

Reliability and Validity of Refractive Error–Specific Quality-of-Life Instruments

Jason J. Nichols, OD, MS, MPH; G. Lynn Mitchell, MAS; Mario Saracino, BA; Karla Zadnik, OD, PhD

Objective: To evaluate the reliability and validity of the National Eye Institute Refractive Error Quality of Life Instrument (NEI-RQL-42) and the Refractive Status and Vision Profile survey (RSVP).

Methods: Eighty-one participants with good visual acuity (better than 20/30 best-corrected acuity in each eye) completed the NEI-RQL-42 and RSVP on 2 occasions. Noncycloplegic, subjective refractions and high-contrast visual acuity assessments were also performed. Statistical analyses addressed internal consistency, test-retest reliability, and validity (ie, concurrent and construct validity) of the 2 instruments.

Outcome Measures: The NEI-RQL-42, RSVP survey, subjective refraction, and visual acuity.

Results: The internal consistency for the overall NEI-RQL-42 was excellent (Cronbach $\alpha=0.91$); and for the overall RSVP, good (Cronbach $\alpha=0.81$). Likewise, the test-

retest reliability for the overall NEI-RQL-42 was excellent (intraclass correlation coefficient [ICC], 0.91; 95% limits of agreement, -9.1 to 10.1); and for the RSVP, fair (ICC, 0.76; 95% limits of agreement, -12.1 to 12.5). The NEI-RQL-42 overall score showed good concurrent validity as it correlated significantly with subjective refraction, whereas the RSVP overall score did not. The NEI-RQL-42 and RSVP showed similar construct validity in terms of refractive error discrimination, but the NEI-RQL-42 showed better construct validity when discriminating by the type of refractive correction used by patients. Between-instrument convergent and divergent validity was good.

Conclusions: The NEI-RQL-42 and RSVP generally have good reliability and validity in this sample of patients with refractive error. However, other factors such as content should be considered in choosing 1 of these instruments for studies of refractive error correction.

Arch Ophthalmol. 2003;121:1289-1296

QUALITY-OF-LIFE (QOL) assessments are important in ophthalmic research. These instruments complement and enhance our understanding of a patient's visual status, supplementing the traditional clinical tests used for assessment. Examples of previous ophthalmic QOL instruments include the National Eye Institute Visual Function Questionnaire, the Visual Function-14, the Glaucoma Symptom Scale, and the Graves' Ophthalmopathy Quality of Life Survey.¹⁻⁴ Recent interest in refractive surgery has led to the development of the following 2 refractive error–specific QOL surveys: the National Eye Institute Refractive Error Quality of Life instrument (NEI-RQL-42) and Refractive Status and Vision Profile survey (RSVP).^{5,6}

Quality-of-life instruments should be reliable, valid, sensitive, and responsive if

they are to be clinically useful.⁷ Although these psychometric properties are interrelated, they are examined in somewhat different ways. For instance, patients whose refractive error and vision have not changed should, theoretically, make similar responses on refractive error–specific QOL instruments each time they undergo assessment. To distinguish responsiveness (the ability to detect change in status over time) and sensitivity (the ability to detect differences among different patient groups) from measurement error, we should assess instrument variability when administered to clinically healthy patients on repeated occasions.⁸ Assessing validity consists of the accumulation of evidence over time and various studies suggesting the scales are rational and respond as predicted. Once psychometric properties are established, an instrument may be ready for use in epidemiological studies.^{9,10} The objectives of this study were

From the College of Optometry, The Ohio State University, Columbus. The authors have no relevant financial interest in this article.

to evaluate the reliability and validity of the NEI-RQL-42 and RSVP.

METHODS

SUBJECTS

Subjects recruited for the study were required to review and sign informed consent documents, which were approved by the Biomedical Institutional Review Board of The Ohio State University, Columbus. Subjects had to be 14 years or older, and without ocular abnormality (other than refractive error) or systemic conditions that could influence refractive error stability (eg, pregnancy or diabetes). Patients with best-corrected visual acuity worse than 20/30 or who had undergone ocular surgery (eg, cataract or refractive surgery) were not enrolled in this study.

OUTCOME MEASURES

The NEI-RQL-42 was scored according to the guidelines set forth by its authors.¹¹ The survey consists of 42 items used to develop 13 subscales, which are rescaled to a 100-point scale. Lower subscale scores indicate worse constructs. The subscales include clarity of vision, expectations, near vision, far vision, diurnal fluctuations, activity limitations, glare, symptoms, dependence on correction, worry, suboptimal correction, appearance, and satisfaction with correction. An overall score was calculated by averaging the subscales.

The RSVP was also scored according to the guidelines set forth by its authors.⁶ This scoring system accounts for the refractive correction used by the patient when answering and scoring the questionnaire (eg, spectacles, contact lenses, or both). Forty-five questions are used to calculate 8 subscale scores, which are rescaled to a 100-point scale. Higher subscale scores indicate worse constructs. The subscales include concern, expectations, physical and social functioning, symptoms, driving, optical problems, glare, and problems with corrective lenses. An overall score is also calculated.

A noncycloplegic, subjective refraction was performed according to standard clinical techniques. The end point of the refraction was the maximum plus sphere associated with the best Snellen visual acuity. Visual acuity was then measured at 4 m using the protocol similar to that of the Early Treatment of Diabetic Retinopathy Study.¹²⁻¹⁴ High-contrast, logarithm of the minimum angle of resolution visual acuity was measured in the right and left eyes independently and binocularly, using the results from the subjective refraction. The total number of correct responses was recorded (of a total of 70 letters).

