of visual functioning at baseline.[7] The generalizability from this study is unknown, and a recent systematic review questioned the strength of current evidence linking clinical evidence and vision-related QOL.[8]

We sought to address this gap by evaluating HR-QOL data collected over 9 years in the Collaborative Initial Glaucoma

Treatment Study (CIGTS), a multicenter, randomized clinical trial designed to compare medical versus surgical therapy as the initial treatment of open-angle glaucoma (OAG).[9] We used the CIGTS data to estimate MIDs for two vision-specific HR-QOL measures: the Visual Activities Questionnaire (VAQ)[10] and the NEI-VFQ.[11] We estimated the MIDs using anchor-based cross-sectional (VAQ and NEI-VFQ), anchor-based longitudinal (VAQ only), and distribution-based methods (VAQ and NEI-VFQ).

Methods

Study Participants

The CIGTS study protocol, eligibility criteria, and recruitment procedures have been described in detail previously.[9] In brief, 607 subjects were enrolled in the CIGTS between October 1993 and April 1997. Participants were newly diagnosed OAG patients from 14 clinical centers across the United States. Eligible participants were between 25 and 75 years of age and agreed to be followed for a minimum of five years, with reconsent obtained for additional follow-up. The CIGTS protocol and informed consent were approved by Institutional Review Boards at the University of Michigan and at all clinical sites, and research adhered to the tenets of the Declaration of Helsinki. Upon giving written informed consent, patients were randomized to initial trabeculectomy ($n = 300$) or initial medical therapy ($n = 307$).

Before randomization, subjects received a comprehensive ophthalmologic examination and a telephone-administered HR-QOL interview. Information on participants' age, race, sex, marital status, and education was obtained at the clinic site, while income level and employment status were collected during the HR-QOL interview. Standardized follow-up ophthalmologic examinations and HR-QOL interviews were conducted every 6 months following initial treatment. The CIGTS HR-QOL interviews were administered from a centralized interviewing facility by trained telephone interviewers.

The HR-QOL interview included generic and condition-specific measures, including the VAQ and NEI-VFQ, and has been described in detail.[12] The VAQ and NEI-VFQ assess patients' perceptions of their visual functioning and the impact of visual problems on their daily activities. The reliability and validity of these two measures have been studied with generally good psychometric results.[10,11,13] When the trial began, the 33-item VAQ was selected as the most relevant condition-specific measure of visual function status available. After enrollment was complete and CIGTS participants were being followed, the NEI released the 25-item version of the NEI-VFQ, which was first administered at the 54-month follow-up and annually thereafter.

The VAQ was developed to assess the effect of treatment on activities of daily living involving vision. The VAQ asks about problems individuals may have performing everyday tasks that involve visual function. Each item describes a vision problem and asks how often this problem occurs on a five-point scale from 1 (never) to 5 (always), or "not applicable." The VAQ includes eight subscales: color discrimination, glare disability, light-dark adaptation, acuity/spatial vision, depth perception, peripheral vision, visual search, and visual processing speed. The VAQ total score and subscale scores were calculated as a mean of all items, or the items in that subscale, respectively. Higher scores indicated worse visual functioning.

The NEI-VFQ was developed for people who have cataract, macular degeneration, diabetic retinopathy, primary OAG, cytomegalovirus retinitis, or low vision from any cause.[11] Since its development, it has been applied to patients with

these and a number of other ophthalmic conditions in clinical trials and population-based studies.[14-17] The NEI-VFQ yields a total score that includes all items except one that asks about general health. In addition, 12 subscale scores can be computed: general health, general vision, near vision, distance vision, driving, peripheral vision, color vision, and ocular pain as well as vision-specific role limitations, dependency, social function, and mental health. The NEI-VFQ is scored as an unweighted average of the included items transformed to a 0 to 100 scale, where 0 represents the worst possible score and 100 represents the best.

Clinical measures in the CIGTS included the MD from Humphrey 24-2 full threshold VF testing, and best-corrected VA measured using the Early Treatment for Diabetic Retinopathy Study (ETDRS) logarithmic charts.[18] These clinical measures were obtained at baseline (before treatment initiation) and every six months thereafter.

For cross-sectional analyses, we chose the 66-month follow-up to achieve the largest number of subjects with data from the VAQ and NEI-VFQ ($n = 420$). Longitudinal analysis on the VAQ required subjects having at least 12 months of follow-up from baseline with clinical and HR-QOL data, which included 97% (590) of participants. The HR-QOL measures were more highly correlated with the MD or VA of the better eye than the worse eye (the eye with better MD also had the better VA in almost all cases); therefore, all analyses involving clinical measures were performed using data from the better eye. For ease of interpretation, the VAQ was rescored to be on the same 0 to 100 scale as the NEI-VFQ, where a high score represents better functioning.

