Consistency Between Visual Acuity Scores Obtained at Different Test Distances

Theory vs Observations in Multiple Studies

Li Ming Dong, PhD; Barbara S. Hawkins, PhD; Marta J. Marsh, MS

Objective: To investigate the consistency of visual acuity (VA) scores measured at 2 different distances in patients with or at risk for choroidal neovascularization.

Methods: Best-corrected VA scores measured at 2 distances for the same eyes at the same examinations were collected from 4 sets of randomized clinical trials among patients with or at risk of choroidal neovascularization. Within each trial, the pairs of VA scores were compared and their relationship was explored.

Results: After adjustment for test distance, VA scores obtained at the closer distance were found to be systematically lower than those obtained at the farther distance in all data sets. In the Submacular Surgery Trials pilot study, the average discrepancy between 2- and 0.5-m VA scores was 7.5 letters. In an ancillary study of the Macular Photocoagulation Study, the discrepancy between 10-ft and 5-ft VA scores was 3.1 letters. In the Laser to Drusen Trial pilot study, the discrepancy between 3.2- and 1-m VA scores was 7.3 letters. In the Treatment of Age-Related Macular Degeneration With Photodynamic Therapy Study, in which the VA scores at the closer test distance were censored, the estimated discrepancy between 2- and 1-m VA scores was 8.2 letters. Reduction in visual angle at closer test distance did not explain the discrepancy completely. Features of the macular lesion, poor accommodation of the elderly population with age-related macular degeneration, or the test charts did not account for the discrepancies.

Conclusion: The VA scores at distances less than 2 m were lower than expected in all 4 studies. The observed discrepancy was consistent with findings from a study among healthy young subjects, suggesting that the phenomenon is real and common.

Arch Ophthalmol. 2002;120:1523-1533

IN OPHTHALMOLOGIC patient care and research, particularly in retinal and choroidal diseases that affect central vision, visual acuity (VA) or change in VA is often used to measure the effectiveness of interventions and to monitor adverse reactions. A survey of multicenter randomized clinical trials in ophthalmology in the United States sponsored by The National Eye Institute (NEI), National Institutes of Health, found that VA or change in VA was the primary or secondary outcome measurement in 23 (96%) of 24 sets of clinical trials surveyed (unpublished survey, completed August 2002). Typically, VA is measured monocularly by testing a patient's ability to recognize letters printed on a chart using first one eye and then the other. Various test charts have been used in clinical practice and in clinical studies, such as the Snellen chart, Sloan charts, or Bailey-Lovie charts. The Snellen chart has been used in the clinic setting for more than 100 years. During the past 20 years, the Early Treatment Diabetic Retinopathy Study (ETDRS) charts (modified Bailey-Lovie charts) have been adopted for many clinical research studies because of their good properties in facilitating quantitative use of VA test results and wide availability.

Ferris et al incorporated design features of the original Bailey-Lovie charts to develop a new set of charts for use in the ETDRS, a multicenter randomized controlled clinical trial sponsored by the NEI. As shown in Figure 1 (1 of the 3 charts in the set), the optotypes are a subset of 10 letters from the English alphabet; each chart has 14 rows, with 5 letters on each row. Letters on the same row were selected to have approximately the same level of difficulty. The size of letters is reduced at a proportion of \( \rho = 10^{-1/0.7943} \), equivalent to a reduction to one-half in size, every 3 lines down the chart. The spacing between letters in each row is the same as the size of letters. The spacing between rows is the same as the size of let-

From the Wilmer Ophthalmological Institute, School of Medicine, The Johns Hopkins University, Baltimore, Md.
ters on the lower row. The 14 rows permit VA measurement across a range of 1.4 logMAR (ie, logarithm of the minimum angle of resolution) values; each letter read can be considered to represent a −0.02 change in logMAR.²

Although the ETDRS charts and the original Bailey-Lovie charts used different subsets of letters and typology, other features and the logMAR scale are the same. Like any other VA test chart, the ETDRS chart can be used to measure VA only within a restricted range at a given test distance because of the physical constraint of the size of the chart. To accommodate the full range of VA of research subjects, or the full range of VA for the same subject over time, it is often necessary to use more than one VA test distance, especially when evaluating retinal diseases where, for some eyes, VA could begin at 20/50 and deteriorate in 2 years to 20/800 and, for other eyes, VA could deteriorate from 20/400 to 20/1600 in 2 years. The VA test scores obtained at different test distances are then converted to a common scale for comparison. It is important that the VA score obtained at one distance represent the same VA as the score obtained at a different distance after adjustment for test distance. Inconsistency in VA scores obtained at different distances would lead to bias in estimating treatment effect, if it exists, when VA or change in VA is the outcome variable or in assessing changes in the vision of an eye in the clinical care setting.

The design of the ETDRS chart allows the number of letters read at one distance to be converted to equivalent numbers of letters read at another distance by adding or subtracting a constant number. The constant is determined on the basis of the principle that the same visual angle represents the same vision regardless of the test distance. Because the size of letters in each line of ETDRS charts is reduced to half every 3 lines down the chart, the visual angle of the first line at 4 m, for example, is the same as the visual angle of the fourth line at 2 m; the visual angle of the second line at 4 m is the same as the visual angle of the fifth line at 2 m, and so forth. Therefore, if a test subject is
moved to half of the original test distance, he or she is expected to read 3 more lines, or 15 more letters, with the same eye. The number of letters read at 4 m can be converted to the 2-m equivalent by adding 15 letters.

The above conversion rule has been used in many clinical trials and epidemiologic studies whenever different VA test distances were used with the ETDRS charts or Bailey-Lovie charts. According to the conversion rule, scores obtained at different distances theoretically should translate to the same VA fraction. However, no publications from the chart developers have been found to support this theoretical assumption. The purpose of this article is to examine the consistency of VA scores obtained at different test distances with the original and modified Bailey-Lovie charts using data from 4 different studies.