STUDY DESIGN

Subjects first completed 1 refractive error-specific QOL instrument (the NEI-RQL-42 or the RSVP), which was randomly assigned. Following completion of the first questionnaire, subjective refraction and visual acuity were measured. Subjects then completed the second questionnaire. Each subject was asked to return for a second visit within 30 days (± 7 days) to examine the test-retest reliability of each of these outcomes. The QOL questionnaires were completed in the same order at the second visit, and a single examiner performed all measures of subjective refraction and visual acuity.

STATISTICAL ANALYSES

Statistical analyses were conducted using SAS, Version 8.2 software (SAS Institute Inc, Cary, NC). Descriptive statistics were used to summarize refractive error (the average spherical equivalent

of both eyes) and visual acuity results (letters correct) binocularly from visit 1.

Reliability

Internal consistency was assessed for each survey's subscales using the Cronbach α for data from visit 1.¹⁵ It is recommended that the α values for scales be greater than 0.70 to ensure internal consistency.¹⁶ A low α indicates that the items do not come from the same domain, as all items in each subscale should be correlated if they measure the same thing. An excessively high internal consistency (Cronbach $\alpha > 0.90$) could indicate that the items in the instrument are too highly correlated, with too many items measuring the same construct.

The test-retest reliability of the surveys, subjective refraction, and visual acuity measures were assessed using 2 methods.⁸ First, we used the 95% limits of agreement to characterize the test-retest reliability of each outcome.¹⁷ In these analyses, the mean of the differences relative to zero represents the bias between visits, and the width of the 95% limits of agreement represent the test-retest reliability of the scale.¹⁷ Second, we calculated intraclass correlation coefficients (ICC) and their 95% confidence intervals (95% CIs). It is generally recommended that the ICC exceed 0.90 if an instrument is to be used on individual patients in clinical practice and that the ICC exceed 0.70 for discriminating among groups of patients in research.¹⁸ A sample size of 80 patients provided a 95% CI lower bound of 0.70 for an ICC of 0.80.

Validity

Data from visit 1 of the study protocol were used to assess validity. We assessed concurrent validity by examining the correlation of each scale with clinical measures of vision, including refractive error and best-corrected binocular visual acuity. Construct validity was examined by known-groups validation and convergent/divergent validity.¹⁸ Known-groups validation was examined by testing for differences in survey scales by refractive error category (hyperopic, emmetropic, and myopic patients) and by the type of refractive correction (spectacles, contact lenses, or no correction) using the Kruskal-Wallis test. We assessed convergent and divergent validity using a multitrait, multimethod correlation matrix for the subscales and overall scores of the NEI-RQL-42 and RSVP. Because of the distribution of the subscale scores, nonparametric statistics were used (Spearman correlation coefficients and Kruskal-Wallis test). Because of the numerous correlations and comparisons in these analyses, type I errors were minimized by establishing a significance level of .01.

RESULTS

PATIENT CHARACTERISTICS

Eighty-one patients were recruited to participate in the study; however, 2 patients were lost to follow-up, and all analyses reflect a sample size of 79. The average \pm SD time elapsed between visits was 33.1 ± 13.1 days. The average \pm SD age of the sample was 33.3 ± 10.8 years (range, 20.9-61.5 years), and 44 subjects were female. It included 14 hyperopic patients (range, 3.94-0.50 diopters [D]), 26 emmetropic patients (range, 0.44 to -0.38 D), and 39 myopic patients (range, -0.63 to -11.19 D). Twenty-four subjects (30%) wore spectacles; 29 (37%), contact lenses; and 26 (33%), neither mode of correction. The range of binocular visual acuity in this study

was 50 to 68 letters correct (Snellen equivalent, 20/24 to 20/11, respectively) (**Table 1**).

AVERAGE SCORES

The average scores for the NEI-RQL-42 and RSVP across both visits and the mean \pm SD differences between visits can be found in **Table 2**. The only subscale for which the mean scores significantly differed between visits was the RSVP symptoms subscale; the mean of this subscale was 11.6 ± 9.9 and the mean difference between visits was -2.8 ± 9.3 (1-sample *t* test, $P = .01$).

RELIABILITY

Internal Consistency

Table 3 provides reliability estimates for the NEI-RQL-42 and RSVP. The internal consistency of the NEI-RQL-42 subscales was generally good. Three of the subscales had internal consistencies of less than 0.70, including glare (Cronbach $\alpha = 0.49$), activity limitations (Cronbach $\alpha = 0.63$), and appearance (Cronbach $\alpha = 0.65$), and 2 subscales had excessively high reliability (expectations and near vision). Also shown are values obtained from a previous report.¹⁹ The RSVP also generally showed good internal consistency. Two of the 8 RSVP subscales had internal consistencies of less than 0.70, including glare (Cronbach $\alpha = 0.60$) and problems with corrective lenses (Cronbach $\alpha = 0.52$). In a previous report, none of the subscales had an internal consistency of less than 0.70.⁶ The NEI-RQL-42 expectations and RSVP driving subscales had excessively high reliability.