Statistical Methods

As recommended,[4] three methods to estimate MIDs for the HR-QOL questionnaires were used: anchor-based cross-sectional, anchor-based longitudinal, and distribution-based. For each of the anchor-based methods, two clinical anchors were used (MD and VA).

Anchor-Based Cross-Sectional. For each HR-QOL subscale at the 66-month follow-up, linear regression models of HR-QOL score by clinical anchor were fit. The MIDs were calculated by multiplying regression slope estimates by the prespecified clinically relevant difference for the anchor (3 dB for MD,[19] 10 letters for VA[20]).

Anchor-Based Longitudinal. For each HR-QOL subscale, the change from baseline to last follow-up (last time point at which clinical and HR-QOL data were collected on a subject) was regressed on the corresponding change in each anchor. The MIDs then were calculated by multiplying regression slope estimates by clinically relevant changes in each anchor, as defined above. For example, the MID estimate for the MD anchor would be 3 dB multiplied by the regression slope. The MID is interpreted as the difference on the QOL scale associated with a minimally important difference of 3 dB in MD. The NEI-VFQ was available only from 54 months on, and did not provide enough longitudinal data for this method.

Distribution-Based. This approach is based only on the distribution of the HR-QOL values and does not incorporate information from the anchors. Although various multiples of the standard deviation (SD) have been used to estimate the MID, we chose (SD ÷ 3) as recommended by Eton et al.[2] In addition, this choice gave results most similar to the anchor-based methods. The SDs of each subscale of the two HR-QOL questionnaires were calculated based on 66-month data and also based on the change between baseline and last follow-up. Other distribution-based methods as suggested by Crosby et al.[21] are presented in supplementary material.

Plots were used to visually compare the MIDs by anchors, by subscales, and by methods, and also to illustrate the range of

TABLE 1. Descriptive Statistics for the ≥12-Month Sample (Used for the Anchor-Based Longitudinal MID Estimates) and the 66-Month Cross-Sectional Sample (Used for the Anchor-Based Cross-Sectional MID Estimates)*

Continuous Variables	≥12-Month Sample, n = 590		66-Month Sample, n = 420	
	Mean (SD)	Min, Max	Mean (SD)	Min, Max
Age at baseline	58.0 (10.9)	28.8, 75.8	58.3 (10.8)	31.2, 75.8
Better eye baseline MD	−2.5 (2.7)	−18.1, 3.4	−2.2 (2.6)	−18.1, 3.4
Better eye baseline VA	87.9 (4.8)	70.0, 100.0	88.1 (4.7)	70.0, 100.0
Baseline global health	4.3 (1.1)	1.0, 5.0	4.3 (1.0)	1.0, 5.0
Categorical Variables	Frequency	Percent	Frequency	Percent
Sex				
Male	326	55%	228	54%
Female	264	45%	192	46%
Race				
White	330	56%	253	60%
Black	222	38%	148	35%
Asian	10	2%	4	1%
Other	28	5%	15	4%

MD, mean deviation; MID, minimally important difference; SD, standard deviation; VA, visual acuity.

* Distribution-based MID estimates were calculated for both samples.

MID estimates. To further illustrate the MIDs for the VAQ based on the longitudinal method, means and SDs of HR-QOL change within categories of clinical change were calculated. The change from baseline to last follow-up in each anchor measure was categorized into 4 to 5 levels covering the range of improvement, no change, or worsening. Box plots were generated to display the relationship between average HR-QOL change and categorical anchor change, as well as HR-QOL variability within each anchor category. All analyses were performed using SAS v.9.3 (SAS Institute, Inc., Cary, NC, USA).

RESULTS

Table 1 summarizes the baseline characteristics of two sets of subjects: those with at least 12 months of follow-up used in longitudinal MID estimates (n = 590), and the subset with follow-up at 66 months used in cross-sectional and most distribution-based MID estimates (n = 420). Subjects in the two samples were comparable in age (mean 58 years, SD = 11), sex (54%–55% male), better eye VA (88 letters, or approximately 20/20 Snellen VA), and global health impact rating (4.3 of 5.0, SD = 1). The few dropouts before 12 months (n = 17) did not differ from those who were active on any of these characteristics. Those with 12-month follow-up who were inactive at 66 months (n = 170) were more likely to be nonwhite (P = 0.0011, Fisher exact test), reported worse global health impact (P = 0.0036, t-test), and had worse baseline MD in the better eye (P = 0.0010, t-test) than those who were active (n = 420).