Data used in this analysis were collected as part of routine procedures using standard protocols in 4 multicenter studies among patients with choroidal neovascularization (CNV) in age-related macular degeneration (AMD), the ocular histoplasmosis syndrome, and idiopathic causes: the Submacular Surgery Trials (SST), pilot study (3 trials), Macular Photocoagulation Study (MPS), Laser to Drusen Trial (LDT) pilot study, and Treatment of Age-Related Macular Degeneration With Photodynamic Therapy (TAP) investigation. All 4 studies were conducted to compare treatment interventions with observation in managing AMD or CNV.

Details of the studies have been published except those of the LDT pilot study. In all 4 studies, VA was tested at 2 distances when VA was in a range near the minimum that could be measured at the longer test distance. Available pairs of VA scores from both test distances were extracted from study records to compare the VA scores obtained from the 2 test distances. The design, objective of each study, and details of extraction of paired VA scores from each study are described later in this section.

Statistical analysis was conducted separately for each study. Paired t test and linear regression analysis were performed to test the consistency of and to explore the relationship between the 2 VA scores obtained at different test distances. Analysis of censored observations was conducted for data from the TAP Investigation in which the 1-m VA scores were censored.

For convenience, the number of letters read correctly at a specified test distance was used as the unit of analysis unless otherwise specified.

DESIGN AND METHODS OF THE SST PILOT STUDY

The SST pilot study is a set of 4 clinical trials comparing surgery with observation in managing subfoveal CNV in AMD, ocular histoplasmosis, and idiopathic CNV with the sponsorship by the NEI and others. The goal of the pilot study was to explore the feasibility of full-scale trials with the same ocular conditions. Data included in this analysis were from 3 SST pilot trials, from which the full-scale SST was developed and which are ongoing. Best-corrected VA was tested monocularly with Bailey-Lovie charts using data from 4 different studies.

In the SDL pilot study, one eye of each patient, usually the only eye with an eligible lesion, was designated as the study eye. The VA scores from the study eyes tested at both distances at any visit were extracted for analysis from the SST pilot study database. Pairs in which the 2-m VA score was 0 letters were excluded. According to the VA test procedure, this set of study eyes read 1 to 14 letters correctly at 2 m. From this data set, a data set including only the first pair of VA scores from each study eye also was extracted for analysis.

STUDY DESIGN AND METHODS OF THE MPS

The MPS was a set of 8 randomized multicenter trials sponsored by NEI and conducted between 1979 and 1993. The purpose of the trials was to evaluate laser photocoagulation for preventing or delaying severe vision loss in eyes with CNV secondary to age-related macular degeneration, the ocular histoplasmosis syndrome, or idiopathic causes. The primary endpoint was severe (6 or more lines) loss of VA from the baseline level. The VA was tested monocularly with best correction by means of the original Bailey-Lovie charts.

The MPS was a set of 8 randomized multicenter trials sponsored by NEI and conducted between 1979 and 1993. The purpose of the trials was to evaluate laser photocoagulation for preventing or delaying severe vision loss in eyes with CNV secondary to age-related macular degeneration, the ocular histoplasmosis syndrome, or idiopathic causes. The primary endpoint was severe (6 or more lines) loss of VA from the baseline level. The VA was tested monocularly with best correction by means of the original Bailey-Lovie charts.

Two test distances were used in the MPS: 10 ft (3 m) and 5 ft (1.5 m). An eye was tested first at 10 ft; the line with the smallest letters on which all were read correctly and the number of smaller letters read correctly were recorded. Whenever an eye could not read all letters on the top line of the chart correctly at 10 ft, the patient was moved to 5 ft from the test chart and the test was repeated after adjustment of the correction for refractive error. The line with the smallest letters that was read correctly at 5 ft and the number of the smaller letters read correctly were recorded. When the eye could read all the letters on the top line correctly at 10 ft, the line with the smallest letters read correctly and additional letters read on the smaller (lower) line at 10 ft were recorded. The lines and the letters were converted to 5-ft equivalent by adding 3 lines, as an eye was expected to read 15 additional letters when moving from 10 to 5 ft. Then lines and letters were converted to letter score for this analysis by multiplying the number of lines by 5 and adding the number of additional letters.

In the MPS, only 1 VA measurement (smallest line and additional letters at 1 test distance) was recorded for each eye, although VA was tested at both distances for some eyes. Between February 1984 and June 1985, an ancillary study was conducted to investigate the consistency of VA scores from the 2 distances. During this period, vision examiners at a subset of MPS clinical centers recorded both measurements for eyes that were able to read between 1 letter and 4 complete lines (20 letters) on the chart at 10 ft. An average of a 3-letter (0.6-line) difference between the VA measurements from the 2 test distances was reported at the Second NEI Symposium on Eye Disease Epidemiology, June 1985; however, the findings were never published. For our exploration, we retrieved pairs of VA measurements from microfilm for the subgroup of MPS eyes whose VA was tested and recorded at both distances. A data set including only the first pair of VA scores per patient was constructed for analyzing the discrepancy and exploring the influence of lesion size. Comparison of VA scores at both test distances was also performed for the data set including all available pairs of measurements.