Test-Retest Reliability

The test-retest reliability of subjective refraction and visual acuity was good. For subjective refraction, we found an ICC of 0.99 (95% CI, 0.99-1.00) and the 95% limits of agreement were -0.53 to 0.53 D. For visual acuity, the ICC was 0.76 (95% CI, 0.57-0.87) and 95% limits of agreement were -5.20 to 5.80 letters. This finding indicates that refractive error and vision were stable during the 1-month interval.

The overall test-retest reliability of the NEI-RQL-42 was excellent (Table 3), with an ICC of 0.91 (95% CI, 0.82-0.95). The 95% limits of agreement for the overall scale were -9.1 to 10.1 U. Six of the 13 subscales showed wide limits of agreement ($\geq \pm 25$ U), including expectations (± 35 U), diurnal fluctuations (± 27.2 U), glare (± 36 U), dependence on correction (± 36 U), worry (± 32 U), and satisfaction with correction (± 27 U). Four of the 13 subscales had ICC values less than 0.70, including activity limitations (ICC, 0.67; 95% CI, 0.44-0.82), glare (ICC, 0.64; 95% CI, 0.40-0.80), appearance (ICC, 0.49; 95% CI, 0.20-0.70), and satisfaction with correction (ICC, 0.68; 95% CI, 0.46-0.82).

The overall test-retest reliability of the RSVP was good, with an ICC of 0.76 (95% CI, 0.59-0.87). The 95% limits of agreement for the overall RSVP scale were -12.1 to 12.5 U. Two subscales had significant variability ($\geq \pm 25$ U) when we examined the 95% limits of agreement, in-

Table 1. Patient Characteristics, Refractive Error Status, Modes of Correction, and Visual Acuity

Variable	
No. of patients	81
Age, y	
Mean \pm SD	33.3 \pm 10.8
Range	20.9-61.5
No. (%) female	44 (55.7)
Refractive error, No. (%) [*]	
Hyperopia [†]	14 (18)
Emmetropia [‡]	26 (33)
Myopia [§]	39 (49)
Mode of ophthalmic correction, No. (%)	
Spectacles	24 (30)
Contact lenses	29 (37)
Neither	26 (33)
Visual acuity	
OD	
Mean \pm SD	58.4 \pm 5.5 (20/18)
Range [¶]	20/14 to 20/27
OS	
Mean \pm SD	59.5 \pm 3.8 (20/18)
Range [¶]	20/15 to 20/23
OU	
Mean \pm SD	60.3 \pm 3.1 (20/17)
Range [¶]	20/15 to 20/22

Abbreviation: D, diopter.

^{*}Indicates spherical equivalent of the right and left eye from visits 1 and 2.

[†]Defined as a spherical equivalent of at least +0.50 D of refractive error.

[‡]Defined as a spherical equivalent of less than +0.50 D but greater than -0.50 D.

[§]Defined as a spherical equivalent of at least -0.50 D of refractive error.

^{||}Expressed as number of letters correct (Snellen equivalent).

[¶]Expressed as Snellen equivalent.

cluding expectations (± 41 U) and driving (± 33 U). Seven of the 8 subscales had ICC values less than 0.70, including expectations (ICC, 0.63; 95% CI, 0.38-0.79), physical and social functioning (ICC, 0.65; 95% CI, 0.42-0.81), driving (ICC, 0.66; 95% CI, 0.42-0.81), symptoms (ICC, 0.64; 95% CI, 0.39-0.80), optical problems (ICC, 0.64; 95% CI, 0.40-0.80), glare (ICC, 0.60; 95% CI, 0.34-0.77), and problems with corrective lenses (ICC, 0.47; 95% CI, 0.18-0.69). Also shown in Table 3 are ICC values previously reported in the literature.⁶ Three subscales and the overall score had significantly worse ICC values in this study than those reported.⁶

VALIDITY

Concurrent Validity

On the NEI-RQL-42, 4 of the 13 subscales and the overall score significantly correlated with refractive error, including expectations ($r = 0.45$; $P < .001$), far vision ($r = 0.30$; $P = .007$), diurnal fluctuations ($r = 0.30$; $P = .007$), dependence on correction ($r = 0.44$; $P < .001$), and the overall score ($r = 0.39$; $P = .003$). We found no significant correlations between any of the NEI-RQL-42 scores and best-corrected visual acuity. On the RSVP, 2 of the 8 subscales significantly correlated with subjective refraction, including driving ($r = -0.29$; $P = .01$) and glare ($r = -0.28$; $P = .01$). Similar to the NEI-RQL-42, none of the RSVP scores correlated with best-corrected visual acuity.