Anchor-based cross-sectional MIDs estimated from the 66-month sample are summarized in Table 2 for both anchors: better eye MD and better eye VA. Subscales most relevant for each anchor are noted by a footnote in the Table. The VAQ total

TABLE 2. MID Estimates From a Cross-Sectional Anchor-Based Approach Using Data at the 66-Month Follow-up for the Two Anchors, Better Eye MD† and Better Eye VA†

Quality of Life Scale	Better Eye MD MID (SE)	Better Eye VA MID (SE)
VAQ	n = 420	n = 427
VAQ total	2.6 (0.8)**	5.2 (1.1)***
Acuity/spatial vision‡	3.3 (1.0)**	7.3 (1.4)***
Peripheral vision§	3.4 (0.8)***	5.3 (1.2)***
Color discrimination	2.3 (0.8)**	3.5 (1.1)**
Glare disability	3.0 (1.1)**	4.8 (1.6)**
Light/dark adaptation	1.8 (1.1)	4.4 (1.6)**
Depth perception	2.5 (0.8)**	5.5 (1.1)***
Visual search	2.2 (0.9)*	5.0 (1.3)***
Visual processing speed	2.7 (0.8)**	5.6 (1.1)***
NEI-VFQ	n = 394	n = 401
NEI-VFQ total	2.3 (0.3)***	3.8 (0.5)***
Near activities‡	2.1 (0.4)***	3.7 (0.6)***
Peripheral vision§	2.1 (0.5)***	3.1 (0.8)***
General health	2.7 (1.0)**	5.6 (1.5)***
General vision‡	3.5 (0.6)***	6.2 (0.8)***
Ocular pain	1.2 (0.5)*	0.4 (0.8)
Distance activities	2.3 (0.4)***	4.1 (0.6)***
Social functioning	1.1 (0.2)***	1.3 (0.4)***
Mental health	2.9 (0.5)***	4.6 (0.8)***
Role difficulties	4.1 (0.7)***	6.5 (1.0)***
Dependency	1.4 (0.4)**	3.1 (0.6)***
Driving§ ‡	4.4 (0.7)***	8.2 (1.0)***
Color vision	0.9 (0.3)**	1.5 (0.4)**

MID, minimally important difference; MD, mean deviation; VA, visual acuity; SE, standard error; VAQ, Visual Activities Questionnaire; NEI-VFQ, National Eye Institute-Visual Function Questionnaire.

† Anchors were defined as 3 dB (for MD) and 10 letters (for VA).

‡ Subscales more relevant for the VA anchor. All four questions in the acuity/spatial vision subscale ask about near vision.

§ Subscales more relevant for the MD anchor.

Significance of the regression slope of QOL measures versus anchors: no superscript indicates $P > 0.05$; * indicates $0.01 < P \leq 0.05$; ** indicates $0.001 < P \leq 0.01$; *** indicates $P < 0.001$.

score and NEI-VFQ total score, respectively, showed MIDs of 2.6 and 2.3 for the MD anchor, and 5.2 and 3.8 for the VA anchor. For all but one scale, MIDs for the VA anchor were larger than those for the MD anchor. Among the various subscales of the two HR-QOL questionnaires, the peripheral vision and acuity/spatial vision subscales of the VAQ showed the largest MIDs for the MD anchor (3.4 and 3.3, respectively), whereas for the NEI-VFQ, the driving (4.4), role difficulties (4.1), and general vision (3.5) subscales showed the largest MIDs for the MD anchor. With regard to the VA anchor, the VAQ subscale that had the highest MID was the acuity/spatial vision subscale (7.3); for the NEI-VFQ, the same three subscales with the highest MID for the MD anchor also were highest for the VA anchor. Models adjusted for treatment effect (which was significant only in 5 of the models) yielded almost identical MIDs.

Table 3 summarizes the average change from baseline to last follow-up of three summary HR-QOL measures from the VAQ (total, acuity/spatial vision, and peripheral vision) within categories of anchor change during the same time period. The mean changes indicated that, on average, those subjects who had decreasing (worsening) MD over time also had decreasing HR-QOL over time for all three summary measures; likewise, subjects who had increasing MD also had increasing HR-QOL measures (Fig. 1). With regard to VA, loss of 10 or more letters was reflected by a mean decrease of 2.9 units for

TABLE 3. Evidence for Responsiveness of the VAQ Total and Subscale Scores at Various Increments of Change Across Anchor Categories

