Test–Retest Variability of Reading Performance Metrics Using MNREAD in Patients with Age-Related Macular Degeneration

Praveen J. Patel,¹ Fred K. Chen,¹,² Lyndon Da Cruz,¹ Gary S. Rubin,¹ and Adnan Tufail¹

PURPOSE. To determine the test–retest variability of reading ability using the MNREAD charts in patients with stable age-related macular degeneration (AMD).

METHODS. In this prospective study, reading ability was measured at two visits in 124 nontreated eyes of 124 patients with AMD, who were enrolled in an ongoing clinical trial using a standardized MNREAD protocol. Only patients with stable AMD who could perform the reading test at 40 cm at both visits were included in the analysis. Different scoring rules were applied to calculate critical print size and maximum reading speed.

RESULTS. Data from the 59 patients with a mean (SD) age of 78 (7.6) years who met the study criteria were analyzed at a mean (SD) interval of 45 (6) days between measurements. The 95% coefficient of repeatability (CR) was 0.30 logMAR for reading acuity. The CR for critical print size and maximum reading speed varied depending on the analysis method applied.

CONCLUSIONS. This is a report of estimates of the intersession test–retest variability of reading performance metrics in patients with stable AMD. The results are helpful both in defining end points in clinical trials for AMD and in distinguishing patients with stable AMD, who were enrolled in an ongoing clinical trial using a standardized MNREAD protocol. Only patients with stable AMD who could perform the reading test at 40 cm at both visits were included in the analysis. Different scoring rules were applied to calculate critical print size and maximum reading speed.

Visual acuity is the most widely used measure of macular function in clinical trials and clinical practice. However, reading ability is an important component of vision function. Reading difficulty diminishes quality of life,¹,² and improvement in reading performance is one of the main objectives for elderly low-vision patients.³ The MNREAD charts, developed at the Minnesota Laboratory for Low-Vision Research, are a commonly used test in clinical trials and clinical practice to assess reading performance.⁴ The continuous-text reading charts allow measurement of the reading acuity (the smallest letter size correctly read), the maximum reading speed (MRS) and the critical print size (CPS). The reading speed should remain constant over large print sizes, forming a plateau on a plot of reading speed against print size. Measurements over this plateau can be used to define the MRS. As the print size becomes smaller, there is often a sharp demarcation, as the reading speed rapidly falls off. The smallest print size that sustains reading at the MRS is termed the CPS.

Intrinsic to any clinical investigation is test–retest variability. Recent work from our group has shown that test–retest variability of visual acuity⁵ and contrast sensitivity⁶ measures may be higher in eyes with stable age-related macular degeneration (AMD) than in those of age-matched subjects with normal vision, with repeatability approximately 10 Early Treatment of Diabetic Retinopathy Study (ETDRS) letters for visual acuity and 6 letters for contrast sensitivity in AMD patients. Although the intrasession test–retest variability of MNREAD charts has been reported in patients with low vision,⁷ there have not been any estimates of the intersession test–retest variability in patients with stable AMD. This group of patients is of great interest given the wide range of novel therapies for atrophic and neovascular AMD entering clinical trials and clinical practice. It is therefore crucial to establish the test–retest variability of reading ability in this group of patients, to define and identify disease progression or response to therapy.

Different criteria have been applied to calculate the MRS and the CPS. In the work by Virgili et al.,⁸ the CPS was defined as the print size at which subsequent smaller print sizes were read at 1.96 SD slower than the mean of the preceding print sizes (and subsequent print sizes were read at 5% lower than the mean of the logarithm of the reading speeds for the preceding print sizes). Another approach (used in a report of reading performance on Radner reading charts by patients with macular telangiectasia type 2)⁹ is to define the MRS as the single reading speed of greatest magnitude across the range of print sizes. A third alternative suggested by our group is to use the mean of the three highest reading speeds to score the MRS. In these two latter methods of calculating MRS, CPS can then be defined as the largest print size that supports reading speed at either 80% or 90% of this MRS (Table 1, Fig. 1). As there is no agreed method of defining the MRS and CPS, our group has suggested these alternative methods of MRS and CPS calculation, in an effort to simplify the analysis of MNREAD data. An alternative approach described by Cheung et al.¹⁰ involves fitting an exponential decay function to reading speed data. However, the approach that they outlined is designed to estimate MRS and CPS parameters in a cohort of patients who have reading speed functions that do not differ in shape. It is especially useful for patients with noisy or incomplete data sets. However, the method is not designed to estimate parameters for individual patients or for patients whose reading speed