Table 2. Overall Mean and Mean Differences for the NEI-RQL-42 and RSVP

Instrument, Subscales	Previously Reported Score	Present Study, Mean ± SD		P Value*
		Score Across Visits	Difference Between Visits	
NEI-RQL-42†				
Clarity of vision	83.8 ± 18.4	87.7 ± 14.4	0.05 ± 11.4	.97
Expectations	43.6 ± 38.2	63.6 ± 31.8	0.6 ± 17.9	.75
Near vision	83.9 ± 18.0	92.8 ± 12.5	-0.5 ± 5.8	.42
Far vision	83.5 ± 15.9	89.1 ± 10.5	-0.8 ± 7.9	.37
Diurnal fluctuations	74.6 ± 23.1	87.6 ± 17.8	1.1 ± 13.9	.49
Activity limitations	85.3 ± 21.9	94.9 ± 10.5	0.4 ± 10.3	.48
Glare	76.4 ± 26.4	84.0 ± 19.7	0.3 ± 18.3	.88
Symptoms	79.2 ± 16.8	77.3 ± 14.7	0.3 ± 10.0	.79
Dependence on correction	42.4 ± 34.6	69.8 ± 24.8	0.7 ± 18.4	.40
Worry	61.3 ± 26.0	69.2 ± 20.5	1.1 ± 16.2	.74
Suboptimal correction	92.7 ± 17.3	94.5 ± 14.1	-2.3 ± 11.8	.09
Appearance	79.3 ± 27.0	89.7 ± 16.1	2.2 ± 18.9	.30
Satisfaction with correction	74.9 ± 22.6	82.9 ± 14.9	3.3 ± 13.1	.03
Overall	NA	89.3 ± 11.0	0.5 ± 4.9	.41
RSVP‡				
Concern	44.0	16.8 ± 14.1	-1.8 ± 8.0	.05
Expectations	58.7	27.6 ± 21.7	1.6 ± 20.8	.50
Physical and social functioning	17.2	7.5 ± 11.6	0.8 ± 10.8	.51
Driving	25.9	15.2 ± 18.6	2.7 ± 16.9	.16
Symptoms	20.8	11.6 ± 9.9	-2.8 ± 9.3	.01
Optical problems	12.8	5.9 ± 9.7	0.5 ± 9.2	.61
Glare	22.3	12.4 ± 12.4	1.0 ± 12.5	.48
Problems with corrective lenses	34.0	15.1 ± 9.5	-1.4 ± 11.4	.37
Overall	26.6	13.9 ± 8.5	0.2 ± 6.3	.80

Abbreviations: NA, not applicable; NEI-RQL-42, National Eye Institute Refractive Error Quality of Life Instrument; RSVP, Refractive Status and Vision Profile survey.

*Determined by the 1-sample *t* test comparing the mean difference with 0 for bias.

†Previously reported score (mean ± SD) from Hays.¹⁹

‡Previously reported mean score from Vitale et al.⁶

Construct Validity

As a measure of known-groups validity, we compared the instruments across the 3 refractive error groups (**Table 4**). For the NEI-RQL-42, we found significant differences between refractive error groups on 5 of the 13 subscales and on the overall score. For instance, hyperopic and myopic patients had much higher expectations than emmetropic patients and reported significantly more trouble than emmetropic patients with their far vision. Hyperopic and myopic patients also reported significantly more dependence on their correction than emmetropic patients, and were also more worried about their vision than emmetropic patients. Overall, hyperopic and myopic patients had significantly worse refractive error-specific QOL than emmetropic patients.

On the RSVP, 3 of the 8 subscales and the overall score differed across refractive error categories. On this instrument, hyperopic and myopic patients were more concerned with their vision than emmetropic patients. Hyperopic and myopic patients also had significantly worse physical and social functioning than emmetropic patients, and hyperopic and myopic patients reported more optical problems than emmetropic patients. Similar to the NEI-RQL-42, hyperopic and myopic patients had worse overall refractive error-specific QOL than emmetropic patients.

As a second measure of known-groups validity, both instruments were compared across the mode of refrac-

tive error correction shown in **Table 5** (ie, spectacles, contact lenses, or neither). For the NEI-RQL-42, 6 of the 13 subscales and the overall score showed significant differences between the modes of correction. Those needing no correction believed that their clarity of vision was significantly better than those wearing spectacles or contact lenses. Those without correction also had fewer expectations regarding their vision than those who wore spectacles or contact lenses. Spectacle or contact lens wearers reported significantly worse far vision than those not needing correction; likewise, they also reported more diurnal fluctuations with spectacles or contact lens wear. Those without correction reported less activity limitations and less dependence on correction than those who wore spectacles or contact lenses. Overall, those not needing correction had significantly better refractive error-specific QOL than those who wore spectacles and contact lenses.

On the RSVP, 3 of the 8 subscales and the overall score differed significantly across modes of refractive correction. Spectacle and contact lens wearers had significantly more concern than those without refractive correction. They also had fewer symptoms than spectacle and contact lens wearers. Contact lens wearers reported significantly more optical problems than spectacle wearers and those without refractive correction. Finally, contact lens wearers had significantly worse overall refractive error-specific QOL than those without refractive correction and spectacle wearers.