(REPRINTED) ARCH OPHTHALMOL/VOL 120, NOV 2002 WWW.ARCHOPHTHALMOL.COM

©2002 American Medical Association. All rights reserved.
DESIGN AND METHODS OF THE LDT PILOT STUDY

The LDT pilot study was a randomized clinical trial conducted in preparation for a full-scale study. The objective of the study was to evaluate laser photocoagulation of eyes with drusen with the goal of reducing the incidence of CNV. Patients eligible for the LDT pilot study were patients with large drusen, focal hyperpigmentation, no neovascular AMD (CNV, disciform scar, or laser-treated area), and VA equivalent to 20/40 or better in the study eye, who had evidence of neovascular AMD in the fellow eye. Patients were randomized to observation and 2 laser treatment protocols at a ratio of 2:1:1. Although the outcome of interest was the incidence of CNV in the study eye, best-corrected monocular VA was measured for both eyes at 3 major examinations: enrollment (baseline), 12-month, and 24-month examinations. Best-corrected monocular VA also was measured in study eyes at 3 additional follow-up examinations: 3, 6, and 18 months. The VA testing protocol for this study was similar to that in the SST pilot study described earlier, except that the test distances were 3.2 and 1 m for this trial. The VA test began at 3.2 m from the test chart with one eye occluded. The total number of letters read correctly at this distance was recorded. When an eye read fewer than 15 letters at 3.2 m, the eye was tested again at 1 m, with adjustment of the refractive correction. The total number of letters read correctly at 1 m was recorded as the VA score. When 15 or more letters were read at 3.2 m, no VA measurement at 1 m was required. The number of letters read at 3.2 m was converted to the same scale as the 1-m score by adding 25 letters.

The LDT pilot study enrolled 99 patients total. Pairs of VA scores (total number of letters read correctly) at both test distances were extracted from the LDT pilot study database for analysis; pairs with a 3.2-m score of 0 letter were excluded. The range of visual acuities corresponding to 1 to 14 letters read at 3.2 m was equivalent to 20/400 to 20/200 in Snellen notation.

DESIGN AND METHODS OF THE TAP INVESTIGATION

The TAP investigation consisted of 2 randomized, placebo-controlled clinical trials sponsored by QLT Inc (Vancouver, British Columbia), Novartis Ophthalmics AG (Bu{ch}, Switzerland), and Novartis Ophthalmics (Duluth, Ga). The objective of the study was to determine whether photodynamic therapy with verteporfin could safely reduce the risk of VA loss in patients with subfoveal CNV caused by AMD. The primary end point was the proportion of study eyes with fewer than a 15-letter loss in VA from baseline after 1-year follow-up in the eye identified as the study eye at the time of enrollment.

Monocular best-corrected VA was tested with ETDRS charts at one or both of 2 test distances: 2 and 1 m. An eye was tested at the 2-m distance first. Whenever the eye read fewer than 20 letters at 2 m, the eye was tested for the ability to read only the top 3 lines at 1 m after correction of refractive error for the closer distance. The VA scoring method in the TAP investigation was as follows: when the eye read fewer than 20 letters at the 2-m distance, the final VA score was the sum of the numbers of letters read correctly at both distances; when the eye read 20 letters or more at the 2-m distance, the final VA score was the total number of letters read correctly at 2 m plus 15. Pairs of nonzero VA scores (number of letters read correctly) at both test distances were extracted for analysis.

METHOD USED TO ACCOUNT FOR CENSORING OF TAP INVESTIGATION SCORES AT 1 M

Because eyes were tested only for ability to read on the top 3 lines (15 letters) of the charts at the 1-m distance, the 1-m VA scores recorded were in fact censored at 15. To examine the distribution of the censored VA scores at 1 m, the percentile of the recorded 1-m VA scores was plotted against that of a standard normal distribution (quantile-quantile [QQ] plot, not shown). Points in this plot fell approximately along a line with some deviation at the upper end that probably was due to the censoring, suggesting that the uncensored 1-m VA scores may be approximated by a normal distribution. If we assume the uncensored 1-m VA scores were normally distributed, then the mean and SD of the distribution can be estimated. Let Y be a normally distributed random variable with mean µ and SD σ, and Y is a random variable censored at c from above, that is, Y=Y*, if Y*>c; Y=c, if Y*≤c. The mean of Y (ie, expected value, or E[Y]), derived in the same way as in theorem 22.3 in Greene,9 is:

\[ E[Y] = c[1 - \Phi(\alpha)] + \Phi(\alpha)|\mu + \sigma\lambda|, \]

where \( \Phi(\alpha) = \Phi\left[\frac{(c - \mu)}{\sigma}\right] \), \( \Phi(\alpha) \) is the probability of censoring, \( \alpha = (c - \mu)/\sigma \), \( \lambda = -\Phi(\alpha)/\Phi'(\alpha) \), and \( \Phi'(\cdot) \) and \( \Phi(\cdot) \) are the density and distribution functions, respectively, for a standard normal variable. Since \( \Phi(\alpha) \) can be estimated by the percentage of the uncensored 1-m VA scores, 0.55, then \( \alpha = -1(0.55) = 0.13 \). Use the average of the censored 1-m VA scores, 13.3, to estimate \( E[Y] \) in the above equation, replace \( \alpha \) by 0.13, and let \( c = 15 \), then \( \mu \) and \( \sigma \) can be solved from the above equation and the relationship 0.13 = (15 - \mu)/\sigma.

Tobit analysis10 was performed to estimate the linear relationship between the uncensored 1-m VA scores and the 2-m VA scores with SAS 6.12 software (SAS Institute Inc, Cary, NC).

RESULTS

FINDINGS FROM THE SST PILOT STUDY

A total of 224 study eyes (of 327 patients) were found to have both a 2- and a 0.5-m VA score for at least 1 pilot study examination. A study eye in this data set may have multiple pairs of VA scores from different study examinations, but only 1 pair (the first pair) of VA scores per study eye were included in our analysis. The VA range corresponds to approximately 20/500 to 20/250 with Snellen notation.
Figure 2 displays the number of letters read correctly at 0.5 m (y-axis) vs the number of letters read correctly at 2 m (x-axis). Because an eye is expected to read 30 more letters when moved from 2 to 0.5 m (assuming that the refractive error has been corrected for the closer distance), the data points are expected to cluster around the line $y = x + 30$ (the dotted line). However, most of the data points are below this line, indicating that the VA score measured at 0.5 m is lower than expected in comparison with the VA score measured at 2 m.

