How Sensitive to Clinical Change are ETDRS logMAR Visual Acuity Measurements?

Daniel A. Rosser,1,2 Simon N. Cousens,3 Ian E. Murdoch,1,2 Fred W. Fitzke,1 and David A. H. Laidlaw4

PURPOSE. To determine the sensitivity to change and specificity achieved when published test–retest variability (TRV) data are used to determine whether measured changes in ETDRS logarithm of the minimum angle of resolution (logMAR) visual acuity reflect true clinical change or are attributable to measurement error alone.

METHODS. Various degrees of change in visual acuity were simulated in a group of normal subjects by adjusting test difficulty through manipulation of viewing distance. Sensitivity to simulated change and specificity were determined with change criteria derived from published Bland-Altman 95% ranges for TRV.

RESULTS. The relationship between viewing distance and measured acuity was as predicted theoretically. Simulated acuity change of 0.2 logMAR (two lines of letters) or greater can be reliably distinguished from no change (both sensitivity and specificity >95%) with the ETDRS chart, but a change of 0.1 logMAR or less cannot.

CONCLUSIONS. The use of 95% ranges for TRV to establish the smallest measured visual acuity change that can be reliably detected ensures a high specificity but does not take account of sensitivity. Use of change criteria derived from published 95% ranges results in a sensitivity of approximately 50% (assuming identical levels of TRV). Sensitivity may be improved by using a change criterion that is smaller than the minimum change sought, providing the change criterion is still at least as large as the 95% range for TRV, so that specificity is maintained. Reducing TRV allows smaller changes in acuity to be reliably detected. (Invest Ophthalmol Vis Sci. 2003;44:3278–3281) DOI:10.1167/iovs.02-1100

Visual acuity data are subject to a degree of measurement error. When an individual subject is tested and subsequently retested, the resultant visual acuities tend to differ, even in the absence of any true clinical change. This effect is referred to herein as test–retest variability (TRV). The clinician monitoring a patient over time therefore must disregard apparent changes resulting from TRV, but must recognize any change that reflects a genuine alteration in clinical status. Failure to do the former results in unwanted false-positive results (i.e., loss of specificity), whereas failure to do the latter compromises the test’s sensitivity to change.

The clinician may draw on experience to make this judgment or may refer to published data on the repeatability and TRV of visual acuity tests to establish a change criterion, beyond which measured change is considered to be genuine and below which it is not. The method of Bland and Altman has been widely used to establish such a change criterion.1–8 The Bland-Altman approach9 involves establishing the upper limit of measured change that can be expected when a clinically stable individual undergoes two visual acuity measurements. The 95% range for TRV is generally used to identify the change criterion against which measured differences are judged. The 95% range for TRV is defined as ±1.96 SD of the differences between paired measurements on a group of subjects. Measured changes that lie outside this range are considered to be “real” changes. This approach ensures that, in the absence of true clinical change (and assuming the TRV to be normally distributed), only 5% of measured differences will exceed this level—that is, the approach seeks to fix the specificity of the procedure at 95%. Various 95% ranges have been published, from ±0.07 to ±0.20 log of the minimum angle of resolution (logMAR; see Table 1).

The Bland-Altman method provides the change criterion that is necessary to fix specificity at a level of 95%. However, another important characteristic of a diagnostic test, sensitivity, is not considered. To our knowledge, there has been no investigation of acuity test sensitivity or how sensitivity to change and specificity are influenced by the change criterion used. This may be in part because measuring sensitivity to change of a visual acuity test is difficult, in that there is no independent gold-standard method of establishing whether true clinical change has occurred and, if so, precisely how much. To overcome this problem, we used one of the advantages of contemporary logMAR visual acuity charts: The difficulty of the chart can be adjusted with absolute precision by altering the viewing distance. For example, the 0.7 logMAR letters on a chart viewed at 4 m are precisely equivalent to the 1.0 logMAR letters viewed from 8 m. Hence, although we cannot alter an individual’s acuity by an exact amount, we can alter, with great precision, the difficulty of the task that the test presents.

