BJO Online First, published on November 9, 2007 as 10.1136/bjo.2007.121715

Validation of a Computerised logMAR Visual Acuity Measurement System (COMPlog): Comparison with ETDRS and the Electronic ETDRS testing algorithm in adults and amblyopic children.

Running title: Validating COMPlog against ETDRS and E-ETDRS in adults and amblyopic children.

D Alistair H Laidlaw[1,2], Vijay Tailor[2], Nilpa Shah[1], Silva Atamian[1], Cassie Harcourt[2]

1 Vitreo Retinal Unit, St Thomas' Hospital, London, UK
2 EEMU Maidstone General Hospital, Maidstone UK

Address for correspondence
Mr DAH Laidlaw
Consultant Vitreo Retinal Surgeon
St Thomas' Hospital
Lambeth Palace Rd
London SE1 7EH, UK
Alistair.Laidlaw@gstt,nhs.uk

Competing Interest: None declared

Abstract

Background/Aim: The COMPlog clinical visual acuity measuring system is being developed for both routine and research use. This study aimed to validate its performance in amblyopic children and both normal and diseased adults against the gold standard ETDRS chart and the E-ETDRS computerised acuity measurement algorithm.
Method: Timed test and retest fully interpolated 5 letters per line logMAR visual acuity measurements were taken for 70 adults and 59 amblyopic children using the ETDRS chart and the COMPlog visual acuity measurement system. 39 of the adults also underwent computerised acuity testing using the E-ETDRS testing algorithm. The tested adults included normals as well

as subjects with a range of ocular diseases. The methods of Bland and Altman were employed with test retest variability (TRV) expressed as 95% confidence limits for agreement.
Results: No significant bias was observed between the gold standard ETDRS acuity measurements and those taken with either COMPlog or E-ETDRS. TRVs of +/- 0.12 logMAR and +/- 0.10 logMAR were respectively found for COMPlog measurements in the amblyopic children and adult groups compared to +/- 0.12 logMAR for the ETDRS chart in both groups. The TRV of the E-ETDRS system was slightly greater at +/- 0.16 logMAR. Median testing times for COMPlog and ETDRS were 95 and 85 seconds and 66 and 56 seconds respectively in the paediatric and adult groups and 120 seconds for the E-ETDRS measurements on adults.
Discussion: COMPlog measurements agree well with and are similarly reliable to the gold standard ETDRS chart with comparable test times. E-ETDRS algorithm measurements took approximately twice as long.

Introduction

The COMPlog computerised clinical visual acuity measuring system is being developed as an alternative to hard copy acuity charts for routine and research use. Any new acuity test must be shown to perform well against the current gold standard ETDRS chart[1,2]. Contour interactions present in amblyopia may influence clinical acuity measurement; any letter test designed for clinical use should therefore be shown to be applicable to literate children as well as adults[3]. The only other computerised acuity measurement system of which we are aware is the E-ETDRS algorithm running on the Electronic Visual Acuity tester[4,5,6]. The aim of this method comparison study was to validate the performance of the COMPlog visual acuity measurement system using amblyopic children, normal adults, and adults with a range of acuity impairing ophthalmic diseases as subjects. Performance was determined in terms of 1) agreement (bias) with gold standard ETDRS acuity chart measurements, 2) by comparison of the test retest variability (TRV) of COMPlog, ETDRS and E-ETDRS measurements and 3) test time for each measurement system.

Patients and Methods

59 children undergoing treatment for amblyopia were invited to participate, as were 70 adults with either normal or diseased eyes. Ethical approval for this study was granted and informed consent to participate was obtained. Each subject underwent timed single letter scoring test and retest measurements of the visual acuity of their poorer acuity eye using ETDRS charts 1 and 2 and the COMPlog system (i.e. 4 tests in total). The last 39 of the 70 adult subjects recruited also underwent computerised acuity measurements using the E-ETDRS algorithm. All tests were performed in random sequence in order to control for fatigue and learning effects. All measurements were conducted by one of 3 trained examiners (VT, SA, NS) under consistent lighting conditions. Subjects wore their habitual spectacle correction with their fellow eye occluded.[7]
Responses and computerised test scores were recorded on specifically designed proformata.