Table 3. Reliability of the NEI-RQL-42 and RSVP

Outcome, Subscales (No. of Questions)	Cronbach α		Test-Retest Reliability		
	Internal Consistency	Previously Reported ^{6,19}	ICC (95% CI)	Previously Reported ICC	95% Limits of Agreement, U
NEI-RQL-42*					
Clarity of vision (4)	0.75	0.72	0.73 (0.53 to 0.85)	NR	-22.3 to 22.4
Expectations (2)	0.93	0.90	0.85 (0.73 to 0.92)	NR	-34.5 to 35.7
Near vision (4)	0.91	0.85	0.90 (0.81 to 0.95)	NR	-11.9 to 10.9
Far vision (5)	0.78	0.81	0.75 (0.57 to 0.87)	NR	-16.3 to 14.7
Diurnal fluctuations (2)	0.86	0.73	0.74 (0.54 to 0.86)	NR	-26.1 to 28.3
Activity limitations (4)	0.63	0.76	0.67 (0.44 to 0.82)	NR	-19.8 to 20.6
Glare (2)	0.49	0.75	0.64 (0.40 to 0.80)	NR	-35.7 to 36.2
Symptoms (7)	0.76	0.78	0.79 (0.63 to 0.89)	NR	-19.3 to 19.9
Dependence on correction (4)	0.79	0.74	0.76 (0.58 to 0.87)	NR	-35.4 to 36.8
Worry (2)	0.75	0.80	0.73 (0.53 to 0.85)	NR	-30.7 to 32.9
Suboptimal correction (2)	0.72	0.64	0.71 (0.49 to 0.84)	NR	-25.4 to 20.8
Appearance (3)	0.65	0.66	0.49 (0.20 to 0.70)	NR	-34.8 to 39.2
Satisfaction with correction (1)	Undefined	Undefined	0.68 (0.46 to 0.82)	NR	-22.4 to 29.0
Overall (42)	0.91	NR	0.91 (0.82 to 0.95)	NR	-9.1 to 10.1
RSVP†					
Concern (6)	0.85	0.83	0.85 (0.73 to 0.92)	0.88	-17.4 to 13.8
Expectations (2)	0.76	0.70	0.63 (0.38 to 0.79)	0.91	-39.1 to 42.3
Physical and social functioning (11)	0.86	0.87	0.65 (0.42 to 0.81)	0.63	-20.3 to 21.9
Driving (3)	0.92	0.93	0.66 (0.42 to 0.81)	0.70	-30.4 to 35.8
Symptoms (5)	0.72	0.84	0.64 (0.39 to 0.80)	0.80	-21.1 to 15.5
Optical problems (5)	0.71	0.82	0.64 (0.40 to 0.80)	0.91	-17.5 to 18.6
Glare (3)	0.60	0.75	0.60 (0.34 to 0.77)	0.72	-23.5 to 25.6
Corrective lens problems (7)	0.52	0.82	0.47 (0.18 to 0.69)	0.78	-23.7 to 21.0
Overall (42)	0.81	0.92	0.76 (0.59 to 0.87)	0.88	-12.1 to 12.5

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; NEI-RQL-42, National Eye Institute Refractive Error Quality of Life Instrument; NR, not reported; RSVP, Refractive Status and Vision Profile survey.

*Cronbach α and ICCs are previously reported in Hays.¹⁹

†Cronbach α and ICCs are previously reported in Vitale et al.⁶

A multitrait, multimethod correlation matrix (**Table 6**) presents relations between the instruments and their scales. Of the 126 potential combinations of scales between the instruments, 86 (68%) showed significant relations. The NEI-RQL-42 suboptimal correction subscale did not converge with most RSVP subscales. The RSVP expectations subscale did not converge with any NEI-RQL-42 subscale. The RSVP problems with corrective lenses subscale did not converge with the NEI-RQL-42 activity limitations, suboptimal correction, near vision, or satisfaction with correction subscale.

COMMENT

RELIABILITY

Questionnaires and their scales must be reliable to be useful in research and clinical practice. Two important forms of reliability include internal consistency and test-retest reliability. Internal consistency is high when items that constitute a scale are related to each other. For the NEI-RQL-42, we found that most subscales (77% of them) had good internal consistency (Cronbach $\alpha > 0.70$). Three subscales had poor internal consistency, including activity limitations, glare, and appearance, which also had poor internal consistency in a previous report.¹¹ Two subscales had very high internal consistencies, including expectations and near vision; the items in these subscales

might be redundant. None of the NEI-RQL-42 subscales had internal consistencies of greater than 0.90 in a previous report. For the RSVP, we found that 78% of the subscales had good internal consistency. Two subscales had poor internal consistency, including glare and problems with corrective lenses. The driving subscale had a very high internal consistency; the items in this subscale might be considered redundant. We found much lower internal consistencies for a large number of RSVP subscales than previously reported, including expectations, symptoms, optical problems, glare, and problems with corrective lenses, as well as the overall score.⁶

Subjective refraction and visual acuity were stable during the course of the study. Theoretically, we would expect that patients whose refractive error and visual acuity had not changed to provide similar responses on refractive error-specific QOL instruments each time they undergo assessment. The NEI-RQL-42 performed well in terms of test-retest reliability. Most scales had ICC values greater than 0.70, except activity limitations, glare, appearance, and satisfaction with correction. Six subscales had relatively wide 95% limits of agreement, including expectations, diurnal fluctuations, glare, dependence on correction, worry, and satisfaction with correction. The RSVP had worse test-retest reliability. Only 2 subscales had ICC values greater than 0.70, including concern and the overall score. This differs from a previous report in which all but 1 subscale (physical and so-

Table 4. Comparisons of the NEI-RQL-42 and RSVP in Differentiating Refractive Error