Table 1 summarizes the VA scores obtained from 2 test distances. The average VA score at 0.5 m is 7.5 letters lower than the 2-m presumed equivalent VA score. The 0.5-m scores also have larger variation than 2-m scores, because the range of 2-m scores was constrained between 1 and 14 letters, while the range of 0.5-m scores was unrestricted up to 70 letters.

A regression line was fitted to the number of letters read at 0.5 m (as the dependent variable $y$) vs the number of letters read at 2 m (as the independent variable $x$). The regression line obtained (shown as the solid line in Figure 2) is $y = 20.4 + 1.3x$. No quadratic or higher order of nonlinear relationship was detected. The 95% confidence intervals for the slope and the intercept of the regression line are 1.1 to 1.5 and 18.6 to 22.3, respectively. The observed relationship between the number of letters read at the 2 test distances is significantly different from the expected theoretical relationship $y = 30 + x$ ($P < .001$, Hotelling $T^2$ test). Figure 2 also shows that the discrepancy between the 2-m VA score and the 0.5-m VA score is larger in the poorer vision range: the mean (median) difference between 2- and 0.5-m VA scores was 8.6 (8) letters for eyes that read 1 to 7 letters correctly at 2 m and 6.8 (5) letters for eyes that read 8 to 14 letters correctly at 2 m.

The same analysis was carried out for a larger data set that included all pairs of VA scores at both test distances from any study examination. From 327 patients, 569 pairs of VA scores from 224 study eyes were available at both 2 and 0.5 m from study examinations. Although observations in this data set may be correlated because of the inclusion of multiple pairs of VA scores from the same study eye at different examinations, results were almost identical to the above findings based on a data set with uncorrelated observations. The distribution of the 2-m VA scores, converted to theoretically equivalent 0.5-m score, had a mean (median) of 36.0 (35) letters with an SD of 3.7 letters. The distribution of the 0.5-m VA scores had a mean of 28.3 (29) letters with an SD of 8.6 letters. The average VA score at 0.5 m was 7.7 letters (95% confidence interval, 7.1-8.3 letters) lower than the 2-m presumed equivalent VA score. The estimated regression line between 0.5-m VA scores ($y$) and 2-m VA scores ($x$) was $y = 20.3 + 1.3x$. The closeness of the results to those based on independent observations may suggest that the correlation between VA scores from the same eye may be negligible for purposes of our analysis. Of course, the SEs of the estimates based on the full data set were smaller than those based on the subset that included only 1 pair of VA scores per patient, as they were estimated from a larger sample.

**FINDINGS FROM THE MPS ANCILLARY STUDY**

A total of 487 pairs of VA measurements from 293 eyes of 295 patients with juxtafoveal or subfoveal lesions were...
identified. Analysis was carried out for the set of 259 independent pairs of measurements, one pair from each of 259 patients. The VA range of this group of patients corresponds to approximately 20/250 to 20/80 by means of Snellen notation.

Figure 3 displays the number of letters read correctly at 5 ft vs the number of letters read at 10 ft. The dotted line $y = 15 + x$ is the expected (theoretical) relationship between the VA scores at these 2 test distances, where $y$ stands for number of letters read at 5 ft, and $x$, the number of letters read at 10 ft. A trend similar to that from the SST pilot study was found from the MPS ancillary study data, although not as obvious: most of the points are below the dotted line, indicating that VA scores obtained at 5 ft were lower than VA scores obtained at 10 ft. A summary of the VA scores at the 2 test distances from the MPS ancillary study is displayed in Table 1. On average, VA scores measured at 5 ft were 3.1 letters lower than VA scores measured at 10 ft after adjustment for test distance.

The regression line for the observed data is $y = 9.9 + 1.2x$, where $y$ is the number of letters read at 5 ft and $x$ is the number of letters read at 10 ft. Again, no obvious nonlinear relationship was detected. The 95% confidence intervals were 8.5 to 11.2 and 1.1 to 1.4 for the intercept and slope, respectively. The observed line is statistically significantly different from the expected line $y = 15 + x$ ($P < .001$, Hotelling $T^2$ test). Figure 3 also shows that the discrepancy between the 2 measurements is larger for eyes with poorer vision than for eyes with better vision.

A parallel analysis was also carried out for the full data set: multiple pairs of VA measurements from one patient were included whenever available, whether paired VA scores of an eye from multiple examinations or paired VA scores from both eyes. A total of 487 pairs of VA measurements were analyzed. The estimated mean difference in VA scores and the regression line based on this full data set were almost identical to those based on the data set with independent pairs of VA scores: eyes tested at 5 ft read about 3.2 letters less than at 10 ft after adjustment for test distance. The estimated intercept and the slope from linear regression were 9.8 and 1.2, respectively. Again, the similarity of the results suggests that correlation between the multiple pairs of measurements from the same patient in the full data set are negligible for the purpose of our analysis. The SEs of the 3 estimates based on the full data set were smaller than those based on the subset that included only 1 pair of VA scores per patient.

FINDINGS FROM THE LDT PILOT STUDY

A total of 85 pairs of VA scores from 43 eyes of 42 patients were identified that met the criterion that the 3.2-m VA score be in the range of 1 through 14. The VA range corresponds to approximately 20/400 to 20/200 by means of Snellen notation. Among 85 pairs of VA scores, 82 pairs of VA measurements were from the fellow eyes, ie, the eyes with neovascular AMD. Only 3 pairs were from the study eyes, because most of the study eyes had VA better than 20/160 throughout the study.