COMPlog is a computerised visual acuity measurement system consisting of a laptop PC capable of running Micrsoft Windows XP®, a 21 inch 1600X1200 resolution LCD flat panel secondary monitor and a software programme running within the Microsoft dotnet ® framework. The examiner controls the test through a series of sequential control screens presented on the laptop monitor (figure 1): these enable collection of demographic data, control of various aspects of the testing algorithm (see below), response recording, and presentation of results in 1 of 2 logMAR formats (decimal logMAR and number of ETDRS letters) and 1 of 3 Snellen equivalent format scores (Decimal Snellen, UK Snellen and US Snellen).

The COMPlog testing algorithm consists of 2 phases: ‘range finding’ and ‘thresholding’; both of which require forced choice responses from the subject and input of the response to each letter as correct or incorrect by a technician. Range finding aims to roughly identify threshold and consists of presentation of sequentially smaller single crowded Sloan letter steps starting from 0.8 logMAR, with a step size of 0.2 logMAR. If the 0.8 logMAR letter cannot be correctly identified then sequentially larger letters are presented. The thresholding phase commences 0.4 logMAR larger than the range finding result, unless this exceeds 1.2 logMAR in which case it commences at 1.6 logMAR. The thresholding phase consists of presentation of sequential lines of randomly chosen Sloan letters surrounded by a crowding box. No letters are repeated on any line. The line size increment employed in the thesholding phase is set at 0.1 log units. The other aspects of the test are however user controllable: these include letter spacing, the presence or absence of crowding bars, the number of letters per line, number of letters per line incorrectly identified to terminate the test and number of tests upon which the score is based. In this study 5 letters per line spaced half a letter width apart and surrounded at the same separation by a crowding box of one stroke width thickness (figure 1) were employed. The termination criterion was set at all 5 letters on one line. In the event of letters being incorrectly identified on the first thresholding line sequentially larger lines are presented until an entire line is correctly read, with the programme then descending to threshold, but only presenting lines of each size of letters once. If all 5 letters cannot be fitted simultaneously on to the secondary monitor the line is broken up into fractions with as many letters as possible of each size being presented. Letters larger than 1.6 logMAR are not presented, if a patient is unable to correctly identify any letters of 1.6 logMAR size the programme invites scoring on a count fingers, hand movements, perception of light and no perception of light scale. In this way

with a 21 inch monitor and a single viewing distance of 3 meters an acuity range of up to 1.68 logMAR (1 ETDRS letter, 1.5/71 UK Snellen or 20/957 US Snellen) may be measured without moving the patient.

Once the termination criterion has been met the test automatically terminates with calculation and presentation of a fully interpolated logMAR acuity score and its Snellen equivalent in the selected formats, along with the total test time.

The COMPlog programme was also adapted to run the E-ETDRS test algorithm [4,5,6]. This test measures a 5 letters per line fully interpolated logMAR acuity and also employs 0.1 logMAR line size increments but does so by presenting one letter at a time in a surrounding crowding box spaced half a letter width from the test letter.

ETDRS charts 1 and 2 were used and displayed in the standard Lighthouse ® Low Vision Products light box[2]. The ETDRS charts were read from a distance of 4m unless the subject misnamed any letters on the top line. In this event, the subject was moved to 1m and the remainder of the measurements taken at this distance with 0.6 logMAR being added to the score.

The end point for all tests employed in this study was defined as an entire line of letters being misread. A fully interpolated logMAR acuity score was calculated for each test result as previously described[8,9,10].

Statistical analysis

The methods of Bland and Altman[11] were used to quantify 1) bias (mean and 95% confidence interval of the mean) between ETDRS logMAR and both COMPlog and E-ETDRS algorithm scores 2) test retest variability (TRV) expressed as 95% confidence limits for agreement (mean plus and minus 2 standard deviations) for the paired ETDRS, COMPlog and E-ETDRS algorithm scores. Time scores were compared using paired t tests.