From the ¹National Institutes of Health Research (NIHR) Biomedical Research Centre for Ophthalmology (Moorfields Eye Hospital and UCL Institute of Ophthalmology), London, United Kingdom; and the ²Centre for Ophthalmology and Vision Science, University of Western Australia, Perth, Australia.

Supported by The Special Trustees of Moorfields Eye Hospital. This research has received a portion of its funding from the Department of Health’s NIHR Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital and UCL Institute of Ophthalmology. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health.

Submitted for publication September 19, 2010; revised January 10 and February 21, 2011; accepted February 25, 2011.

Disclosure: P. J. Patel, None; F. K. Chen, None; L. Da Cruz, None; G. S. Rubin, None; A. Tufail, None.

No reprints will be available.

Corresponding author: Praveen J. Patel, Medical Retina Service, Moorfields Eye Hospital, 162 City Road, London EC1V 2PD, UK; praveenjpate@yahoo.co.uk.

Copyright 2011 The Association for Research in Vision and Ophthalmology, Inc.

3854
functions may differ in shape. Given the heterogeneity of our sample, we did not want to make any assumptions about consistency of shape. In addition the analysis method outlined by Cheung et al. relies on access to computer software and may not be easily available to clinicians in an office or outpatient setting.

In this work we report the intersession repeatability of MNREAD-based measures of reading performance in eyes of patients with early AMD in a clinical trial setting. We also report repeatability using the different scoring rules described for CPS and MRS. The results of this study will directly affect clinical trial design for the treatment of late AMD by reporting limits of variability of reading performance allowing the distinction of true clinical change in reading ability from measurement variability.

MATERIALS AND METHODS

Patients
Data from the untreated (better-seeing) eye of patients enrolled in the Avastin (bevacizumab) for Choroidal Neovascularization (ABC) Trial \(^{11,12}\) were used in this repeatability study. This clinical trial was a prospective, double-masked, randomized, controlled trial investigating the safety and efficacy of intravitreal bevacizumab (Avastin; Genentech, Inc., South San Francisco, CA) for the treatment of neovascular AMD \(^{11,12}\). All patients had consented to the assessment of reading performance and the research followed the tenets of the Declaration of Helsinki. In addition, approval for this research had been obtained both from the ABC Trial Steering Committee and the Research Governance Committee of Moorfields Eye Hospital. MNREAD data at baseline and week 6 visits from a total of 124 eyes of 124 patients were available for analysis. These eyes had a spectrum of disease from drusen to geographic atrophy and macular scars, due to neovascular AMD.

<table>
<thead>
<tr>
<th>MRS Method</th>
<th>Definition</th>
<th>CPS Method</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>The mean reading speed for sentences in print larger than the CPS</td>
<td>CPS defined as the print size when two criteria are met: 1. All smaller print size sentences (than the CPS) were read at a speed (log transformed) that was 1.96 times the standard deviation below the average of the previous larger print sizes. 2. All the following sentences were read 5% slower (−25% change in words per minute) than the average of the previous larger print sizes. If these disagreed, the CPS was the print size of the largest sentence.</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Mean of reading speed of fastest three sentences read</td>
<td>B1 Smallest print size* that supports reading speed at 90% of MRS</td>
<td>B2 Smallest print size* that supports reading speed at 80% of MRS</td>
</tr>
<tr>
<td>C</td>
<td>The speed of the fastest sentence read</td>
<td>C1 Smallest print size* that supports reading speed at 90% of MRS</td>
<td>C2 Smallest print size* that supports reading speed at 80% of MRS</td>
</tr>
</tbody>
</table>

* All print sizes smaller than the CPS read at a speed lower than either 90% or 80% of the MRS, as appropriate (to avoid problems with local minima).