Anchor Category	n	Mean Change (SD)		
		VAQ Total	Acuity/Spatial Vision Subscale	Peripheral Vision Subscale
MD				
Loss ≥ 3 dB	89	−3.72 (17.18)	−2.29 (27.15)	−3.96 (21.40)
Loss 1.1–2.9 dB	119	0.94 (15.51)	5.02 (25.41)	−3.28 (18.08)
± 1 dB of baseline	224	1.56 (15.76)	4.28 (24.93)	−0.23 (17.87)
Gain 1.1–2.9 dB	120	4.85 (15.73)	8.35 (25.41)	3.59 (19.60)
Gain ≥ 3 dB	38	12.98 (19.82)	17.87 (31.14)	7.53 (21.50)
VA				
Loss ≥ 10 letters	106	−2.94 (19.12)	0.06 (29.15)	−4.80 (23.17)
Loss 5.1–9.9 letters	91	3.86 (16.27)	8.70 (25.48)	1.15 (18.82)
± 5 Letters of baseline	362	3.03 (15.60)	5.85 (25.25)	0.85 (17.99)
Gain > 5 letters	28	2.73 (18.00)	4.91 (28.23)	0.18 (20.47)

Note: change is calculated as the follow-up measure − baseline measure, in those subjects with at least 12 months of follow-up. VAQ, Visual Activities Questionnaire; SD, standard deviation; MD, mean deviation; dB, decibels; VA, visual acuity.

the VAQ total score. The other loss category (5.1–9.9 letters), the no change category, and the gain category (>5 letters) all showed an increase in the VAQ total score of approximately 3 units. Similar trends in QOL change within categories of VA change also were seen for the VAQ's acuity/spatial and peripheral vision subscales. However, the VA change patterns did not show the ordinal relationships seen with the MD anchor. As shown in the box plots for MD (Fig. 1), there is much variation in these data.

Anchor-based longitudinal MIDs, estimated from change-in-QOL versus change-in-anchor regression models, are summarized in Table 4. These change scores were calculated as the difference between baseline and the last time point where clinical and HR-QOL data were collected on a subject. These subjects had a median of 6.8 years of follow-up (range, 1–9 years). The MIDs were calculated for a 3 dB change in MD and for a 10-letter change in VA. The MIDs for the VAQ total score were 3.4 (MD anchor) and 3.2 (VA anchor). For the acuity/spatial score they were 3.9 and 2.5, and for the peripheral vision score they were 3.1 and 3.7. The MIDs for the MD and VA anchors were similar for the VAQ total and all subscales. Models adjusted for treatment effect (which was not significant in any of the models) yielded almost identical MIDs.

Distribution-based MIDs for the VAQ subscales were similar regardless of whether they were calculated at the 66-month cross-section or based on the change from baseline to last follow-up (Table 5). The 66-month MIDs were 6.5 for the VAQ total score, 8.5 for acuity/spatial vision, and 7.0 for peripheral vision. The MIDs for the NEI-VFQ total score and subscales were smaller than those from the VAQ at 66 months. Medicine and surgery specific MIDs were similar to
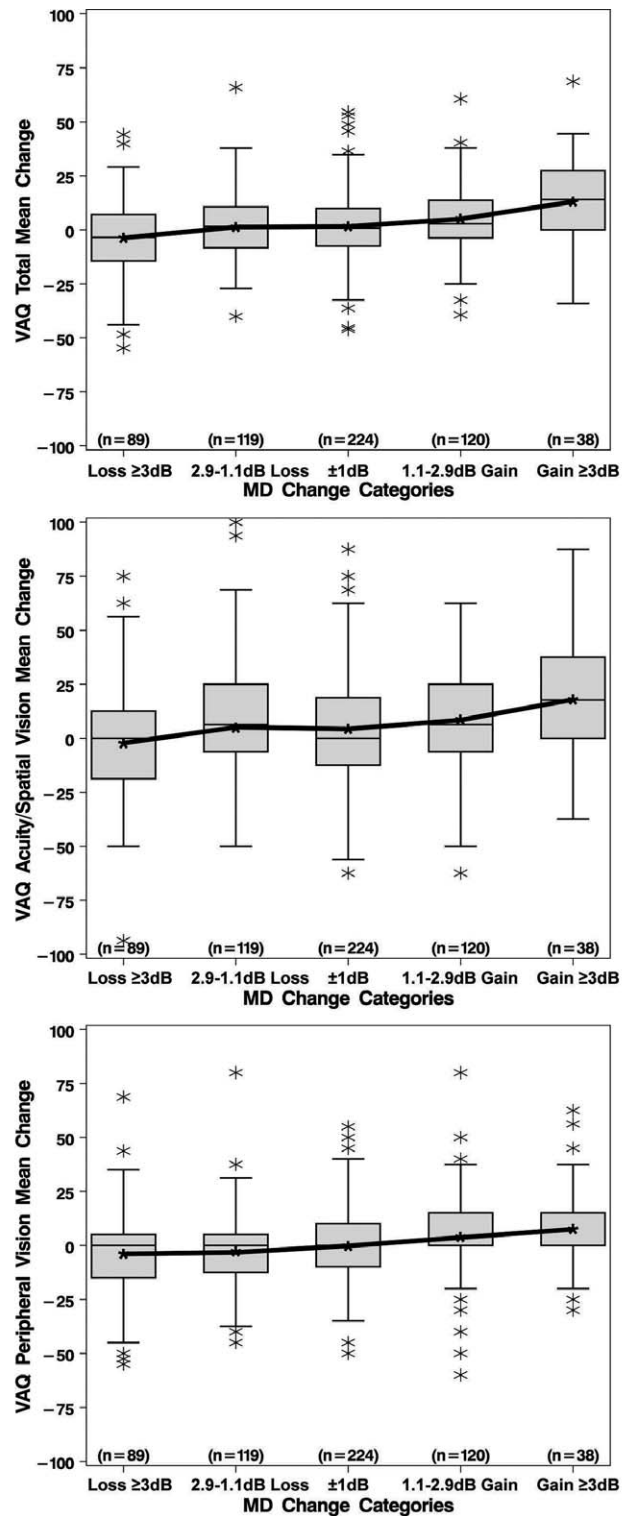
FIGURE 1. Boxplots showing change in VAQ subscales from baseline to last follow-up stratified by categories of MD change. Medians in the last panel of boxplots (for the VAQ peripheral vision) are zero for all categories of MD change. VAQ, Visual Activities Questionnaire; MD, mean devaiation; dB, decibels.