Outcome, Subscales	Score, Mean ± SD			P Value*
	Hyperopic	Emmetropic	Myopic	
NEI-RQL-42				
Clarity of vision	86.0 ± 18.4	92.5 ± 16.4	85.0 ± 14.9	.02
Expectations	66.1 ± 32.2†	88.5 ± 20.3	45.5 ± 27.4†	<.001
Near vision	93.9 ± 8.9	96.2 ± 10.0	90.7 ± 15.6	.24
Far vision	90.5 ± 12.6†	94.6 ± 6.8	85.7 ± 11.1†	.002
Diurnal fluctuations	91.4 ± 13.8	93.1 ± 17.8	81.5 ± 21.7	.02
Activity limitations	95.5 ± 10.2	97.8 ± 6.1	92.4 ± 14.2	.11
Glare	84.8 ± 22.0	85.6 ± 22.3	82.4 ± 20.2	.66
Symptoms	74.5 ± 20.2	77.8 ± 12.3	77.7 ± 16.9	.85
Dependence on correction	68.2 ± 34.5†	89.3 ± 17.6	56.7 ± 22.1†	<.001
Worry	66.1 ± 23.2†	77.4 ± 25.5	63.8 ± 18.1†	.01
Suboptimal correction	96.4 ± 13.4	NA	92.8 ± 18.3	.10
Appearance	86.7 ± 19.4	87.9 ± 23.9	89.7 ± 15.9	.42
Satisfaction with correction	83.1 ± 18.0	83.1 ± 20.2	79.5 ± 16.2	.45
Overall	83.4 ± 12.6†	89.5 ± 10.1	78.7 ± 9.7†	<.001
RSVP				
Concern	19.9 ± 15.7†	11.1 ± 15.4	21.4 ± 12.9†	.002
Expectations	27.9 ± 29.4	32.2 ± 26.3	22.8 ± 20.3	.29
Physical and social functioning	3.3 ± 3.6†	3.1 ± 6.2	11.0 ± 12.7†	.009
Driving	7.7 ± 11.5	10.7 ± 18.6	17.9 ± 21.1	.05
Symptoms	13.0 ± 12.7	8.6 ± 11.1	15.7 ± 10.8	.02
Optical problems	7.9 ± 12.2†	2.6 ± 8.3	6.8 ± 8.5†	.007
Glare	11.3 ± 14.5	7.0 ± 10.1	15.2 ± 12.5	.02
Problems with corrective lenses	16.1 ± 15.2	10.9 ± 13.0	16.9 ± 11.1	.26
Overall	14.0 ± 8.3†	10.3 ± 7.4	16.0 ± 8.0†	.008

Abbreviations: NA, not applicable; NEI-RQL-42, National Eye Institute Refractive Error Quality of Life Instrument; RSVP, Refractive Status and Vision Profile survey.

*Calculated by the Kruskal-Wallis test.

†Significantly different from emmetropic patients ($P < .01$).

cial functioning) had test-retest ICC values of greater than 0.70.⁶ Several of our test-retest ICC values differed significantly from those of the previous report, including expectations, optical problems, problems with corrective lenses, and the overall score.⁶ However, the inclusion of patients with only good vision (better than 20/30 visual acuity) in this study might be considered a limitation in that the generalizability of the test-retest results is limited to healthy patient samples.

VALIDITY

Although a scale may look correct (face validity) and may cover the right things (content validity), other important aspects of validity must be assessed (ie, concurrent and construct validity). A valid refractive error-specific QOL instrument should also theoretically be associated with clinical measures of vision, such as refractive error and visual acuity. For instance, individuals with high levels of ametropia should be very dependent on their correction, limited at times in their activities and functioning, and have more symptoms and problems with their vision than someone with low ametropia. Therefore, the degree of ametropia should be globally related to refractive error-specific QOL. This was true for the overall NEI-RQL-42 score, but not the overall RSVP score, using a conservative α of .01. Refractive error-specific QOL might also be related to best-corrected visual acuity, especially if individuals do not have visual acuity that is correct-

able to levels associated with good functional ability. As we limited the inclusion of individuals in this study to those with visual acuity of better than 20/30, it is not necessarily surprising that there were no correlations with the refractive error-specific QOL instrument and best-corrected visual acuity, as the 20/30 level of visual acuity is still very functional.

Another important component of validity is construct validity. An instrument measuring refractive error-specific QOL should be able to distinguish among hyperopic, myopic, and emmetropic patients. The NEI-RQL-42 and RSVP showed good construct validity in terms of their overall scores; emmetropic patients in general had fewer problems than myopic and hyperopic patients. On the NEI-RQL-42, 4 of the 13 subscales showed differences among refractive groups, including expectations, far vision, dependence on correction, and worry. In a previous report, emmetropic patients scored significantly better than myopic patients on 12 subscales and better than hyperopic patients on all 13 subscales.¹⁹ On the RSVP, 3 subscales showed differences, including concern (similar to worry on the NEI-RQL-42), optical problems, and physical and social functioning. Again, hyperopic and myopic patients scored significantly worse than emmetropic patients.

Refractive error-specific QOL instruments should also be able to distinguish between patients with varying forms of refractive correction, as each presumably has an impact on different aspects of daily living, symp-

Table 5. Comparisons of the NEI-RQL-42 and RSVP in Differentiating Patients by Refractive Error Correction