FIGURE 1. Illustrative example of the methods used to calculate MRS and CPS (o be viewed along with Table 1). Methods A, B, and C are shown for MRS with Methods 1 and 2 for CPS.
ical examination, fluorescein angiography, and optical coherence tomography (OCT) imaging were used to detect disease progression at each visit, and eyes with active neovascular AMD or progressive atrophic AMD were excluded from the repeatability analysis, as were eyes that were unable to perform the reading test at 40 cm at both visits. Patients with disease progression were excluded from analysis as the assessment of repeatability of measurements taken over 6 weeks assumes no change in disease status over this period that caused true change in reading ability. Patients with incomplete data or with other ocular diseases that could affect reading performance were also excluded from the analysis.

**Refraction and Reading Performance Measurement**

All patients were refracted at each visit using a standardized protocol. At the baseline visit, the patient's distance viewing spectacle prescription was measured with a lensometer and was used as the beginning approximate refraction. If no spectacles for distance vision were worn, retinoscopy or autorefraction was used as a starting point for refraction. Subjective refraction was then performed refining both the spherical and cylindrical components of the refractive error to give the final correction (at 4 m) for both eyes.

After refraction, reading performance was measured by optometrists accredited for clinical trials work using a standardized protocol with the MNREAD acuity charts. These contain 19 sentences of different print sizes ranging from $1.3$ to $0.5$ logMAR with each sentence containing 60 characters. The test was performed monocularly, with the right eye tested followed by the left eye, using charts with different sentences with a reading correction added to the patient’s refraction to optimize reading at 40 cm. The same charts were used at the next visit. The patients were instructed not to change their reading distance, and the examiner watched carefully to ensure that this was the case. Chart luminance was 120 cd/m².

Patients were asked to read each sentence as quickly and accurately as possible, starting at the top of the chart when they heard the examiner say “start.” Each sentence was revealed one at a time, and the examiner used a stopwatch to time each sentence. The time taken to read the sentence was recorded in seconds, as was the number of words read incorrectly using a scoring sheet. The reading parameters were then calculated by one of the investigators (PJP) applying a range of scoring rules for MRS and CPS.

**Calculation of Reading Performance Parameters**

The reading acuity (RA) in logMAR was calculated with the formula

$$RA = 1.4 - (\text{sentences read} \times 0.1) + \left(\text{number of words read incorrectly}\times 0.01\right).$$

Reading speed was calculated for each sentence by dividing the number of words read correctly by the time taken (in seconds) to read the sentence and multiplied by 60 (to give units of words per minute; wpm). The MRS was calculated using three methods (Table 1). The CPS, defined as the smallest print size that supports the MRS, was calculated by using the methods described in Table 1.

Reading assessment formed only part of the clinical trial visit assessment. In addition to this MNREAD assessment, the patients also had visual acuity and contrast sensitivity measurement at baseline and visual acuity measurement at week 6. These were performed before reading assessment.

**Data Analysis**

Summary statistics (mean and SD) for demographic and reading performance data were calculated (SPSS, ver. 16.0; SPSS, Chicago, IL). The difference between CPS measures across the two visits was analyzed with the Wilcoxon signed-rank test, as Kolmogorov-Smirnov statistics showed that the CPS values were not normally distributed ($P < 0.05$) at the two visits. Paired $t$-tests were used to compare the RA and MRS between baseline and week 6. $P < 0.05$ was considered statistically significant.

Assessment of repeatability of reading performance parameters (RA, MRS, and CPS) was performed in line with methods outlined by Bland and Altman. For each parameter and scoring method, within-subject differences were plotted against within subject means (Figs. 2–4). We
confirmed that the differences between measurements did not increase, decrease, or change variability with the mean and that the differences were approximately normally distributed. The mean intra-subject SD ($s_w$) was used to calculate the coefficient of repeatability (CR), defined by Bland and Altman as $1.96 \times \sqrt{2 s_w^2}$ or $2.77 s_w$. The difference between two measurements for the same subject is expected to be less than the coefficient of repeatability for 95% of pairs of observations. The term $s_w^2$ is the within-subject residual mean square in the one-way ANOVA table. In addition, 95% confidence intervals (CIs) of each estimated CR were calculated. This confidence interval provides an upper and lower limit for each estimate of repeatability and depends on the sample size and the number of repeated measurements taken per subject. All patients with possible disease progression between study visits were excluded from the analysis, as were patients with incomplete data or other ocular diseases.