each other. In addition, Supplementary Table S1 gives MID estimates using ½ SD, as well as the standardized response means, effect sizes, and responsiveness statistics for the VAQ total and all VAQ subscales.

TABLE 4. MID Estimates From a Longitudinal Anchor-Based Approach Using the Differences Between Baseline and Last Follow-up Time Points ($n = 590$ With at Least 12 Months of Follow-up; Anchors Are the Better Eye MD† and Better Eye VA†)

Quality of Life Scale	Better Eye MD MID (SE)	Better Eye VA MID (SE)
VAQ		
VAQ total	3.4 (0.6)***	3.2 (0.7)***
Acuity/spatial vision‡	3.9 (0.9)***	2.5 (1.1)*
Peripheral vision§	3.1 (0.7)***	3.7 (0.8)***
Color discrimination	3.2 (0.7)***	2.6 (0.8)**
Glare disability	3.8 (1.0)***	4.2 (1.1)***
Light/dark adaptation	2.7 (0.9)**	2.4 (1.1)*
Depth perception	3.8 (0.7)***	3.3 (0.8)***
Visual search	3.2 (0.7)***	3.5 (0.9)***
Visual processing speed	3.6 (0.7)***	3.7 (0.8)***

Note: Longitudinal anchor-based MIDs could not be calculated for the NEI-VFQ, which was first used in the study at the 54-month follow-up, because there was insufficient subsequent follow-up to provide good estimates. MID, minimally important difference; MD, mean deviation; VA, visual acuity; SE, standard error; VAQ, Visual Activities Questionnaire; NEI-VFQ, National Eye Institute-Visual Function Questionnaire.

† Anchors were defined as 3 dB (for MD) and 10 letters (for VA).

‡ Subscales more relevant for the VA anchor. All four questions in the acuity/spatial vision subscale ask about near vision.

§ Subscales more relevant for the MD anchor.

Significance of the regression slope of QOL measures versus anchors: no superscript indicates $P > 0.05$; * indicates $0.01 < P \leq 0.05$; ** indicates $0.001 < P \leq 0.01$; *** indicates $P < 0.001$.

Figure 2 displays the MIDs by anchor and method for the VAQ and NEI-VFQ total scores, and the 2 subscales that are similar in content between the questionnaires (acuity/spatial vision for VAQ and near activities for NEI-VFQ; peripheral vision for VAQ and NEI-VFQ). The range of estimated MIDs for the VAQ total score was 2.6 to 6.5 units and for the NEI-VFQ total score it was 2.3 to 3.8 units. For the acuity/spatial subscale of the VAQ, MIDs ranged from 2.5 to 8.5, and for the near activities subscale of the NEI-VFQ MIDs ranged from 2.1 to 3.7. For the peripheral vision subscale, MIDs ranged from 3.1 to 7.0 for the VAQ and 2.1 to 4.3 for the NEI-VFQ. Distribution-based MIDs were somewhat larger than those from anchor-based methods.

DISCUSSION

Anchor-based methods to estimate MID values relate differences in relevant clinical measures of disease status and/or function to self-reported change in HR-QOL. This can be achieved using information collected from people who have differing severities of disease at a designated time (cross-sectional) or changes within the same person over time (longitudinal). This study used cross-sectional and longitudinal anchor-based approaches, and two anchors (a 3-dB change in MD and a 10-letter change in ETDRS VA) as well as a distribution-based approach to estimate MID values for the VAQ and the NEI-VFQ in newly diagnosed glaucoma patients with up to 9 years of follow-up. For the VAQ, cross-sectional and longitudinal data were available, whereas for the NEI-VFQ, which was only available after five years of follow-up, analyses were limited to cross-sectional data.

