Functional Reading Independence (FRI) Index: A New Patient-Reported Outcome Measure for Patients With Geographic Atrophy

Miriam Kimel,1 Nancy Kline Leidy,1 Elizabeth Tschosik,2 Chantal Dolan,3 Eric H. Souied,4 Rohit Varma,5 and Neil M. Bressler6

1Evidera, Bethesda, Maryland, United States
2Genentech, Inc., South San Francisco, California, United States
3CMD Consulting, Inc., Sandy, Utah, United States
4Centre Hospitalier Intercommunal, Créteil, France
5USC Eye Institute, Keck School of Medicine, University of Southern California, Los Angeles, California, United States
6Johns Hopkins University School of Medicine, Baltimore, Maryland, United States

Correspondence: Elizabeth Tschosik, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA; tschosik.elizabeth@gene.com.
Submitted: July 20, 2016
Accepted: September 18, 2016
Citation: Kimel M, Leidy NK, Tschosik E, et al. Functional Reading Independence (FRI) Index: a new patient-reported outcome measure for patients with geographic atrophy. Invest Ophthalmol Vis Sci. 2016;57:6298–6304. DOI:10.1167/iovs.16-20361

PURPOSE. To develop and validate the Functional Reading Independence (FRI) Index, a new patient-reported outcome measure assessing reading activities in individuals with geographic atrophy (GA) due to age-related macular degeneration.

METHODS. The Index was developed through expert consultation and qualitative patient interviews. Reliability, validity, and responsiveness were tested with data from the Mahalo study (NCT01229215) of lampalizumab in patients with GA.

RESULTS. Qualitative interviews (n = 40) yielded a 10-item FRI Index, which was refined to seven items in quantitative testing (n = 100). Strong internal consistency (marginal reliability = 0.90) and reproducibility (intraclass correlation coefficient = 0.86) were shown. Known-group validity testing for baseline mean FRI Index scores showed differences (mean [SD]) between patients with Minnesota Low-Vision Reading test reading speed ≥80 vs. <80 words per minute (3.0 [0.7] vs. 1.9 [0.7]; P < 0.001), and between patients above vs. below median values on the National Eye Institute Visual Function Questionnaire-25 (NEI-VFQ-25) score (2.9 [0.7] vs. 2.1 [0.8]; P < 0.001). Convergent validity with binocular measures was strong (Spearman’s correlation = 0.72 for reading speed, 0.66 for NEI-VFQ-25). Analysis of sensitivity to change revealed mean FRI Index score changes for patients with GA lesion size growth ≥2.5 mm²/18 months of −0.41 (0.70) vs. −0.13 (0.61) for patients with lesion growth <2.5 mm²/18 months (P = 0.07).

CONCLUSIONS. The FRI Index demonstrated good reliability and validity in patients with GA. Further study in a broader GA population is warranted to confirm responsiveness.

Keywords: geographic atrophy, functional reading independence, patient-reported outcomes, visual function
second greatest aging-related concern, behind only concerns about ill health.\textsuperscript{16}

Currently, there is no approved treatment that repairs, halts, or slows the worsening of GA and any associated visual function loss. Research and clinical trials are underway to address this need, including Chroma and Spectri, which are two identically designed, multicenter, randomized, double-masked, sham-controlled phase III pivotal studies of lampalizumab, an inhibitor of the alternative complement pathway under investigation for the treatment of GA secondary to AMD.

To understand the impact of new therapies from the perspective of the patient, it is important to expand assessment of visual function beyond BCVA to include other outcomes that are relevant to patients. These include low luminance visual acuity, microperimetry, and functional measures related to reading and activities of daily living. Some patient-reported outcome (PRO) measures capture related aspects of visual functioning, such as the National Eye Institute Visual Function Questionnaire-25 (NEI-VFQ-25).\textsuperscript{17} visual function Index,\textsuperscript{18} and the activities of daily vision scale.\textsuperscript{19} However, none of the available measures were developed or validated specifically for use in patients with GA, nor do they focus solely on patients’ ability to complete everyday reading activities. A measure of a specific concept of interest in a target patient population may complement a more broad generic measure, and may be more sensitive to change than a generic measure. In addition, evidence of content validity and other measurement properties in the specific target population is recommended by the US Food and Drug Administration (FDA) for PRO measures intended to support a labeling claim.\textsuperscript{20}