Figure 4 displays a scatterplot of the number of letters read correctly at 1 m vs the number of letters read correctly at 3.2 m. The dotted line is the expected theoretical relationship between the VA scores from these two distances, $y = 25 + x$, where $y$ is the number of letters read at 1 m and $x$ is the number of letters read at 3.2 m. As observed in the SST pilot study and MPS ancillary study, most of the data points in Figure 4 are below the expected line, indicating that the VA score at 1 m tends to be lower than the VA score at 3.2 m after conversion to the same scale. Table 1 displays the VA scores tested at 3.2 m and at 1 m from the LDT pilot study. On average, eyes read 7.3 letters less at the 1-m distance than they did at 3.2 m after conversion.

The regression line for the observed data is $y = 12.8 + 2.0x$, where $y$ is the number of letters read at 1 m and $x$ is the number of letters read at 3.2 m. The 95% confidence interval for the intercept is 10.5 to 15.1 and for the slope is 1.6 to 2.4. Multivariate $t$ test (Hotelling $T^2$) suggests that the observed line differs from the expected line $y = 25 + x$ to a statistically significant degree ($P < .001$). This line has a steeper slope than that found in either the SST pilot study or the MPS ancillary study.
To rule out the effect of potential correlations among pairs of VA scores from the same patient, the analysis was repeated for a subset of 42 pairs of VA scores from 42 eyes, each from different patients. The regression line obtained is \( y = 13.9 + 1.9x \), and the mean difference between the VA scores measured at 2 distances was 6.5 (SD, 7.6), very close to those from the full data set. Therefore, we view the correlation among multiple pairs of measurements from the same patient to be negligible.

**FINDINGS FROM THE TAP INVESTIGATION**

Among 609 patients enrolled in the TAP investigation, 2387 pairs of VA measurements (number of letters read correctly) made at 2 test distances were identified. These VA measurements came from 552 eyes of 422 patients. The VA range corresponds to approximately 20/500 to 20/250 by means of Snellen notation.

Unlike the other studies discussed, the VA score at the 1-m distance was censored. That is, even if an eye might have read more than 15 letters at the 1-m test distance, the number of letters read at that distance was still recorded as 15. It is not known how many letters would have been read had the patient been allowed to read below the top 3 lines of the chart; 45% (1070/2387) of 1-m VA scores were “15 letters” in the database.

**Figure 5** is a scatterplot of the paired VA measurements from the TAP investigation. Because the 1-m VA scores are censored at the horizontal line \( y = 15 \), mean and SD for the uncensored 1-m VA measurements could no longer be estimated with the sample mean and SD as in the other 3 studies. However, assuming the uncensored 1-m VA scores to be from a normal distribution, we estimated the mean and the SD of uncensored 1-m scores. The QQ plot (not shown) of the observed 1-m VA scores against a standard normal suggested that the normal assumption was reasonable. The QQ plots of VA scores at the closer test distances for the other 3 studies (not shown) indicated that the (uncensored) VA scores at the closer distance were reasonably well approximated by a normal distribution.

The summary statistics for VA scores at 2-m and uncensored VA scores at 1-m in the TAP investigation are presented in Table 1. The VA scores at 2 m were converted to equivalent VA scores at the 1-m test distance by adding 15 letters. The mean difference between the 2- and 1-m VA scores was estimated by the difference in the means of the 2- and 1-m VA scores. On average, VA scores at 2 m were about 8.2 letters larger than VA scores at 1 m. The general trend of lower 1-m VA scores was also apparent in Figure 5: for given values of 2-m VA scores where less than 50% of the 1-m VA scores were censored (1-6 letters read at 2 m), the medians of 1-m VA scores were well below the expected line. The SD of the difference between the paired VA scores cannot be estimated.

A line also was fitted to the data with a further assumption that the distributions of 1-m VA scores at each given value of the 2-m VA scores were also approximately normal with constant variance. The linear relationship estimated through the Tobit model was \( y = 10.4 + 0.5x \), where \( y \) is the number of letters read at 1 m without censoring and \( x \) is the number of letters read at 2 m. These estimates of the regression line, the population mean, and SD for the uncensored 1-m VA scores from the TAP investigation rely heavily on the normality and homoscedasticity assumptions and on the percentage of censoring as well. With as much as 45% censoring in the data, the estimated regression line should be interpreted as an approximation to the overall trend.

**CONSISTENCY OF FINDINGS ACROSS STUDIES**

In all 4 sets of randomized clinical trials, VA scores at a shorter test distance were, on average, lower than those expected in comparison with scores at a longer distance on the basis of theoretical assumptions. **Table 2** summarizes the relationships between the VA scores from 2 test distances estimated from observations in the 4 data sets for the SST Pilot Study, MPS Ancillary Study, LDT Pilot Study, and TAP Investigation.

<table>
<thead>
<tr>
<th>Study</th>
<th>Ratio of 2 Test Distances</th>
<th>Discrepancy Letters†</th>
<th>Observed Relationship</th>
<th>Expected Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST Pilot study</td>
<td>4</td>
<td>7.5</td>
<td>( y = 20.4 + 1.3x )</td>
<td>( y = 30 + x )</td>
</tr>
<tr>
<td>MPS ancillary study</td>
<td>2</td>
<td>3.1</td>
<td>( y = 9.9 + 1.2x )</td>
<td>( y = 15 + x )</td>
</tr>
<tr>
<td>LDT pilot study</td>
<td>3.2</td>
<td>7.2</td>
<td>( y = 12.8 + 2.0x )</td>
<td>( y = 25 + x )</td>
</tr>
<tr>
<td>TAP investigation</td>
<td>2</td>
<td>8.2</td>
<td>( y = 10.4 + 0.5x )</td>
<td>( y = 15 + x )</td>
</tr>
</tbody>
</table>

*SST indicates Submacular Surgery Trials; MPS, Macular Photocoagulation Study; LDT, Laser to Drusen Trials; TAP, Treatment of Age-Related Macular Degeneration With Photodynamic Therapy; \( y \) is the number of letters read at the shorter test distance; and \( x \) is the number of letters read at the longer test distance.†Average of visual acuity scores at the shorter test distance minus visual acuity scores at the shorter test distance.
sets. The discrepancy varied from 3.2 letters to 8.2 letters, depending on the pair of the test distances used in the study.

