Repeatability of visual function measures in age-related macular degeneration

**Tariq Aslam · Sajjad Mahmood · Konstantinos Balaskas ·
Niall Patton · Rajeev G. Tanawade · Shi Zhuan Tan ·
Stephen A. Roberts · Jeremy Parkes · Paul N. Bishop**

Abstract

Purpose To assess repeatability of visual function measures in patients with early, intermediate or late age-related macular degeneration (AMD) without active neovascular disease in the study eye, but active neovascular AMD in the fellow eye.

Methods One hundred subjects from an ongoing trial were screened for this study in which their LogMAR acuity, contrast sensitivity and reading performance were assessed using standardised protocols by trained optometrists. The same measures were repeated one month later and repeatability of the visual functions assessed.

Results Data from 83 subjects satisfied inclusion criteria for analysis. Coefficient of repeatability was 14.9 letters for LogMAR visual acuity , 7.2 letters for Pelli Robson contrast sensitivity, 0.72 for LogMAR reading acuity, 110.4 words/min for reading speed and 0.67 for LogMAR critical print size. Intraclass correlation coefficients allowed comparison between measures and were found to be 0.96 for LogMAR visual acuity, 0.93 for contrast sensitivity, 0.75 for LogMAR reading acuity, 0.79 for reading speed and 0.74 for LogMAR critical print size. Coefficients of variation were 9.4 %, 10.7 %, 48.4 %, 28.4 % and 31.8 % respectively.

Conclusions We found coefficients of repeatability that concurred with previous studies demonstrating variability of visual functions in patients with AMD. In addition, we found intraclass correlation coefficients to be better with visual acuity and contrast sensitivity than with measures of reading performance.

Keywords Age-related macular degeneration · contrast sensitivity · distance visual acuity · reading performance · repeatability

T. Aslam · S. Mahmood · K. Balaskas · N. Patton ·
R. G. Tanawade · S. Z. Tan · J. Parkes · P. N. Bishop
Manchester Royal Eye Hospital, Manchester Academic
Health Science Centre, Manchester, UK

P. N. Bishop
Centre for Advanced Discovery and Experimental Therapeutics
(CADET), Central Manchester University Hospitals NHS
Foundation Trust, Manchester Academic Health Science Centre,
Manchester, UK

T. Aslam · P. N. Bishop
Centre for Ophthalmology and Vision Research, Institute of Human
Development, University of Manchester, Oxford Road,
Manchester, UK

S. A. Roberts
Centre for Biostatistics, Institute of Population Health,
University of Manchester, Oxford Road, Manchester, UK

K. Balaskas (✉)
Moorfields Eye Hospital NHS Foundation Trust,
City Road, London, UK
e-mail: konstantinos.balaskas@gmail.com

Introduction

Management of patients with active neovascular age-related macular degeneration (AMD) involves regular patient visits with administration of intravitreal anti-vascular endothelial growth factors (anti-VEGF) as required, largely guided in clinical practice by optical coherence tomography (OCT) scans. An important additional aspect of management is monitoring of fellow eyes that are at risk of neovascular disease. In addition to formal vision testing at each hospital visit, patients are advised to return urgently should they experience a change in vision in these fellow eyes between hospital visits. It may sometimes be practical for such perceived changes to be confirmed objectively by home vision testing or testing by an optometrist or other medical practitioner before repeat access to formal ophthalmological examination and OCT is arranged. Distance visual acuity continues to be the standard

visual measure used to guide decisions on therapeutic management of AMD, despite limitations for the assessment of functional deterioration. [1–5] Other psychophysical tests of vision, such as contrast sensitivity [6] and reading performance [7, 8] have additionally been shown to be useful tools for monitoring patients with early AMD and have been used in clinical trials [9, 10].

Knowledge of the expected reliability of testing different visual functions in patients with stable AMD would be invaluable in deciding whether measured outcomes suggest the eye pathology is no longer in the category of stable AMD and further investigations are prudent. Reliability of one test compared to another could guide carers as to which particular visual functions would be most appropriate to measure.

In order to determinethe reliability of vision tests in patients with eyes that have stable AMD, recent papers have used the untreated fellow eye of patients receiving active therapy for neovascular AMD [11–13] with separate papers providing evidence of reliability of visual acuity and contrast sensitivity in patients with AMD. However, direct comparisons have not been made of the reliability between the different measures of visual function, nor over the periods of time commonly used in a clinical practice setting.

In this study we assessed the reliability of various visual function measures in eyes of patients who had no active neovascular AMD in their study eyes, but were being managed for active neovascular AMD in the fellow eye. We aimed to add to existing knowledge of reliability of key assessments of visual acuity and contrast sensitivity in this setting and in addition compare these visual function measures with each other and with reading measures.

We assessed and compared the repeatability of tests for visual acuity and contrast sensitivity, and for three parameters of reading; reading acuity, reading speed and critical print size. Previous published studies assessed patients up to 12 weeks apart and showed no apparent effect of deterioration bias on reliability measures; [11–13] in this study, we were able to assess visual functions at 4 weeks apart with clinical and imaging confirmation of no neovascular disease. We report our findings in accordance with guidelines for reporting reliability and agreement studies (GRRAS) [14].