This study was conducted to develop a new PRO measure of treatment impact that is appropriate for clinical trials of therapies for patients with GA. Specifically, the objectives of this study were to: (1) develop a new measure for capturing the effects of GA on functional reading activities from the patient’s perspective, the Functional Reading Independence (FRI) Index; (2) establish content validity of the FRI Index; and (3) assess its psychometric properties to demonstrate the utility of the FRI Index as a measure that captures a patient-centered construct which complements and extends other clinical measures of visual function in patients with GA.

METHODS

The research was conducted in two stages. Stage I addresses content validity through qualitative research methods, while stage II establishes the scoring and tests the psychometric properties through quantitative methods. Data from both stages of research were gathered in accordance with the tenets of the Declaration of Helsinki, with study protocols approved by an institutional review board, the Ethical Review Committee, Inc., and participants providing written informed consent prior to data collection.

Stage I: Qualitative Research for Content Validity

Methods used in the development of the new outcome measure were based on current FDA guidelines and International Society for Pharmacoeconomics and Outcomes Research good research practices.\textsuperscript{20–22} Functional reading was identified as a concept of interest for further qualitative investigation based on literature review and expert clinician input.

The qualitative research involved two studies: concept elicitation and cognitive interviews, described below. For each study, participants were recruited from US ophthalmology clinics using the following inclusion criteria: clinician-diagnosed GA (unilateral or bilateral) with no choroidal neovascu-
sensitivity to change analysis and some known-group validity assessments, which were determined post hoc based on the distribution of the data. The sample for these analyses included the PRO intent-to-treat population, which was defined as all randomized patients who received at least one treatment dose and had at least one FRI Index assessment. Analyses were performed with data from the lampalizumab and sham treatment groups combined, with the analyst blind to treatment group assignment. We used statistical software (SAS Version 9.2; SAS Institute, Cary, NC, USA) for all psychometric analyses, with the exception of item-response theory (IRT) analyses, which were performed with analytical software (Mplus; Muthén and Muthén, Los Angeles, CA, USA).

Measures

Baseline demographic characteristics were used to characterize the analytical sample. Data from the following measures at baseline and months 6 and 18 were used to evaluate and test the FRI Index: FRI Index; NEI-VFQ-25; BCVA testing based on Early Treatment of Diabetic Retinopathy Study (ETDRS) chart in the study eye; Minnesota Low-Vision Reading (MNREAD Acuity Charts) speed test in both eyes; contrast sensitivity based on the Pelli-Robson chart in the study eye; and GA lesion size measured by fundus autofluorescence in the study eye.

Statistical Analyses

We used IRT analysis to assess item performance and determine which items should be retained for the final instrument. Specifically, analyses were conducted on each item at baseline using Samejima’s graded response model, which is used to fit observed responses of subjects when item responses can be characterized as ordered categorical responses. Two types of scoring algorithms were developed: continuous (mean FRI Index score) and ordinal (FRI Level score), with the latter designed to facilitate interpretation.

Internal consistency of the FRI Index was assessed using marginal reliability—a measure of the overall reliability of the test based on average conditional standard errors, estimated at different points on the scale, for all subjects. Test-retest reliability of FRI Index scores at baseline and month 6 was assessed by an intraclass correlation coefficient (ICC) with data from a prespecified subset of stable patients (n = 27; defined as an NEI-VFQ-25 composite score change <4 points). Good test-retest reliability is indicated by an ICC ≥0.7, an ICC of 0.4 to 0.7 is moderate, and an ICC <0.4 is low. Test-retest reliability for FRI Level scores was assessed using Cohen’s weighted k statistic (i.e., the ratio of agreements to disagreements) in relation to expected frequencies and ranges between 0 (no agreement) and 1 (perfect agreement).