59 amblyopic children (37 male) aged between 5 and 10 years (mean age 7.3 yrs) participated in the study. 22 subjects had strabismic amblyopia, 15 anisometropic amblyopia, 17 had a combination of strabismic and anisometropic amblyopia and 5 had various types of meridional amblyopia. The median visual acuity of the tested eyes was 0.56 logMAR range 0.00 to 0.96 logMAR (UK Snellen equivalents, 6/22, 6/6 to 6/55).

Seventy adult patients (39 Male) aged 23 to 81 years (median age 65 years) were also recruited. Their tested eyes had had a median acuity of 0.32 logMAR, range –0.12 to 1.30 logMAR (UK Snellen equivalents, 6/13, 6/4.6 to 3/60). 8 were normal, 6 had cataract or corneal disease, 27 surgical retinal disease and 3 glaucoma or other optic neuropathy. Mixed pathology was present in 26.

The COMPlog secondary monitor had a screen luminance of 41 cd/m^2 and the contrast of letters was measured to be 94%.

Histograms of the distribution of the test-retest and between test acuity variability data suggested that the data sets conformed reasonably to a normal distribution. Scatter plots of the observed test retest variation plotted against the average of the test and retest measurements suggested that there were no systematic association between TRV and the underlying acuities (figure 2). The parametric methods of Bland and Altman were therefore employed. The results of the studies are presented in tables 1, 2 and 3.

In the paediatric study the median (range) of test times for ETDRS logMAR and COMPlog measurements were respectively 85 (40-251) seconds and 92(38-233) seconds ($p=0.52$). The corresponding median test times in the adult study for ETDRS and COMPlog acuity tests were 56 (range 17-148) and 66 (range 27-275) seconds ($p<0.001$) and for E-ETDRS 120 (range 60-300) seconds ($p<0.001$).

Table 1 *COMPlog and ETDRS inter test and test retest agreement: Results from 59 amblyopic children.*

Single letter Scoring	Mean difference (Standard error) logMAR	95% CI mean difference, logMAR	S.D. differences logMAR	TRV (95% confidence limits for agreement) logMAR	Min/max difference logMAR
ETDRS-ETDRS	0.01 (0.01)	-0.01, 0.02	0.06	+/- 0.12	-0.14, 0.16
COMPlog-COMPlog	0.02 (0.01)	0.00, 0.04	0.06	+/- 0.12	-0.08, 0.14
ETDRS-COMPlog	0.01 (0.01)	-0.01, 0.03	-	-	-

Table 2 *COMPlog agreement with ETDRS measurements: Results from 70 adults*

	Mean difference (standard error)	95% CI mean difference	S.D. differences	TRV (95% confidence limits for agreement)	Min/Max difference
ETDRS-ETDRS	0.00 (0.01)	-0.02, 0.02	0.06	+/-0.12	-0.10, 0.16
COMPlog-COMPlog	0.01 (0.01)	0.00, 0.02	0.05	+/-0.10	-0.12, 0.16
ETDRS-COMPlog	-0.02 (0.01)	-0.04, 0.00	-	-	-

Table 3 *E-ETDRS algorithm and COMPlog agreement with ETDRS measurements: Results from 39 adults. These subjects are a consecutive sub set of those reported in table 2.*

	Mean difference (standard error)	95% CI mean difference	S.D. differences	TRV (95% confidence limits for agreement)	Min/Max difference
ETDRS-ETDRS	0.00 (0.01)	-0.02, 0.02	0.07	+/-0.14	-0.10, 0.16
COMPlog-COMPlog	0.02 (0.01)	0, 0.04	0.05	+/-0.10	-0.12, 0.12
E ETDRS-E ETDRS	0.02 (0.01)	-0.01, 0.05	0.08	+/-0.16	-0.22, 0.18
ETDRS-COMPlog	-0.02 (0.01)	-0.05, 0.01	-	-	-
ETDRS- E ETDRS	-0.03 (0.02)	-0.06, 0.00	-	-	-

Discussion

The aim of these studies was to validate the performance of a computerised visual acuity measurement system which we have named COMPlog. There are many potential advantages of such a computerised visual acuity measurement system: These include:

Controllable and standardised measurement rigor, thereby minimizing training and technician based errors,
Automated score calculation and presentation in any of the routinely encountered acuity formats (decimal, UK and US Snellen; decimal logMAR and number of ETDRS letters read),
Random letter generation avoiding memorization effects,
Measurement from one testing distance of 6/3 to 1/60 acuity.