**RESULTS**

**Patient Characteristics and Reading Performance**

Of the 124 eyes of 124 patients with data available for analysis, reading performance data from 59 patients were included in the analysis, after excluding patients with missing data (9 patients), inability to attempt the MNREAD assessment at 40 cm (43 patients), clinical or OCT evidence of disease progression between the two measurement visits (three patients), or active CNV (four patients) or other ocular disease affecting vision (six patients). The 59 patients included 33 women and 26 men, with a median age of 78 (range, 59–88) years. There were 37 right and 22 left eyes, and the mean (SD) number of days between the test sessions was 43 (4) days, with a range of 35 to 56. The mean (SD) visual
acuity was an Early Treatment of Diabetic Retinopathy Study (ETDRS) letter score of 78 (8.5 letters), approximately 6/9 Snellen equivalent, with a median score (range) of 79 letters (44–92 letter score). The CR for visual acuity measurement in this cohort across the two visits was 8 ETDRS letters.

**Reading Performance Parameters**

The mean (SD) RA at baseline was 0.22 (0.19) logMAR and 0.23 (0.21) logMAR at the week 6 measurement. The mean (SD) difference in RA between the two visits was 0.01 (0.16) logMAR (P = 0.5). The mean (SD) difference MRS and CPS using the different scoring rules described previously are summarized in Tables 2 and 3, respectively, with no significant differences between values at the two visits. The Bland Altman plots for RA, MRS, and CPS using the different scoring rules (Figs. 2–4) showed no obvious relationship between difference and magnitude. The coefficient of repeatability (with 95% CI) for reading acuity was 0.30 (0.28–0.32). The coefficients of repeatability for MRS and CPS are presented in Tables 4 and 5. Method A yields the lowest (most repeatable) CR for MRS (Table 4) with method B2 yielding the lowest CR for CPS in this cohort of patients (Table 5). For MRS, method B was less repeatable (CR 17% greater than for method A) but with overlapping 95% CI for CR estimates) and method C proved to be the least repeatable (CR 42% greater than for method A). For CPS, Table 5 shows that methods A, B1, and C2 yielded similar CR estimates with overlapping 95% CI. However, method C1 provided the least repeatable method to calculate CPS in this cohort of patients with stable AMD (CR 52% greater than for method B2).

**DISCUSSION**

Reading ability is an important component of vision function. It has a central role in daily-living activities and work, and thus improvement in reading performance is an important objective for most elderly low-vision patients. In this work, we investigated the test–retest variability of reading performance metrics in patients with stable AMD using the MNREAD charts and applying different criteria to calculate the MRS and CPS. It is important to note that we excluded patients with disease progression from our analysis and confirmed no significant differences in the paired measurements across the two visits. We report estimated 95% coefficients of repeatability of 0.30 logMAR for RA, between 0.44 and 0.67 logMAR for CPS, and between 66 and 94 words per minute using different scoring rules for these two latter parameters. These values are higher than those reported in previous studies. We found no relationship between the differences in reading performance metrics at the two visits and the size of the measurement across the range of RA, MRS, and CPS included in the study. As most of the patients included in this analysis had mild disease with few cases of advanced AMD, these findings are most applicable to patients with early AMD.

There have been several reports in different languages of the repeatability of other methods of assessing reading performance in patients with macular disease; however, to our knowledge, there has been only one report of the use of MNREAD charts in a study. This study, by Subramanian and Pardhan, was conducted in a small cohort of patients with a range of macular diseases. The authors reported the repeatability of MNREAD measures, with repeated measures taken on the same day, by the same observer in a laboratory setting. They reported much better repeatability with 95% CR of 0.1 logMAR for RA; 0.2 logMAR for CPS, and 0.1 logRS for MRS.