Convergent validity was assessed by correlating FRI Index scores (Spearman for both continuous and ordinal scoring) with GA lesion size, NEI-VFQ-25, BCVA testing, MNREAD Acuity Charts reading speed, and contrast sensitivity at baseline. Correlation coefficients >0.3 indicate acceptable convergent validity. The a priori expectation was that there would be stronger correlation with binocular measures compared with monocular measures, as subjective assessment of vision-dependent activities may be more closely associated with binocular vision. A stronger relationship with monocular measures may be expected if the monocular measure was from the better-seeing eye. However, in this study, the monocular measures were from the treated eye, which was most often the worse-seeing eye.

Known-group validity was tested using analysis of variance, comparing mean FRI Index scores of subjects grouped by GA lesion size (<4 disc areas [≈10 mm²] and ≥4 disc areas, a stratification factor for the trial); MNREAD Acuity Charts reading speed (≥80 words per minute [WPM] and <80 WPM, where 80 WPM represents minimum fluency); NEI-VFQ-25 scores (>median and <median [60.6; range, 29–96]); and BCVA (>median and <median letter score [48.0 (approximate Snellen equivalent 20/125); range, 68–20 (approximate Snellen equivalent 20/50–20/400)]). Because there were no specific interpretation guidelines for establishing known groups for the NEI-VFQ-25 or BCVA, distribution-based groupings were used. We used χ² statistics to assess FRI Level scores by stratified groups or GA lesion size and NEI-VFQ-25. Analyses were performed at baseline.

The ability of the FRI Index score to detect change was assessed by comparing changes in mean FRI Index scores from baseline to month 18 between subjects with less vs. more growth in GA lesion size. For our analysis, less growth was defined as GA lesion size gain ≥2.5 mm²/18 months, whereas more growth was GA lesion size gain <2.5 mm²/18 months. Geographic atrophy lesion size growth of 2.5 mm² is approximately equal to 1 disc area and is close to the median increase in GA lesion size over 18 months in the subpopulation of patients who completed the FRI Index (2.77 mm²). Analysis of covariance was used to examine the difference in the mean changes of the mean FRI Index scores (continuous) from baseline to the month 18 visit between less GA lesion growth and more GA lesion growth, controlling for age, sex, and FRI Index baseline score. We used χ² statistics to assess FRI Level scores by responder groups.

RESULTS

Stage I: Content Validity

A total of 23 patients participated in the concept elicitation interviews; mean age was 78 years and 16 (69.6%) were women. The average time of diagnosis was 7 years prior to enrollment (Supplementary Table S1). Ten functional reading activities were identified as problematic for nearly 50% of participants with GA (Fig. 1). These 10 problematic reading tasks formed the basis of a 10-question interviewer-administered draft FRI Index.

The draft FRI Index was administered to 17 patients with GA who participated in the cognitive interviews to refine the

![Figure 1](http://tvst.arvojournals.org/ on 01/08/2018)
The majority of patients rated most items as moderately or extremely important, ranging from 47% (reading to use the telephone) to 82% (reading signs). Patients understood the 7-day recall period and considered it acceptable. The majority of patients reported performing the functional reading activities within that time frame.

The findings from the IRT analysis supported a seven-item measure and the initial conceptual framework was revised. Based on feedback from patients, clinical experts, interviewers, and an expert in translations and linguistic validation of questionnaires, the FRI Index was finalized and prepared for implementation and validation in English-speaking patients enrolled in Mahalo (n = 100; Supplementary Table S2).

Stage II: Final Scoring and Psychometric Assessment

Final Scoring. The seven-item FRI Index yields two scores, each representing functional reading independence: the mean FRI Index score and the FRI Level score (Fig. 2). The mean score is computed using the average of the FRI item Level scores for each of the seven questions, and reflects the patient’s independence performing functional reading activities on a continuous scale. Supplementary Figure S1 shows a scoring example. Mean FRI Index scores range from 1 to 4, with higher scores indicating greater independence. The FRI Level score is an ordinal classification scheme to aid interpretation, created by rounding the mean FRI Index score to the nearest integer from 1 to 4, assigning respondents to one of four functional reading independence Levels: Level 1 (unable to do); Level 2 (help some or most of the time); Level 3 (moderately independent); and Level 4 (totally independent).