COMPlog employs the Sloan letter set and follows the standard ETDRS chart design format of one stroke width components and a 5X5 stroke width design grid [1,2]. The between letter spacing and the separation of the letters to the surrounding crowding box however is based on half a letter width rather than the full letter width separation on an ETDRS chart. We have previously shown in both adults and amblyopic children that measurements made on hard copy logMAR chart using half letter width letter, line and crowding bar separation agree well and are comparably reliable to the gold standard ETDRS chart[3,7]. A notable difference is however that these charts presented multiple lines of letters rather than a single line in a crowding box.

With the use of a 1600X1200 pixel 21 inch monitor and a viewing distance of 3 meters it is possible to present a single letter in a crowding box of 1.6 logMAR size, this is equivalent to the size of the top line of an ETDRS chart viewed from 1m. This means that the COMPlog system is potentially able measure the entire range of clinically encountered acuities from one test distance, as opposed to the 2 test distances usually employed with hard copy charts. It also means that an acuity of Count Fingers (CF) is known to be worse than one ETDRS letter or 1.68 logMAR (6/288 Snellen); in routine clinical practice CF can mean anything from worse than 6/60 to worse than 1/60 depending on the rigor with which the test was performed.

The data presented here are the results of the first of a series of experiments which aim to rigorously determine and document the performance of the COMPlog visual acuity measurement system. The ETDRS chart has been used for comparison purposes as it is the current gold standard and the E-ETDRS algorithm employed as the only other available computerised acuity measurement system of which we are aware. Limited performance data from one other computerised acuity measurement system have been published. This system is no longer available and was limited to measuring acuity of 0.6 logMAR (Snellen 6/24 or better from its designed test distance [12] .

The results of these studies have shown that COMPlog measurements agree well (i.e. show no evidence of bias) when compared to the ETDRS chart when measuring the acuity of the defective eye of amblyopic children and in adults with widely ranging acuity and ocular

pathology. This suggests that the method of presenting a single line of letters spaced half a letter width apart and surrounded by a crowding box does not result in either adverse or beneficial contour interaction. COMPlog also differs from the ETDRS chart in terms of letter selection, screen luminance and letter contrast: COMPlog letters are randomly selected with the exception that no repeats occur on any line; whereas the combinations of letters on each line of an ETDRS chart was deliberately chosen in an attempt to ensure equal legibility of each line [2]. COMPlog screen luminance and letter contrast are also both lower than that on the ETDRS chart (35cd/m2 vs 111cd/m2 and 92% vs 100%). The observed lack of bias between COMPlog and ETDRS measurements suggest that these differences do not exert a clinically important effect on acuity measurements.

The TRV of ETDRS measurements was found to be +/- 0.12 logMAR in both our paediatric amblyopia and adult groups. (The 39 patients reported in table 3 are a consecutive sub set of the 70 described in table 2). These results are comparable to other studies on the repeatability of ETDRS measurements in which TRV values ranging between +/-0.07 and +/-0.20 logMAR have been reported[3,4,7,13-18.] In this respect our subjects appear to be similarly reliable to those reported elsewhere. The TRV of the 5 letters per line COMPlog measurements in the 59 paediatric amblyopes which we studied was found to be +/-0.12 logMAR and +/-0.10 logMAR in the 70 adults. This suggests that the TRV of COMPlog measurements might reasonably be considered to be at least comparable to those taken using ETDRS. The TRV of E ETDRS measurements was found to be +/- 0.16 compared to +/-0.14 for ETDRS and +/-0.10 for COMPlog in the 39 adults on whom all 3 test and retest measurements were taken. It was not our aim to test the significance of these differences.