Both measurement method and patient-related factors may underlie the apparent greater variability reported in the present study when compared to the results from the work by Subramanian and Pardhan. Patient fatigue and concentration may contribute to the variability of measurements in our study as reading performance metrics were assessed as part of a range of other clinical trial assessments (including other vision function assessments and imaging investigations). This better reflects a clinical trial or clinical practice setting and differs from the laboratory setting in the study by Subramanian and Pardhan. In addition, the measurements in our study were separated in time which contrasts with their study, in which measurements were taken on the same day. It is therefore not surprising that the estimates of the variability of reading performance metrics are larger in our study when compared to those in their report.

Furthermore, measurement-related factors may also have contributed to the variability of reading performance metrics. In our study, measurements were taken in the context of a clinical trial by a group of optometrists accredited for clinical trial work with no knowledge that the repeatability of measurements would be examined. The use of different examiners at different visits in our study provides an additional source of noise in the measurement of reading performance metrics. This contrasts with the previous report by Subramanian and Pardhan wherein a single observer performed all the measurements in a laboratory setting with the specific aim of

**Table 4. Coefficient of Repeatability for MRS Using Different Scoring Rules**

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>CR (wpm)</th>
<th>95% CI for CR (wpm)</th>
<th>Log Reading Speed CR</th>
<th>95% CI for CR (wpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>66</td>
<td>54–77</td>
<td>0.22</td>
<td>0.18–0.25</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>63–91</td>
<td>0.22</td>
<td>0.18–0.26</td>
</tr>
<tr>
<td>C</td>
<td>94</td>
<td>77–111</td>
<td>0.25</td>
<td>0.20–0.29</td>
</tr>
</tbody>
</table>

**Table 5. Coefficient of Repeatability for CPS Using Different Scoring Rules**

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>CR (logMAR)</th>
<th>95% CI for CR (logMAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.55</td>
<td>0.45–0.65</td>
</tr>
<tr>
<td>B1</td>
<td>0.51</td>
<td>0.41–0.60</td>
</tr>
<tr>
<td>B2</td>
<td>0.44</td>
<td>0.36–0.51</td>
</tr>
<tr>
<td>C1</td>
<td>0.67</td>
<td>0.55–0.79</td>
</tr>
<tr>
<td>C2</td>
<td>0.50</td>
<td>0.41–0.59</td>
</tr>
</tbody>
</table>
assessing and optimizing repeatability. Also, the method used to measure and record reading performance in our study differs from those used to obtain laboratory estimates. Optometrists used a stopwatch to measure the time taken to read each sentence and noted any errors made at the time of reading. This approach may be inherently less repeatable than when sentences are recorded digitally or on tape, with calculations of time taken and errors made after the test has been completed, as is done in laboratory-based estimates of repeatability, and may be an additional source of variability in our study. It may be that the performance of older patients with AMD in reading tasks was significantly influenced by such patient- and measurement-related factors when compared to younger subjects with normal vision. It is therefore particularly important to control for these factors in this patient cohort, to minimize variability, for example, by using a single observer to take readings and to take steps to minimize patient fatigue.

Although it is important to report laboratory-based repeatability estimates with repeated measurements taken on the same day by the same observer, estimates taken under these conditions may not apply to clinical trial or clinical practice settings in which patients are followed up longitudinally, often with different observers assessing vision function at different visits.

The use of alternative criteria to derive the CPS and MRS was an attempt to simplify the calculations of these parameters over the methods outlined by Virgili et al. and also used in the study by Subramanian and Pardhan (referred to as method A in our study). Defining the MRS based on the mean speed of the fastest read three sentences (method B) or just the fastest single speed (method C) and then further defining the CPS as the largest print size that supports reading at either 90% (method 1) or 80% (method 2) of the MRS would be a simpler approach than method A. The single fastest speed has been used as the MRS in a recent study of reading performance in patients with macular telangiectasia type 2. Our results suggest the method B2 may yield improved repeatability of the CPS over methods A and C, although potentially at the price of reduced repeatability of MRS with method B over method A. Further studies are needed to investigate the impact of alternative scoring rules on repeatability of CPS and MRS in patients with macular disease.

In summary, this is the first study to report the intersession repeatability of reading performance using MNREAD charts in a large cohort of patients with age-related macular degeneration in a clinical trial setting. The values obtained in this study may be used to guide future clinical trial design and provide insight into the repeatability of reading performance measures in this important group of patients.

References