The observed discrepancies in VA scores in the first 3 studies are consistent with each other in the sense that the discrepancy increases with the ratio of the longer test distance to the shorter test distance. The VA testing protocols for these 3 studies were very similar except for the test distances used.

The smallest discrepancy observed was 3.2 letters from the MPS data set. The MPS scoring algorithm was different from that in the SST pilot study and the LDT pilot study. In the MPS, the VA score at both distances was the line with smallest letters read perfectly plus additional smaller letters read. Letters larger than the line of smallest letters read completely were counted as having been read correctly, even when some were missed, as long as 1 or more letters were read on each line of the chart. This procedure could reduce the discrepancy between scores at 10 and 5 ft.

The discrepancy between VA scores measured from 2 test distances was estimated to be 8.2 letters by means of TAP data, in comparison with the 3.2-letter discrepancy observed in the MPS ancillary study, although in both studies the longer test distance was twice the shorter test distance. The larger estimated discrepancy in the TAP investigation is likely due to the censoring mechanism of the 1-m VA scores. Recall that in the TAP investigation the eye was tested with the use of only the top 3 lines (15 letters) of the test charts, while in the MPS (and the other 2 studies), eyes were tested with the whole chart at the closer test distance. To estimate the average uncensored VA score at 1 m in the TAP investigation, we treated any 1-m VA score of 15 letters as censored. The VA scores less than 15 letters at 1 m were considered uncensored. In fact, VA scores less than 15 at 1 m also may have been censored in the TAP investigation. For example, a VA score of 14 letters at 1 m could represent an eye that read only 1 letter incorrectly on the 3 lines of the chart. If the patient had been allowed to read more lines with that eye, as permitted in the other 3 studies, he or she might have read another 2 letters on the fourth line, ie, the VA score would have been 16 letters instead of 14 letters. Because of the way the VA test was conducted, censoring of this type could not be recognized. The unrecognized censoring of this type might cause the mean and SD of the uncensored VA score at 1 m to be underestimated. Consequently, the size of the discrepancy between the 2- and 1-m VA scores may be overestimated with the data from the TAP investigation.

The observed linear relationships between VA scores at 2 test distances exhibited consistency in 3 of 4 studies, although a departure from the expected linear relationship was observed across all 4 studies. As shown in Table 2, the slope of the regression line from the MPS ancillary study data (1.2) was about the same as that from the SST pilot study data (1.3). The slope of the regression line in a smaller study, the LDT pilot study, was somewhat larger (2.0). Importantly, the lines were steeper than the theoretically expected lines in all 3 trials, and the intercepts of the observed lines were lower than those of the expected lines in all 4 studies. Interestingly, the ratio of the observed intercept to the expected intercept in each study was nearly constant: 0.68 in the SST pilot study, 0.65 in the MPS ancillary study, 0.51 in the LDT pilot study, and 0.69 in the TAP investigation.

The observed linear relationship between VA scores made at 2 test distances in the TAP investigation was the only one that differed from the other 3 studies. The observed line (Table 2) was flatter than the theoretically expected line, possibly explained by undercounted censoring of 1-m VA scores less than 13, as in the example mentioned earlier. The high 1-m VA scores, say 11 to 14, were more likely to be censored scores than the low 1-m VA scores, say 1 to 4, and could not be recognized, thus flattening the regression line.

The discrepancy between VA scores from 2 test distances was observed in all 4 studies. We explored possible factors that may contribute to the observed discrepancy between scores at the 2 test distances.

**EFFECT OF LESION SIZE**

Because almost all data came from eyes with subfoveal or parafoveal lesions, an immediate question is whether the discrepancy is associated with a CNV lesion, which often results in a central scotoma and greater use of eccentric vision. To explore the effect of lesion size, we defined the discrepancy for each pair of VA scores to be the observed VA score at the longer distance minus the VA score at the shorter distance.

Analysis of correlation between discrepancy and lesion size was carried out with the use of the SST pilot study and the MPS ancillary study databases. No strong linear or nonlinear association was found between the discrepancy and the size of the lesion. For the SST pilot study, the Pearson correlation coefficient was $r = 0.12$ ($P = 0.09$) and the Spearman correlation coefficient was $r = 0.09 (P = 0.20)$ for 203 pairs of VA scores from 203 study eyes for which lesion size was available. For the MPS ancillary study, the Pearson correlation coefficient was $r = -0.03 (P = 0.61)$ and the Spearman correlation coefficient was $r = 0.04 (P = 0.48)$ (248 eyes, 1 eye per patient).

Because almost all VA scores used in our analysis of the LDT pilot study data were from fellow eyes, no lesion size data were available. For the TAP investigation, the same analysis could not be carried out because the discrepancy for each pair of VA scores cannot be defined with the censored 1-m VA scores.

**EFFECT OF VISUAL ANGLE**

The theoretical rule for converting VA scores measured at different distances described earlier assumes implicitly that the same visual angle will result in the same VA score after adjustment for test distance. When an eye is tested at a pair of distances, for example, 0.5 and 2 m, as in the SST pilot study, one may ask whether the visual angle of line 1 at 2 m is the same as the visual angle of line 7 at 0.5 m on the ETDRS charts. The exact visual angle of letters on any line of the ETDRS chart at a given test distance can be calculated by means of the length of chart, the test distance, letter size of that line, and the location of that line (details of calculation are available...
from the authors; also available at http://www.jhu.edu/wctb/pub_supps/va2dist_appendix.pdf). Figure 6 illustrates the visual angle of line \( i \) when VA is tested at 2 m and the visual angle of line \( i+6 \) when VA is tested at 0.5 m; the eye is assumed to be leveled with the midpoint of the ETDRS chart (24.25 cm from the top of the first row).