The mean (± SD) seven-item FRI Index score at baseline was 2.5 (0.8; Table 1), and decreased slightly over 18 months to 2.2 (0.9). Over time, more patients were unable to do tasks independently (FRI Level 1), and fewer patients were totally independent (FRI Level 4; Table 2).

Internal Consistency and Test-Retest Reliability. Internal consistency assessment revealed a high marginal reliability of 0.90 for the final seven-item FRI Index, demonstrating that individual items are highly related to each other and to the scale as a whole. The mean FRI Index score showed good test-retest reliability over 6 months in stable patients defined by the NEI-VFQ-25 (n = 27; ICC = 0.86). Similarly, the FRI Level score showed substantial agreement (i.e., good test-retest reliability) between the stable groups defined by the NEI-VFQ-25 (weighted κ = 0.626; P < 0.001) based on interpretation of score described by Landis and Koch. 32

Validity. As shown in Table 3, strong correlations were observed between both FRI scores (mean FRI Index score and FRI Level) and the NEI-VFQ-25 (r = 0.66 and 0.61, respectively) and maximum MNREAD Acuity Charts reading speed (r = 0.72 and 0.68, respectively). Low-to-moderate correlations were observed between both FRI scores (mean FRI Index and FRI Level scores) and GA lesion size (0.26 and 0.21, respectively). The higher convergence with binocular measures compared with monocular measures was consistent with a priori expectations.

Known-group validity at baseline for various categories are presented in Table 4. For evaluation of known-group validity for mean FRI Index scores for maximum MNREAD Acuity Charts reading speed, those reading ≥ 80 WPM had higher mean FRI Index scores compared with those reading <80 WPM (P < 0.0001). A similar trend was demonstrated between...
FRI Index scores and GA lesion size groups at baseline (P = 0.06 for ≤4 disc areas vs. ≥4 disc areas). For median-based assessment of known-group validity, differences were observed for NEI-VFQ-25 (P < 0.001), but not for BCVA ETDRS letter score (P = 0.24; Table 4). Known-group validity for FRI Levels, evaluated for groups based on GA lesion size and NEI-VFQ-25 scores followed similar trends (data not shown).

**Sensitivity.** The FRI Index score changes from baseline at month 18 for subjects with more lesion growth (n = 44, defined by GA lesion area change from baseline ≥2.5 mm²/18 months) was a mean (SD) of 0.41 (0.70) compared with −0.13 (0.61) for subjects with less lesion growth (n = 31; difference −0.28, P = 0.07 for main effect of lesion size change and P = 0.01 for overall model).

Changes in FRI Levels over 18 months also trended consistent with expectations based on changes in GA lesion size, such that a higher proportion of subjects with less growth in lesion size would be expected to remain in the same FRI Level over time, while a higher proportion of subjects with more growth in GA lesion size would be expected to move down at least one FRI Level over time. For the 31 subjects classified as having less growth at month 18 (based on GA lesion size change <2.5 mm²/18 months), only 19.4% (n = 6) declined by one or more FRI Levels by month 18, compared with 40.9% (n = 18) declining by one or more FRI Levels among the 44 subjects with more growth in GA lesion size, 21.5% difference (95% confidence interval [CI], 1.44% to 41.67%).

**DISCUSSION**

Patients with GA experience variable visual field abnormalities. As an example, patients may have good BCVA, but have impaired reading ability due to GA surrounding the fovea. As such, alternative measures of visual function beyond BCVA are needed.11 Loss of reading function is a major issue for GA patients, who may become reliant on others to complete everyday tasks. The functional reading difficulties and adaptive visual behaviors performed by these patients may not be captured by measurement of BCVA alone. This underscores the importance of considering alternative visual function endpoints to fully understand the disease and to effectively measure treatment outcomes.

To address this need, a new measure of visual function that captures the impact of GA on functional reading independence was developed for use in clinical trials. Qualitative methods were used to assure content validity of the FRI Index; quantitative data supported score reliability and validity.