For anchor-based methods, the relevant anchor should clinically measure a similar aspect to that assessed by the HR-QOL measure. Anchors that do not match a particular HR-QOL measure may yield underestimates of the MID.[22] We chose two

TABLE 5. MID Estimates From a Distribution Based Approach

Variable	MID 1/3 SD of QOL Scale at 66-mo Cross-Section*	MID 1/3 SD of Change From Baseline to Last Follow-up in QOL Scale
VAQ		
VAQ total	6.5	5.5
Acuity spatial vision	8.5	8.7
Peripheral vision	7.0	6.4
Color discrimination	6.7	6.3
Glare disability	9.2	8.9
Light/dark adaptation	9.3	8.4
Depth perception	6.4	6.5
Visual search	7.9	6.9
Visual processing speed	6.7	6.4
NEI-VFQ		
NEIVFQ total	2.8	–
Near activities	3.7	–
Peripheral vision	4.3	–
General health	8.2	–
General vision	5.0	–
Ocular pain	4.3	–
Distance activities	3.5	–
Social functioning	2.1	–
Mental health	4.5	–
Role difficulties	5.8	–
Dependency	3.4	–
Driving	5.8	–
Color vision	2.5	–

Note: MIDs were not calculated for the NEIVFQ based on changes in scores from baseline because the questionnaire did not become available until 54 months into the study. MID, minimally important difference; SD, standard deviation; VAQ, Visual Activities Questionnaire; NEI-VFQ, National Eye Institute-Visual Function Questionnaire; QOL, quality of life.

* The distribution-based MIDs alternatively could have been calculated using baseline SDs; these were similar to those reported using 66-month SDs. Correlations between baseline and 66-month measures ranged from 0.45 to 0.61.

anchors (MD and VA) commonly used by vision care providers to evaluate visual function. In glaucoma applications, wherein early VF loss affects peripheral vision, the use of a visual field anchor, such as the MD, was predicted to be most relevant for the VAQ and NEI-VFQ peripheral vision subscales. Our findings showed a strong impact of the MD anchor on driving and role difficulty, which are activities requiring peripheral vision,[23] as well as on the 4-question VAQ peripheral vision subscale. However, the MID estimate based on the single-question NEI-VFQ peripheral vision subscale was relatively small, suggesting that this question may have provided inadequate coverage of the subject content. The findings for our VA anchor were consonant with predicted subscales, with the largest MIDs for the VAQ acuity/spatial vision and the NEI-VFQ general vision subscales. Although in general the relevant subscales had somewhat larger MIDs, a surprising finding was the relative uniformity of MID estimates across all the VAQ and NEI-VFQ subscales (within MD/VA and cross-sectional/longitudinal sets). This pattern is consistent with the need for visual field and visual acuity for a wide variety of visual activities.

The generally larger cross-sectional MIDs for the VA anchor compared to the MD anchor may indicate that a 10-letter difference in VA is more noticeable than a 3-dB difference in
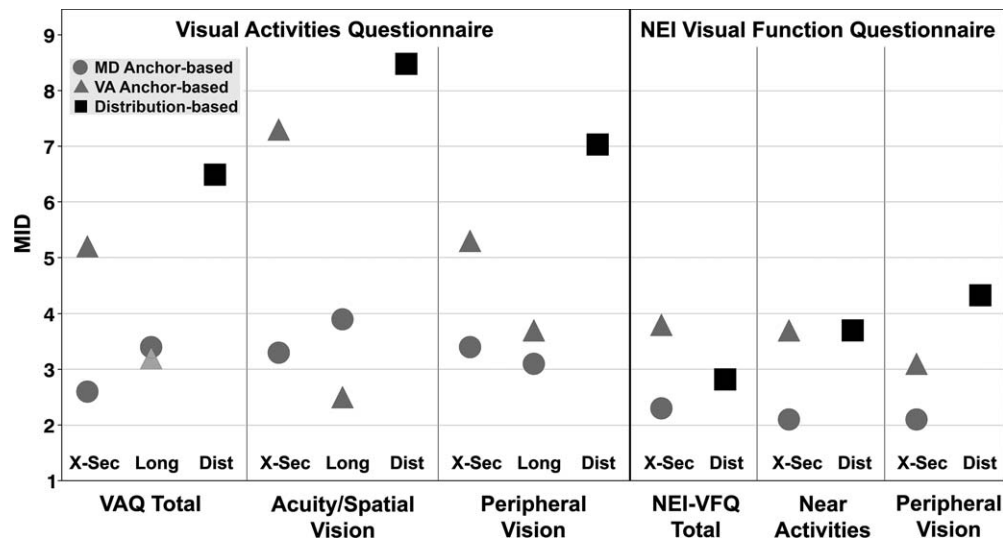
FIGURE 2. Comparison of the MIDs between anchors (MD and VA), for each quality of life measure (VAQ and NEI-VFQ) and method (cross-sectional [X-Sec], longitudinal [Long], and distributional [Dist]). MID, minimally important difference; MD, mean deviation; VA, visual acuity; VAQ, Visual Activities Questionnaire; NEI-VFQ, National Eye Institute-Visual Function Questionnaire.