The purpose of our study was to determine the level of sensitivity of the ETDRS logMAR chart to simulated change and the specificity achieved when using change criteria derived from published TRV data. Various degrees of acuity change were simulated by altering the difficulty of the test through manipulating the viewing distance. Experimental conditions were designed to avoid any factors that were thought to have the potential to increase TRV, such as uncorrected refractive error, ocular disease, variable lighting conditions, and multiple examiners.
**MATERIALS AND METHODS**

**Subjects**
Fifty subjects were recruited from the staff of Moorfields Eye Hospital and the students of City University, London. The tenets of the Declaration of Helsinki were adhered to in full. Inclusion criteria were: age less than 50 years; absence of any ocular abnormality, including media opacity; ability to understand and comply with the testing protocol; and Snellen acuity of 6/9 or better.

One eye of each subject was assessed. When both eyes met the inclusion criteria, the right eye was used as the study eye. All subjects wore full distance refractive correction, as established by formal subjective refraction immediately before the study measurements were taken.

**Equipment**
The ETDRS logMAR chart (Lighthouse International, New York, NY) is taken. The ETDRS logMAR chart (Lighthouse International, New York, NY) is based on the design suggested by Bailey and Lovel16 but incorporates the recommendations of the U.S. National Academy of Sciences–National Research Council (NAS-NRC).1,17 The chart has been described in detail by Ferris et al.12

**Design.** The letter stimuli are printed on a translucent panel and lit from behind. The chart has five letters per row ranging in size from +1.0 to −0.30 logMAR at 4 m. There are three versions of the chart, differing only in the letter combinations used.

**Testing Paradigm.** Subjects were required to identify each letter on the chart until they identified a full row of letters incorrectly, at which point the test was terminated, and the acuity calculated.

**Scoring.** Interpolated acuity measurements were derived using the method described by Ferris et al.12 In this method, visual acuity is scored by the letter rather than by the line, as is common in clinical practice. The resultant reduction in the size of the scale increment has been shown to reduce the level of TRV.3,6

**Testing Procedure**
Each subject underwent acuity measurements at five different distances: 4.0, 4.5, 5.0, 6.3, and 8.0 m. Taking the 4-m distance as the reference, each increase in distance increased the difficulty of the test by a precise amount (Table 2). The increments in relative difficulty were chosen to cover the range of TRV data reported in the literature. An additional measurement was conducted at the baseline distance of 4 m, to allow the test’s specificity to be calculated. The six measurements were performed in random order, using a randomly selected ETDRS chart, to prevent learning effects from influencing the results. As the six ETDRS measurements were conducted using three charts, the only proviso to this was that the same ETDRS chart not be used for two consecutive measurements.

**Analysis**

**Categorizing Subjects.** For a given degree of simulated change (or for no simulated change when both measurements were conducted at the same distance), subjects were categorized as having changed in visual acuity if the measured acuity change exceeded the relevant change criterion, otherwise they were categorized as having not changed in acuity. The change criteria investigated were those that could be identified from the published literature (95% ranges for TRV, Table 1).

**Sensitivity to Change.** Sensitivity was defined as the percentage of individuals having undergone a simulated change whom the test correctly identified as having a change.

**Specificity.** Specificity was defined as the percentage of individuals who, on the basis of the two test results at 4 m, were not identified as having a change.

**RESULTS**

Of the subjects recruited, 15 (30%) were emmetropes, 29 (58%) myopes, and 6 (12%) hypermetropes. The refractive error (spherical equivalent) ranged from −8.00 to +2.00 with a mean of −1.51 D. Eight subjects (16%) had astigmatism of 0.75 D (maximum 1.75 D) or more.

The distribution of measured acuity change for each degree of simulated change (0, 0.05, 0.10, 0.20, and 0.30) was assessed for normality with the Shapiro-Wilk W-test. There was no evidence (P > 0.05) of departure from a normal distribution for any of the degrees of simulated change.

To assess the validity of the technique of simulating acuity change through manipulation of the viewing distance, acuities measured at 4.5, 5.0, 6.3, and 8.0 m were subtracted from that measured at the reference distance of 4 m. Table 3 shows the mean differences in measured acuity at each distance, alongside the differences theoretically expected (Table 3). The 95% confidence interval for the means in the table should be distinguished from the 95% ranges referred to elsewhere in this article and described in the introduction. In contrast to the range of values obtained, the 95% confidence interval is an index of the precision of our estimate of the population mean. The confidence interval allows us to assess whether the measured differences in acuity are compatible with those theoretically expected. The observation that, in each case, the 95% confidence interval includes the expected value (Table 3) indicates that the relationship between measured acuity and viewing distance is compatible with that predicted from theory.