When the test distance changes from 2 to 0.5 m as in the SST pilot study, each eye is expected to read 6 additional lines. Table 3 presents the calculated minimum angle of resolution in logMAR units for 3 pairs of lines at 2 and 0.5 m. Table 3 also gives letter size \( l_i \) and the distance from the midpoint of line \( i \) to the midpoint of the chart \( h_i \), which are needed to calculate the visual angle of line \( i \).

The logMAR at 2 m (column 4 in Table 3) for lines 1, 2, and 3 agree well with the expected logMAR obtained with the conversion rule. However, values of logMAR at 0.5 m (column 8 in Table 3) are all lower than their corresponding values in column 4, suggesting that at 0.5 m the visual angles for lines 7, 8, and 9 are slightly smaller than the visual angles for lines 1, 2, and 3 at 2 m. The reductions in logMAR, or visual angles, at 0.5 m (the last column of Table 3) translate into reading 1.5, 2, and 3 letters less than 2-m VA scores. The reduced visual angle at 0.5 m may account in part for the lower VA scores at 0.5 m compared with VA scores at 2 m observed in the SST pilot study. However, it does not explain completely the observed total discrepancy of 7.5 letters. In addition, the last column in Table 3 suggests that the discrepancy is larger for better vision (line 3 at 2 m), although it was observed in the SST pilot study, the MPS ancillary study, and the LDT pilot study that the discrepancy was larger for the poorer scores.

In summary, the visual angle at 0.5-m test distance is reduced in comparison with longer test distances and may contribute to the lower observed VA scores at 0.5 m in the SST pilot study. However, these observations would not account for a significant portion of the discrepancy.

### Table 3. Calculated LogMAR for Lines in the ETDRS Chart at 2 Test Distances

<table>
<thead>
<tr>
<th>Line</th>
<th>( l_i ), cm</th>
<th>( h_i ), cm</th>
<th>LogMAR at 2 m</th>
<th>Line</th>
<th>( l_i ), cm</th>
<th>( h_i ), cm</th>
<th>LogMAR at 0.5 m</th>
<th>Difference, LogMAR</th>
<th>Difference, No. of Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.818</td>
<td>21.344</td>
<td>1.309</td>
<td>7</td>
<td>1.461</td>
<td>14.486</td>
<td>1.281</td>
<td>0.03</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>4.621</td>
<td>11.503</td>
<td>1.213</td>
<td>8</td>
<td>1.161</td>
<td>16.958</td>
<td>1.169</td>
<td>0.04</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3.671</td>
<td>3.685</td>
<td>1.114</td>
<td>9</td>
<td>0.922</td>
<td>18.921</td>
<td>1.058</td>
<td>0.06</td>
<td>3</td>
</tr>
</tbody>
</table>

*LogMAR indicates logarithm of the minimum angle of resolution; ETDRS, Early Treatment Diabetic Retinopathy Study; \( l_i \), letter size at line \( i \); and \( h_i \), distance from the midpoint of line \( i \) to the midpoint of the chart.

In all 4 studies analyzed in this investigation, VA scores obtained with original or modified Bailey-Lovie charts at test distances shorter than 2 m were less than predicted from scores at longer test distances based on the
properties of VA test charts with logMAR scales. The numbers of pairs of VA measurements analyzed ranged from 85 in the smallest data set to 2387 in the largest data set. Because most of the eyes in the analysis had juxtafoveal or subfoveal neovascular lesions or scars, the effect of lesion size was evaluated but was not found to be related to the size of discrepancies in vision. Difference in the size of the visual angle at different test distances explained a portion of the discrepancy.

Our findings are from patients who had CNV in the macula. However, a few articles11-13 reported similar phenomena in small studies among subjects without CNV. A larger study was reported by Giese14 in 1946. Findings from these studies were summarized by Sloan.15 In Giese’s study, 400 young (median age, 19 years; age range, 17-36 years) healthy subjects with at least 20/20 VA were tested at 10, 5, and 1 m and closer distances. The VA was measured by means of a multiple-choice checkerboard illuminated with 8 foot-candles and reported in visual decimal (Snellen fraction).

Table 4 displays the average VA reported in Giese’s article at 1-m and closer distances. Giese’s data showed a steady drop in average VAs as the test distance decreased below 1 m (Figure 7). We selected results from 3 pairs of test distances in Giese’s study and compared them with findings from the SST pilot study, MPS ancillary study, and LDT pilot study, respectively: 1 vs 0.25 m, 1 vs 0.5 m, and 1 vs 0.33 m. These 3 pairs of test distances, although not the same pairs of test distances used in the 3 studies, were selected because the ratio of the 2 test distances within each pair is the same or nearly the same as in the 3 studies, respectively (4 for the SST pilot study, 2 for the MPS ancillary study, and 3.2 for the LDT pilot study). From Table 4, the difference in average VA (in logMAR) between 1 and 0.25 m is $-0.212 - (-0.021) = -0.191$, ie, VA measured at 0.25 m would have translated to about 9.5 letters fewer than VA measured at 1 m. In comparison, the 0.5-m VA score was 7.5 letters fewer than the 2-m VA score in the SST pilot study. The same calculation was performed for the other 2 pairs of test distances. For the 1- and 0.5-m pair, the difference is $-0.069$ in logMAR or 3.5 letters fewer at 0.5 m. In the MPS ancillary study, VA scores at 5 ft (1.5 m) were 3.2 letters fewer than VA scores at 10 ft (3 m). For the 1- and 0.33-m pair, the difference obtained from Giese’s study is $-0.105$ in logMAR, about 5 letters’ difference. In comparison with the LDT pilot study, the observed difference in VA scores was 7.3 letters. All 3 VA differences from the 3 pairs of the test distances in Giese’s study were close to observed differences in VA scores in the SST pilot study, the MPS ancillary study, and the LDT pilot study.