Outcome, Subscales	Score, Mean ± SD			P Value*
	Spectacles	Contact Lenses	Neither	
NEI-RQL-42				
Clarity of vision	79.5 ± 18.0†	88.8 ± 14.9†	94.3 ± 12.7	.001
Expectations	50.0 ± 31.3†	52.5 ± 28.1†	89.0 ± 25.1	<.001
Near vision	86.7 ± 17.4	95.5 ± 10.6	96.3 ± 8.2	.04
Far vision	84.2 ± 12.6†	89.2 ± 9.7†	94.8 ± 7.7	.001
Diurnal fluctuations	78.3 ± 22.1†	86.4 ± 20.6†	96.3 ± 11.3	.002
Activity limitations	89.9 ± 14.8†	95.0 ± 11.4†	99.0 ± 5.00	.001
Glare	76.6 ± 24.3	83.3 ± 20.3	91.5 ± 16.0	.04
Symptoms	73.5 ± 20.5	80.1 ± 12.9	77.2 ± 14.4	.54
Dependence on correction	57.1 ± 28.0†	60.7 ± 12.3†	91.8 ± 17.5	<.001
Worry	66.1 ± 23.7	64.6 ± 20.2	76.0 ± 20.1	.18
Suboptimal correction	97.8 ± 10.4	90.8 ± 20.2	100.0 ± 0.0	.03
Appearance	83.3 ± 22.1	92.4 ± 16.6	89.2 ± 18.8	.20
Satisfaction with correction	74.2 ± 20.0	84.7 ± 12.5	84.2 ± 19.5	.09
Overall	76.6 ± 11.7†	81.8 ± 9.1†	90.8 ± 8.9	<.001
RSVP				
Concern	21.9 ± 13.4†	20.8 ± 16.8†	10.0 ± 10.6	.003
Expectations	19.8 ± 16.5	27.9 ± 26.6	32.3 ± 26.3	.22
Physical and social functioning	6.1 ± 5.3	11.1 ± 13.3	3.4 ± 9.1	.02
Driving	15.2 ± 12.7	17.2 ± 23.5	8.7 ± 18.4	.02
Symptoms	12.3 ± 14.4†	17.4 ± 10.4†	8.4 ± 7.5	.008
Optical problems	8.1 ± 10.4	6.2 ± 8.8†	2.6 ± 8.3	.008
Glare	13.9 ± 13.6	13.8 ± 11.6	7.7 ± 12.0	.05
Problems with corrective lenses	19.7 ± 14.7	13.7 ± 7.9	10.4 ± 14.6	.14
Overall	14.8 ± 8.0	15.8 ± 7.6†	10.4 ± 8.3	.01

Abbreviations: NEI-RQL-42, National Eye Institute Refractive Error Quality of Life Instrument; RSVP, Refractive Status and Vision Profile survey.

*Calculated by the Kruskal-Wallis test.

†Significantly different from those with no correction ($P < .01$).

Table 6. Multitrait, Multimethod Correction Matrix for NEI-RQL-42 and RSVP Scores*

NEI-RQL-42 Subscales	RSVP Subscales, Correction Coefficients								
	Concern	Expectations	Physical Social Functioning	Symptoms	Glare	Corrective Lens Problems	Driving	Optical Problems	Overall
Worry	-0.70	0.02	-0.32	-0.33	-0.36	-0.26	-0.38	-0.50	-0.54
Expectations	-0.57	-0.04	-0.48	-0.16	-0.28	-0.37	-0.39	-0.45	-0.49
Activity limitations	-0.48	0.12	-0.44	-0.07	-0.18	-0.21	-0.30	-0.35	-0.35
Symptoms	-0.20	0.03	-0.27	-0.53	-0.45	-0.41	-0.30	-0.35	-0.43
Glare	-0.29	-0.10	-0.20	-0.48	-0.57	-0.28	-0.33	-0.47	-0.52
Suboptimal correction	-0.23	0.04	-0.10	-0.33	0.009	-0.26	-0.07	-0.50	-0.13
Clarity of vision	-0.57	0.08	-0.36	-0.29	-0.29	-0.53	-0.43	-0.52	-0.49
Diurnal fluctuations	-0.46	0.11	-0.31	-0.38	-0.34	-0.42	-0.43	-0.48	-0.50
Near vision	-0.45	-0.0009	-0.39	-0.11	-0.13	-0.27	-0.42	-0.39	-0.43
Far vision	-0.28	0.12	-0.26	-0.74	-0.44	-0.49	-0.74	-0.53	-0.50
Dependence on correction	-0.53	0.14	-0.46	-0.37	-0.40	-0.03	-0.33	-0.43	-0.42
Appearance	-0.35	-0.19	-0.28	-0.03	-0.22	-0.19	-0.13	-0.28	-0.30
Satisfaction with correction	-0.43	-0.01	-0.40	-0.07	-0.26	-0.18	-0.29	-0.37	-0.33
Overall	-0.79	0.03	-0.58	-0.41	-0.49	-0.43	-0.50	-0.63	-0.69

Abbreviations: NEI-RQL-42, National Eye Institute Refractive Error Quality of Life Instrument; RSVP, Refractive Status and Vision Profile survey.

*Correlation coefficient of greater than -0.28 represents a statistically significant relation between variables ($P < .01$).

toms, and functioning. On the NEI-RQL-42, we found differences between the groups on 6 of the 13 subscales, including clarity of vision, expectations, far vision, diurnal fluctuations, activity limitations, and dependence on correction, and the overall score. We detected no differences among refractive correction groups on the symptoms subscale, although it has been reported that those who wear glasses or contact lenses commonly experi-

ence symptoms.^{20,21} This finding may be related to the subscale's content. In a previous report, those requiring no correction scored significantly better than those wearing spectacles or contact lenses on 8 of the 13 NEI-RQL-42 subscales. The RSVP showed differences among groups on 4 of 9 scales in this report (overall score, concern, symptoms, and optical problems). Similar to our study, others have found differences between refractive

correction groups using the RSVP on the overall score, symptoms, and optical problems in addition to physical and social functioning and problems with corrective lenses.