Functional Reading Independence Index scores showed strong evidence of internal consistency and reproducibility. Internal consistency is based on correlations between different items on the same test and measures how well items on a test measure the same construct or idea. The IRT marginal reliability of 0.90 exceeded acceptable threshold of 0.7 for internal consistency.28 Test-retest reliability is the degree to which the results are consistent over time, and the intraclass correlation coefficient of 0.86 for the mean FRI Index scores exceeded the acceptable threshold of 0.7. The weighted κ of 0.63 for test-retest reliability of ordinal FRI Level scores also demonstrated agreement (i.e., good test-retest reliability) across time points.32

**Table 2.** Frequency of FRI Levels at Baseline and Month 18

<table>
<thead>
<tr>
<th>FRI Level</th>
<th>Baseline</th>
<th>Month 18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 94)</td>
<td>(n = 81)</td>
</tr>
<tr>
<td>Level 1</td>
<td>16 (17.0)</td>
<td>24 (29.6)</td>
</tr>
<tr>
<td>Level 2</td>
<td>28 (29.8)</td>
<td>23 (28.4)</td>
</tr>
<tr>
<td>Level 3</td>
<td>39 (41.5)</td>
<td>28 (34.6)</td>
</tr>
<tr>
<td>Level 4</td>
<td>11 (11.7)</td>
<td>6 (7.4)</td>
</tr>
</tbody>
</table>

For each patient, the FRI Level score is based on mean FRI Index scores.

**Table 3.** Convergent Validity of the Mean FRI Index Score at Baseline

<table>
<thead>
<tr>
<th>Measure</th>
<th>n</th>
<th>Mean FRI Index</th>
<th>FRI Level</th>
<th>Correlation Coefficient (P Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA lesion size, mm²</td>
<td>78</td>
<td>−0.26 (0.0205)</td>
<td>−0.21 (0.0667)</td>
<td></td>
</tr>
<tr>
<td>NEI-VFQ-25 mean score</td>
<td>94</td>
<td>0.66 (&lt;0.0001)</td>
<td>0.61 (&lt;0.0001)</td>
<td></td>
</tr>
<tr>
<td>Mean BCVA. ETDRS letter score</td>
<td>94</td>
<td>0.16 (0.1187)</td>
<td>0.15 (0.2205)</td>
<td></td>
</tr>
<tr>
<td>MNREAD Acuity Charts reading speed, WPM</td>
<td>89</td>
<td>0.72 (&lt;0.0001)</td>
<td>0.68 (&lt;0.0001)</td>
<td></td>
</tr>
<tr>
<td>Contrast sensitivity, letters read per minute</td>
<td>94</td>
<td>0.16 (0.1309)</td>
<td>0.08 (0.4212)</td>
<td></td>
</tr>
</tbody>
</table>

Spearman’s correlation coefficients based on measure distributional properties; correlation coefficients >0.3 indicate acceptable convergent validity.
Validity refers to the extent to which scores on an instrument measure or represent what they are intended to measure. Convergent validity (correlated with measures of theoretically related constructs) was evident in the moderate- to high correlations between the FRI Index score and NEI-VFQ-25 and MNREAD Acuity Charts reading speed, indicating that the FRI Index is measuring the construct intended.

Assessment of known-group validity indicates that mean FRI Index scores discriminate between groups with expected differences. Consistent with expectations, patients with higher MNREAD Acuity Charts reading speed (>80 WPM) had higher mean FRI Index scores. Mean FRI Index scores also discriminated between groups above and below the median for NEI-VFQ-25 scores. The known-group validity assessment for GA lesion size demonstrated a trend toward higher mean FRI Index scores in subjects with smaller lesions, but was not statistically significant. Likewise, known-group validity based on BCVA ETDRS letter score groups was not statistically significant. This is consistent with expectations given the limitations of BCVA as an outcome measure in GA.