Despite the similarities with our findings, 2 caveats are appropriate. First, as pointed out earlier, the pairs of test distances selected from Giese’s report were not the same as those used in the 3 studies. The discrepancy between VA scores obtained from 2 test distances may depend not only on the ratio of test distances but also on the test distances themselves. Second, VA was presented in Giese’s article on the visual decimal scale, which is the decimal value of Snellen fraction. Therefore, his average VA score was the average of VA in Snellen fraction. Although we can convert VA in Snellen fraction to equivalent logMAR by taking log10 of the reciprocal of Snellen fraction, the converted average Snellen fraction to logMAR is not the same as the average VA in logMAR. The former is the arithmetic mean and the latter is the geometric mean, which is no greater than the arithmetic mean. The comparable average VA from Giese’s study would have been the average of transformed visual decimal, which could not be calculated without the original data. Despite the 2 caveats, the comparable results from Giese’s study and the 3 studies under comparison represent at least a rough approximation of the VA discrepancy at different test distances.

Giese’s data suggest that the phenomenon we have described is real, common, and not restricted to eyes with central macular lesions for which the patients rely on eccentric fixation to read test charts. Furthermore, the young age of the subjects in Giese’s study would appear to eliminate the factor of difficulties with accommodation at closer distances after correction of refractive errors appropriate for the closer distances used.

In randomized trials, the discrepancy that may be introduced has the potential not only to affect interpre-

---

**Table 4. Visual Acuity at 1 m or Closer Test Distances**

<table>
<thead>
<tr>
<th>Test Distance, m</th>
<th>Average Visual Acuity, Visual Decimal</th>
<th>Equivalent LogMAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.63</td>
<td>-0.212</td>
</tr>
<tr>
<td>0.50</td>
<td>1.39</td>
<td>-0.143</td>
</tr>
<tr>
<td>0.40</td>
<td>1.37</td>
<td>-0.137</td>
</tr>
<tr>
<td>0.33</td>
<td>1.28</td>
<td>-0.107</td>
</tr>
<tr>
<td>0.25</td>
<td>1.05</td>
<td>-0.021</td>
</tr>
<tr>
<td>0.20</td>
<td>0.95</td>
<td>0.022</td>
</tr>
</tbody>
</table>

*Test distance and average visual acuity data from Giese.14 LogMAR indicates logarithm of the minimum angle of resolution.*

---

**Figure 7.** Visual acuity means and SDs obtained on the multiple-choice checkerboard test (Figure 2 of Giese**).
tation of low VA scores but also to bias the estimation of treatment effect, especially when the outcome is the change in VA scores. The general impact is that the loss of VA may be exaggerated. Whether the biased estimates of change in VA will change the conclusion of a trial is a difficult question to answer. The answer will depend on not only the size of discrepancy but also the testing method, scoring algorithm, magnitude and direction of changes observed, and the between-treatments effect size of interest. Thus, these effects should be evaluated within each study, either theoretically by modeling possible discrepancies and outcomes given the testing and scoring protocol and the study design or empirically by collection and analysis of sufficient data.

In summary, theoretical algorithms for calculating VA scores at different test distances may be satisfactory for monitoring of patients for clinical care purposes. However, investigators charged with designing and analyzing VA data for changes over time should be aware of the discrepancies we observed in 4 different multicenter studies that complied with published recommendations for measuring and reporting VA in clinical research. Whenever test distances of less than 2 m are necessary, whether in a randomized trial or a study with another design, provisions should be made to evaluate the consistency of scores obtained at different test distances. Researchers and readers of research publications should be aware of the potential for bias that may lead to misstatement of VA or exaggerated magnitude of changes in VA over time. Our findings support the need for more investigation of these issues as well as continuing improvement in methods to measure VA clinically.

Submitted for publication March 5, 2002; final revision received May 31, 2002; accepted July 10, 2002.

This research was supported by funding award U10 EY11557 to the Submacular Surgery Trials from the National Eye Institute of the National Institutes of Health, Bethesda, Md, and the Michael B. Panitch Stop AMD [Age-Related Macular Degeneration] Fund, Baltimore, Md.

The Submacular Surgery Trials Pilot Study was supported by awards R21 EY10823 and U10 EY11547 from the National Eye Institute and by the Submacular Surgery Trials Research Fund (donors to this fund are listed in the Submacular Surgery Trials Pilot Study Report Number 1). The Macular Photocoagulation Study ancillary study was supported by the National Eye Institute. The Laser to Drusen Trial pilot study was supported by IRIS Medical, Mountain View, Calif, and Coherent Medical Group, Santa Clara, Calif.

The Treatment of Age-Related Macular Degeneration With Photodynamic Therapy study was supported by QLT Inc, Novartis Ophthalmics AG, and Novartis Ophthalmics.

We are grateful to the investigators and sponsors of the 4 studies from which we obtained data.

Corresponding author: Li Ming Dong, PhD, Wilmer Clinical Trials and Biometry, 550 N Broadway, Ninth Floor, Baltimore, MD 21205-2010.

Reprints: Barbara S. Hawkins, PhD, Wilmer Clinical Trials and Biometry, 550 N Broadway, Ninth Floor, Baltimore, MD 21205-2010.

REFERENCES

8. Maguire MG. Correlation between visual acuity measurements taken at five and ten feet: Biometry and Epidemiology Program, National Eye Institute, National Institutes of Health. Paper presented at: Second NEI Symposium on Eye Disease Epidemiology; June 6, 1985, Bethesda, Md.