An important part of construct validity is convergent and divergent relations between scales, and we found good convergent validity between the instruments. There was a strong correlation between the individual scales from each instrument and the overall scale from the alternative instrument. The RSVP expectations subscale did not converge with any NEI-RQL-42 subscale. This subscale is composed of 2 questions that pertain to an individual's acceptance of functional but less than perfect vision. Perhaps the RSVP expectations subscale assesses an area of refractive error-specific QOL that is not assessed by the NEI-RQL-42; alternatively, this domain may not be related to refractive error-specific QOL at all. It is unclear to us why this subscale behaved in this manner.

Although both of these instruments satisfy aspects of reliability and validity to one degree or another, the investigators choosing such an instrument for studies of refractive error correction should thoroughly evaluate content of the instrument they choose in addition to the other psychometric properties of that instrument. One purpose of these instruments is to predict outcomes associated with refractive surgery, and there are clearly significant changes in the NEI-RQL-42 and RSVP subscales after refractive surgery.^{11,22} However, the constructs relevant to patients before and after refractive surgery are not necessarily the same as those necessary for patients successfully wearing contact lenses and spectacles, or attempting new types of these forms of visual correction. As seen in this sample of patients with a limited range of scores, floor or ceiling effects may limit the utility of these instruments. Furthermore, the content of both instruments may not be appropriate if they are to be used in trials that include wearers of spectacles and contact lenses, in which patients are presumed to be successfully wearing these modes of refractive correction.²³ These types of issues will be important for future researchers to consider when selecting or developing an appropriate refractive error-specific QOL instrument.

Submitted for publication September 5, 2002; final revision received March 11, 2003; accepted March 25, 2003.

Corresponding author: Jason J. Nichols, OD, MS, MPH, College of Optometry, The Ohio State University, 320 W 10th Ave, PO Box 182342, Columbus, OH 43218 (e-mail: nichols.142@osu.edu).

- Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD, for the NEI-VFQ Field Test Investigators. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). *Arch Ophthalmol*. 1998;116:1496-1504.
- Lee BL, Gutierrez P, Gordon M, et al. The Glaucoma Symptom Scale: a brief index of glaucoma-specific symptoms. *Arch Ophthalmol*. 1998;116:861-866.
- Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14: an index of functional impairment in patients with cataract. *Arch Ophthalmol*. 1994;112:630-638.
- Terwee CB, Gerding MN, Dekker FW, Prummel MF, Wiersinga WM. Development of a disease specific quality of life questionnaire for patients with Graves' ophthalmopathy: the GO-QOL. *Br J Ophthalmol*. 1998;82:773-779.
- Koch DD. Measuring patient outcomes after refractive surgery. *J Cataract Refract Surg*. 2001;27:645-646.
- Vitale S, Schein OD, Meinert CL, Steinberg EP. The Refractive Status and Vision Profile: a questionnaire to measure vision-related quality of life in persons with refractive error. *Ophthalmology*. 2000;107:1529-1539.
- Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, XII: how to use articles about health-related quality of life. *JAMA*. 1997;277:1232-1237.
- Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. *Control Clin Trials*. 1991;12(suppl):142S-158S.
- Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40:171-178.
- Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1966;58:295-300.
- Hays RD, Spitzer KL. *National Eye Institute Refractive Error Quality of Life Instrument (NEI-RQL-42), Version 1.0: A Manual for Scoring and Use*. Los Angeles, Calif: Rand Health Sciences; 2002.
- Ferris FL III, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *Am J Ophthalmol*. 1982;94:91-96.
- Ferris FL III, Bailey I. Standardizing the measurement of visual acuity for clinical research studies: guidelines from the Eye Care Technology Forum. *Ophthalmology*. 1996;103:181-182.
- Gordon MO, Schechtman KB, Davis LJ, McMahon TT, Schornack J, Zadnik K, Collaborative Longitudinal Evaluation of Keratoconus (CLEK) Study Group. Visual acuity repeatability in keratoconus: impact on sample size. *Optom Vis Sci*. 1998;75:249-257.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-332.
- Bland JM, Altman DG. Cronbach's alpha. *BMJ*. 1997;314:572.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
- Fayers P, Machin D. *Quality of Life: Assessment, Analysis, and Interpretation*. New York, NY: John Wiley & Sons Inc; 2000.
- Hays RD. *Development of the National Eye Institute Refractive Error Quality of Life Instrument (NEI-RQL-42)*. Los Angeles, Calif: Rand Health and the University of California-Los Angeles; 2002.
- Begley CG, Caffery B, Nichols KK, Chalmers R. Responses of contact lens wearers to a dry eye survey. *Optom Vis Sci*. 2000;77:40-46.
- Vajdic C, Holden BA, Sweeney DF, Cornish RM. The frequency of ocular symptoms during spectacle and daily soft and rigid contact lens wear. *Optom Vis Sci*. 1999;76:705-711.
- Schein OD, Vitale S, Cassard SD, Steinberg EP. Patient outcomes of refractive surgery: the refractive status and vision profile. *J Cataract Refract Surg*. 2001;27:665-673.
- Nichols JJ, Mitchell GL, Zadnik K. The performance of the refractive status and vision profile survey in a contact lens clinical trial. *Ophthalmology*. 2001;108:1160-1166.