Beck et al reported no bias when comparing EVA E-ETDRS algorithm measurements to hard copy ETDRS chart measurements in 265 adults[4]. 2 further validation studies have been carried out using the EVA and E-ETDRS algorithm on children; 100/311 of whom had had amblyopia.[5,6]. Neither study documents agreement between hard copy ETDRS and E-ETDRS measurements so potentially important bias could conceivably have been present. The TRV of EVA E-ETDRS measurements in 265 adults reported was +/-0.14 logMAR[4] and between +/- 0.12 and +/- 0.14 logMAR in the paediatric subjects[5]. These TRV data and the ones we present above are broadly comparable to published data for the TRV of the ETDRS chart in adults and children and provide further validation of computerised visual acuity testing devices.

A 5 letters per line COMPlog measurement was found to take slightly longer than an ETDRS measurement (median test times 92 compared to 85 seconds and 66 compared to 56 seconds respectively) in the paediatric and adult populations. All test times include the ETDRS score calculation time, it is however our experience that the calculation of an ETDRS score may be prone to technician error. The adult time difference reaches statistical significance ($p<0.001$) but the paediatric does not ($p=0.52$). The clinical significance of the 10 second time penalty of using COMPlog as opposed to ETDRS is up to the user to decide. E-ETDRS measurements were found to take nearly twice as long at a median of 120 seconds in the tested adult population. This is because the E-ETDRS algorithm presents a single crowded letter at a time ($p<0.001$). We believe that this might be considered a clinically significant time penalty. We are not aware of any other data on test times with the E-ETDRS algorithm.

COMPlog allows user control of testing rigor in terms of numbers of letters per line and termination criteria. The influence of these factors on testing time and TRV will be explored in future studies. In general it might reasonably expected that fewer letters per line and less stringent termination criteria would result in faster but less repeatable measurements[3,7].

In summary these studies have shown that COMPlog visual acuity measurements agree well, show similar test retest variability and take a similar length of time to the gold standard ETDRS chart. E-ETDRS measurements took approximately twice as long.

References

1 Bailey IL, Lovie JE. New design principles for visual acuity letter charts. *Am J Optom Physiol Opt* 1976; 53: 740–5.

2 Ferris FL 3rd, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *Am J Ophthalmol* 1982; 94: 91–6.

3 Laidlaw DA, Abbott A, Rosser DA. Development of a clinically feasible logMAR alternative to the Snellen chart: performance of the "compact reduced logMAR" visual acuity chart in amblyopic children. *Br J Ophthalmol* 2003; 87:1232-4.

4 Beck RW, Moke PS, Turpin AH, Ferris FL 3rd, SanGiovanni JP, Johnson CA, Birch EE, Chandler DL, Cox TA, Blair RC, Kraker RT. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol* 2003;135:194-205.

5 Cotter SA, Chu RH, Chandler DL, Beck RW, Holmes JM, Rice ML, Hertle RW, Birch EE, Moke PS. Reliability of the electronic early treatment diabetic retinopathy study testing protocol in children 7 to <13 Years Old. *Am J Ophthalmol* 2003;136:655-61.

6 Moke PS, Turpin AH, Beck RW, Holmes JM, Repka MX, Birch EE, Hertle RW, Kraker RT, Miller JM and Johnson CA. Computerized method of visual acuity testing: adaptation of the amblyopia treatment study visual acuity testing protocol. *Am J Ophthalmol* 2001;132:903-9.

7 Rosser DA, Laidlaw DA, Murdoch IE. The development of a 'reduced logMAR' visual acuity chart for use in routine clinical practice. *Br J Ophthalmol* 2001;85:432-6.

8 Vanden Bosch ME, Wall M. Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye* 1997;11:411–7.

9 Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci* 1991;32:422–32.

10 Arditi A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol Vis Sci* 1993;34:120–9.

11 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986:307-10.

12 Rosser DA, Murdoch IE, Fitzke FW, Laidlaw DA. Improving on ETDRS acuities: design and results for a computerised thresholding device. Eye. 2003 Aug;17:701-6.

13 Siderov J, Tiu AL. Variability of measurements of visual acuity in a large eye clinic. *Acta Ophthalmol Scand* 1999;77: 673–6.