From the paired measurements taken at 4 m, the 95% range for TRV was calculated to be ±0.11 logMAR according to the
**DISCUSSION**

We measured the specificity and sensitivity of the ETDRS chart to various degrees of simulated acuity change. When combined with the Bland-Altman approach to determine a setting-specific change criterion of 0.11 logMAR, the ETDRS performed well at detecting changes of 0.2 logMAR (estimated sensitivity of 100%, 95% confidence interval [CI] 93%–100%; estimated specificity 90%, 95% CI 86%–99.5%). However, for an acuity change of 0.10 logMAR, test sensitivity was poor (38%, 95% CI 25%–53%). Furthermore, this study used normal subjects wearing full refractive correction, and all the measurements were taken by a single examiner in a short time under identical conditions, using an interpolated scoring method. These are all factors that may have helped to limit the degree of TRV in this study, and may therefore expect the test to perform less well in day-to-day use. Subsequent to the main analysis, the data were rescored using the line-assignment method favored in routine clinical practice (defined as the logMAR value of the smallest letter size at which at least three of five letters were correctly identified). Aucities scored this way were subject to greater TRV (95% range ±0.11 logMAR).

**SUMMARY**

The degree of visual acuity change that may be reliably detected using the ETDRS chart may be not as small as previously thought. This is because the usual method of establishing a change criterion (using Bland-Altman 95% ranges for TTV) measures the size of change required to fix specificity at 95%, but does not consider sensitivity. Our results suggest that using the 95% TRV range as the change criterion

### Table 4. Specificity and Sensitivity for Each Degree of Simulated Change Using Change Criteria Derived from Published 95% Confidence Ranges for TRV

<table>
<thead>
<tr>
<th>Change Criterion</th>
<th>Sensitivity (0.05 change)</th>
<th>Sensitivity (0.10 change)</th>
<th>Sensitivity (0.20 change)</th>
<th>Sensitivity (0.30 change)</th>
<th>Specificity (0 change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.07±0.5</td>
<td>58 (43–72)</td>
<td>60 (45–74)</td>
<td>100 (93–100)*</td>
<td>100 (93–100)*</td>
<td>84 (71–93)</td>
</tr>
<tr>
<td>0.09±0.5</td>
<td>44 (30–59)</td>
<td>50 (36–65)</td>
<td>100 (95–100)*</td>
<td>100 (95–100)*</td>
<td>90 (78–97)</td>
</tr>
<tr>
<td>0.10±0.5</td>
<td>18 (9–31)</td>
<td>38 (25–53)</td>
<td>100 (95–100)*</td>
<td>100 (95–100)*</td>
<td>96 (86–100)</td>
</tr>
<tr>
<td>0.11±0.5</td>
<td>4 (0–14)</td>
<td>12 (5–24)</td>
<td>66 (55–79)</td>
<td>100 (95–100)*</td>
<td>100 (95–100)*</td>
</tr>
<tr>
<td>0.16±0.5</td>
<td>2 (0–11)</td>
<td>4 (0–14)</td>
<td>54 (40–68)</td>
<td>98 (89–100)</td>
<td>100 (95–100)*</td>
</tr>
<tr>
<td>0.18±0.6</td>
<td>2 (0–11)</td>
<td>4 (0–14)</td>
<td>54 (40–68)</td>
<td>98 (89–100)</td>
<td>100 (95–100)*</td>
</tr>
</tbody>
</table>

Data in parentheses are 95% confidence intervals.

* One–sided 97.5% confidence intervals.

† TRV derived from the present study.
results in a low level of sensitivity. Sensitivity may be improved by using a change criterion that is smaller than the minimum change sought, providing the change criterion is still at least as large as the 95% range for TRV, so that specificity is maintained. Reducing TRV allows smaller changes in acuity to be reliably detected. This can be achieved clinically by reducing the size of the test's scale increment.\textsuperscript{1,4,6}

**References**