MD among different people. Despite the precedent for a 10-letter loss as a minimally significant clinical difference, these results may argue that the VA anchor of 10 letters may overestimate a minimally important difference. Changing the VA anchor from 10 to 8, for example, would simply change each VA anchor-based estimate by a multiple of 0.8. Thus, values presented in this study can be easily modified to any desired alternative VA anchor. However, the longitudinal MIDs showed less difference between the VA and MD anchors, perhaps indicating that gradual loss of VA in a person is less noticeable. The concept of an MID is inherently a within-person effect, and we suggest that the longitudinal MIDs are the more appropriate measures.

McKean-Cowdin et al.,[5] in their cross-sectional evaluation of LALES participants with glaucoma, reported that a 5-dB difference in better eye MD related to approximately a 2.5-point difference in the NEI-VFQ total score. This estimate scales to an MID of 1.5 for a 3-dB difference, somewhat smaller than our MID of 2.3 for a 3-dB MD difference. In a subsequent evaluation of change data over a 4-year period, however, they found that in those with visual function loss over time, those with little initial loss of function did not show a substantial worsening of their vision-specific HR-QOL, whereas those who started with moderate loss of function and suffered further loss did show substantial worsening of their vision-specific HR-QOL. In the CIGTS data, we found a similar significant interaction between baseline level of MD severity and change in MD. Using the longitudinal (change score) anchor-based method, those subjects with worse baseline MD and a 3-dB reduction in MD over time showed greater losses in HR-QOL than those with better baseline MD. For example, for two subjects with 3-dB loss in MD over follow-up, one with baseline MD of −2 dB had an estimated MID for the VAQ total of 3.1, whereas the subject with a baseline MD of −10 dB had an estimated MID of 6.7. Three additional subscales showed significant interactions with similarly discrepant MIDs for baseline MD of −2 and −10 dB (glare disability, 3.3 and 10.4; visual search, 2.8 and 7.3; visual processing speed, 3.2 and 7.5). However, significant interactions were not found for our primary subscales of peripheral vision and acuity/spatial vision. These results give some evidence from the CIGTS that the change in MID is greater with worse baseline MD, as reported

in LALES. The VA anchor did not show any interaction effects, possibly due to the narrow range of baseline VA in our study subjects (range, 70–99; interquartile range [IQR] 82–90). Cross-sectional interaction models based on the MD anchor did not necessarily support the longitudinal model results. While our results are not definitive, they suggest that interaction effects should be explored in future estimation of MIDs.

We also compared our VA anchor-based, cross-sectional MID estimate for the NEI-VFQ total score to that found in patients followed for two years after submacular surgery for subfoveal choroidal neovascularization (the SST).[24] With a 10-letter criterion for the VA anchor, the MID estimates were 3.8 in CIGTS and 3.4 in the SST report. In a study of responsiveness of the NEI-VFQ to changes in VA in patients with neovascular AMD from the two pivotal trials of ranibizumab treatment,[25] a 15-letter change in best-corrected VA was associated with a 3.9- to 4.3-point change in the NEI-VFQ total score. Differences between these MID estimates and those from CIGTS can be explained by different anchor values (10- or 15-letter loss), different average patient levels of functional loss, and use of the cross-sectional method in CIGTS versus longitudinal in SST.

Distribution-based MID estimates take into account variability in measurements of QOL, but they lack a reference to the clinical context.[2] They are most applicable when the MID does not have to be "minimal."[1] As there is no accepted standard for selecting what portion of the QOL variability should be used in estimating the MID, we present several estimates for comparative purposes.

Several methods of estimating MID are useful for different applications or populations. Reporting a range of MIDs with SDs rather than a single number is preferred.[22] Figure 2 provides a visual depiction of the range of MID estimates across the various approaches, and SDs are given in several tables. Advice to select the higher end of the MID range to interpret score differences for a single patient versus the lower end of the range for interpreting group differences[22] should be considered.

In conclusion, MID estimates can help clinicians and researchers understand the relevance of differences in QOL scores between treatment groups or within patients over time. Statistically significant improvements that meet or exceed these MID thresholds should be considered clinically impor-

tant, keeping in mind that these are average values, and individual patients may be more or less sensitive to changes in disease or function. Although fairly robust to estimation methods, variability between our MID estimates for the same QOL scale can be attributed to the anchor chosen, whether cross-sectional or longitudinal analyses were performed, and differences between questionnaires. These MID estimates should be considered in the context of accumulating evidence from multiple studies with different types of visual disabilities.