Sensitivity to change is another type of validity in health outcomes measurement, and refers to the extent to which the instrument can detect change in the predicted direction when there has been a notable change in the status of a related parameter. Exploratory analyses were conducted based on a change of 2.5 mm²/18 months, representing the approximate median GA lesion growth for all patients in the study. Results of these analyses suggest that mean FRI Index score may be sensitive to change in GA lesion size. A previous analysis evaluating a different cut-point of lesion size change had similar findings (Tschosik E, et al. IOVS 2015;56:ARVO E Abstract 4789). Although the main effect was not significant, more loss of functional reading independence was seen for subjects with more lesion growth (n = 62; GA lesion area change from baseline ≥0.94 mm²/18 months) compared with subjects with less lesion growth (n = 13; GA lesion area change from baseline <0.94 mm²/18 months; mean [SD]: −0.34 [0.70] vs. −0.08 [0.47], difference −0.26, P = 1.51 main effect of lesion size change and P = 0.02 for overall model). These findings suggest that a larger study with a range of GA lesion sizes would be needed to draw definitive conclusions.

There are several limitations to this study that should be noted. First, there were relatively small sample sizes for evaluating sensitivity to change. With such small numbers of subjects in each GA lesion growth group (i.e., less lesion growth or more lesion growth), statistical inferences are limited. Second, sensitivity to change in GA lesion size and known-groups assessments for certain criteria (NEI VFQ-25 composite score and BCVA ETDRS letter score) were based on distribution-based groupings since there were no specific interpretation guidelines for clinically relevant thresholds. Third, there may be unknown confounders that were not controlled that could have influenced the change in FRI Index over time, such as low vision rehabilitation services. Fourth, participants in a clinical trial, the source of these data, may not be fully representative of all patients with GA. Although this is appropriate for testing measures designed for use in clinical trials, score performance in nontrial research or clinical settings may differ. Therefore, future research may seek to confirm the validity of the FRI Index in broader GA patient populations. In addition, future research may investigate the relevance and psychometric properties of the FRI Index in other low vision patient populations that experience reading difficulty.

In conclusion, to our knowledge, the FRI Index is the first PRO measure of functional reading independence specifically designed for patients with GA. The seven-item interviewer-administered questionnaire can be summarized with a single score, and interpreted based on maintenance or decline in four discrete Levels of functional reading independence. Qualitative methods were used to assure content validity, while quantitative tests showed evidence of reliability and validity. The FRI Index may have utility in the assessment of potential therapies on visual function in patients with GA, capturing the impact on patient independence related to reading activities. Also, as a binocular measure, the FRI Index could complement monocular measures, resulting in better understanding of real-life patient experiences. The FRI Index is being used as an outcome measure in the ongoing Chroma (NCT02247479) and Spectri (NCT02247531) phase III clinical trials investigating lampalizumab for the treatment of GA, and the observational Proxima A and B studies (NCT02479386 and NCT02399072, respectively), which seek to better characterize the relationship between GA progression and visual function. A letter of support was issued by the European Medicines Agency in January 2016 to facilitate further study of the instrument and promote data sharing.

The authors encourage other groups to use the FRI Index in order to gain better understanding of both the tool and reading independence in patients with GA.

**Note**

The FRI Index is available for use through the Mapi Research Trust (proinformation@mapi-trust.org).

**Acknowledgments**

Supported by Genentech, Inc. (South San Francisco, CA, USA). Genentech, Inc. participated in study design and conduct; and data collection, management, and interpretation. Third-party writing assistance was provided by Michael P. Bennett, PhD, of Envision Scientific Solutions, and funded by Genentech, Inc.

Disclosure: **M. Kimel**, Evidera (E, C); **N.K. Leidy**, Evidera (E, C); **E. Tschosik**, Genentech, Inc. (E); **C. Dolan**, CMD Consulting (E), Genentech, Inc. (C), Iconic Therapeutics (C), Relypsy, Inc. (C), Medivation, Inc. (C); **E.H. Souied**, Novartis (C), Allergan (C), Bayer (C), Thea (C); **R. Varma**, Genentech, Inc. (C, F), Aerie (C), Allergan (C), Bausch and Lomb (C), Genentech, Inc. (C), Research to Prevent Blindness, New York (S); **N.M. Bressler**, Bayer HealthCare Pharmaceuticals, Inc. (F), Genentech, Inc. (F), Lumenis (F), Novartis Pharma AG (F), Optovue (F), Regeneron Pharmaceuticals, Inc. (F)
References