14 Lovie-Kitchin JE. Validity and reliability of visual acuity measurements. *Ophthalmic Physiol Opt* 1988; 8:363–70.

15 ves BC, Wood JM, Hill AR. Vistech VCTS 6500 charts—within- and between-session reliability. *Optom Vis Sci* 1991; 68:728–37.

16 iott DB, Sheridan M. The use of accurate visual acuity measurements in clinical anti-cataract formulation trials. *Ophthalmic Physiol Opt* 1988; 8: 397–401.

17 iti A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol Vis Sci* 1993; 34: 120–9.

18 Reeves BC, Wood JM, Hill AR. Reliability of high- and low-contrast letter charts. *Ophthalmic Physiol Opt* 1993; 13:17–26.

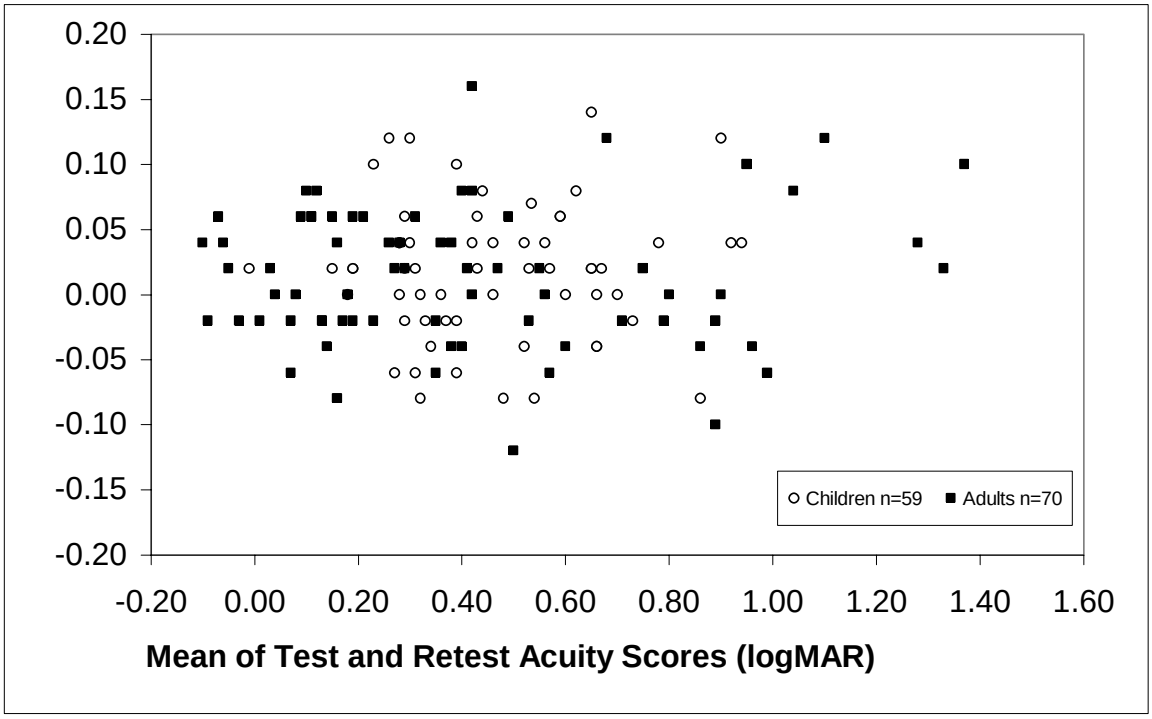

Figure 2
Scatterplot of the difference between COMPlog test and retest logMAR acuity measurements measurements in 59 children (circles) and 70 adults (squares) plotted against the mean of the test and retest measurements.

Validation of a Computerised logMAR Visual Acuity Measurement System (COMPlog): Comparison with ETDRS and the Electronic ETDRS testing algorithm in adults and amblyopic children.

Alistair Laidlaw, Nilpa Shah, Silva Atamian, et al.

Br J Ophthalmol published online November 9, 2007
doi: 10.1136/bjo.2007.121715