References

1. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences. *J Clin Epidemiol*. 2008;61:102–109.

2. Eton DT, Cella D, Yost KJ, et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *J Clin Epidemiol*. 2004;57:898–910.

3. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimally clinically important difference. *Contr Clin Trials*. 1989;10:407–415.

4. Guyatt GH, Osoba U, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77:371–383.

5. McKean-Cowdin R, Varma R, Wu J, et al. Severity of visual field loss and health-related quality of life. *Am J Ophthalmol*. 2007; 143:1013–1023.

6. McKean-Cowdin R, Wang Y, Wu J, et al. Impact of visual field loss on health-related quality of life in glaucoma. The Los Angeles Latino Eye Study. *Ophthalmology*. 2008;115:941–948.

7. Patino CM, Varma R, Azen SP, et al. The impact of change in visual field on health-related quality of life. The Los Angeles Latino Eye Study. *Ophthalmology*. 2011;118:1310–1317.

8. Johns Hopkins University Evidence-based Practice Center. Treatment for glaucoma: comparative effectiveness, a review of the research for adults. AHRQ Publication No. 12-EHC038-EF. April 2012. Available at: www.effectivehealthcare.ahrq.gov/glaucomatreatment.cfm.

9. Musch DC, Lichter PR, Guire KE, Standardi CL. The Collaborative Initial Glaucoma Treatment Study (CIGTS): study design, methods, and baseline characteristics of enrolled patients. *Ophthalmology*. 1999;106:653–662.

10. Sloane ME, Ball K, Owsley C, et al. The Visual Activities Questionnaire: developing an instrument for assessing problems in everyday visual tasks. *Tech Dig Noninvasive Assess Vis Sys*. 1992;1:26–29.

11. 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.

12. Janz NK, Wren PA, Lichter PR, et al. Quality of life in newly-diagnosed glaucoma patients: the Collaborative Initial Glaucoma Treatment Study. *Ophthalmology*. 2001;108:887–897.

13. Mangione CM, Lee PP, Pitts J, et al. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ field test investigators. *Arch Ophthalmol*. 1998;116:1496–1504.

14. Frick KD, Drye LT, Kempen JH, et al. Associations among visual acuity and vision- and health-related quality of life among patients in the Multicenter Uveitis Steroid Treatment Trial. *Invest Ophthalmol Vis Sci*. 2012;53:1169–1176.

15. Chang TS, Bressler NM, Fine JT, et al. Improved vision-related function after ranibizumab treatment of neovascular age-related macular degeneration. Results of a randomized clinical trial. *Arch Ophthalmol*. 2007;125:1460–1469.

16. Cusick M, SanGiovanni JP, Chew EY, et al. Central visual function and the NEI-VFQ-25 near and distance activities subscale scores in people with type 1 and 2 diabetes. *Am J Ophthalmol*. 2005;139:1042–1050.

17. Scott IU, Smiddy WE, Schiffman J, et al. Quality of life of low-vision patients and the impact of low-vision services. *Am J Ophthalmol*. 1999;128:54–62.

18. Ferris FL III, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *Am J Ophthalmol*. 1982;94: 91–96.

19. Lichter PR, Musch DC, Gillespie BW, et al. Interim clinical outcomes in the Collaborative Initial Glaucoma Treatment Study comparing initial treatment randomized to medications or surgery. *Ophthalmology*. 2001;108:1943–1953.

20. Beck RW, Maguire MG, Bressler NM, et al. Visual acuity as an outcome measure in clinical trials of retinal disease. *Ophthalmology*. 2007;114:1804–1809.

21. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56:395–407.

22. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof*. 2005;28:172–191.

23. Janz NK, Musch DC, Gillespie BW, Wren PA, Niziol LM. Evaluating clinical change and visual function concerns in drivers and nondrivers with glaucoma. *Invest Ophthalmol Vis Sci*. 2009;50:1718–1725.

24. Submacular Surgery Trials Research Group. Evaluation of minimum clinically meaningful changes in scores on the National Eye Institute Visual Function Questionnaire (NEI-VFQ) SST report number 19. *Ophthalmic Epidemiol*. 2007;14: 205–215.

25. Suner IJ, Kokame GT, Yu E, et al. Responsiveness of the NEI-VFQ-25 to changes in visual acuity in neovascular AMD: validation studies from two phase 3 clinical trials. *Invest Ophthalmol Vis Sci*. 2009;50:3629–3635.

Investigative Ophthalmology & Visual Science