Methods

Visual function data was derived from patients in a randomised prospective trial, the Greater Manchester Avastin for Choroidal Neovascularisation Trial (GMAN). The trial compares the efficacy of two different treatment regimes using bevacizumab (Avastin; Genentech, Inc., South Franciso, CA) for neovascular AMD. The GMAN trial is registered under ISRCTN (ISRCTN34221234). All patients had consented to vision tests and the protocols were in accordance with the declaration of Helsinki and local research ethics committee approval. AMD was classified in the study eye according to criteria defined by the International Classification study group [15] including early, intermediate and advanced (atrophic or neovascular) AMD. Vision measures from fellow eyes of patients being treated for neovascular AMD in the GMAN trial were considered for this reliability study, taken one month apart. Inclusion criteria were that patients should have in their study, for the purposes of the present work, (untreated) eye early moderate or advanced AMD, including disciform scarring, without evidence of active CNV and not having previously been treated with anti-VEGF injections. The first 100 sequential patients who had been recruited into the GMAN trial were assessed and those who fulfilled entry criteria and had full datasets at the time of statistical analyses were included. Recruitment of these patients took place between 03/02/2008 and 30/04/2011. All patients underwent clinical assessment and OCT imaging to determine disease activity in either eye at both visits, while a subgroup also underwent FFA as per the GMAN trial protocol.

BCVA was measured by registered optometrists experienced in trial work following a standard protocol, using EDTRS LogMAR charts R, 1 and 2 at 2 m (Precision Vision, USA). Chart R was used for refraction; charts 1 and 2 were used for right and left eyes respectively. The charts were presented on an internally illuminated light box (Precision Vision). The tubes were "burnt in" for 96 hours before the start of the study and were replaced annually with similarly "burnt in" tubes to ensure consistent illumination. The room background room illumination was reduced to below 15 foot-candles by the use of blinds; the same room was used on each visit to ensure consistency of ambient conditions. Background illumination was measured with the use of a photometer on first study visit for each patient.

Optometrists were masked to previous measurements of acuity. Patients were refracted at 2 m on each visit using a standardized protocol. The starting point was the previous refraction, using distance spectacle prescription or retinoscopy if no spectacles were available.

The visual acuity was measured by asking the patient to read down the side of the test chart, i.e., read the initial letter on each line, until they started to miscall letters, this gave an approximate threshold. The examiner then directed the patient to the initial letter on the line three lines above the approximate threshold and asked the patient to read each line, progressing down the chart, letter by letter until they made 4 miscalls on a line. If the first three lines the patient was asked to read were not read correctly the patient was directed to the start of the chart and asked to read all of the letters in turn. This approach was adopted to reduce patient fatigue using a protocol similar to that suggested by Carkeet [16]. Patients were always encouraged to read as many letters as possible and encouraged to guess until at least four

mistakes were made in reading the lowest line seen. If a patient miscalled a letter but corrected themselves before moving onto the next letter the correction was scored, however if they went back after moving onto the next letter and tried to correct the error the correction was not scored.

The visual acuity was converted to a 1-m ETDRS letter score by adding 15 to the number of letters read at the 2-m working distance. If less than four letters were read at 2 m, the test was repeated at a 1-m working distance adding +0.50 DS to the prescription. The number of letters correctly read at 1 m (up to 15, as more than 15 would result in double-counting letters) was added to the number of letters read at 2 m to give the visual acuity number of letters.

Two measurements of VA were obtained for each eye; the sequence of presentation of charts being chart 1 RE, then chart 2 LE back to chart 1, and then chart 2 again. The results were averaged to improve precision. The examination was done twice by the same person and the average used as the final measure for each visit.

Contrast sensitivity (CS) was determined with the Pelli Robson charts (Clement Clarke International, Essex, UK) at 1 m, using the previously determined 2-m refraction. The illumination was 500 Lux. The 2-m refraction obtained for the visual acuity measure was used for the contrast sensitivity. Different charts were used for each eye. The contrast sensitivity score was the total number of letters correctly identified. The patient was encouraged to guess until no further letters could be identified. The scores were letter by letter as advocated by Elliott, Sanderson and Conkey [17].

Reading performance was measured with MN reading charts (Lighthouse, New York, USA, http://gandalf.psych. umn.edu/groups/gellab/MNREAD/) at 40 cm using a +2.00 addition to the 2-m refraction, which gave the optimum focus, as these charts are calibrated for a 40-cm working distance. Illumination was 450 lux, and the cards were evenly illuminated and displayed on a reading stand. These charts are continuous-text reading acuity charts; they consist of 19 sentences of 60 characters. Each sentence is printed as three lines with the sentences using progressively smaller typeface, (ranging from +1.3 to −0.5 logMAR). The task was for the patient to read each sentence aloud as quickly and as accurately as possible; this was done monocularly. Each sentence was uncovered sequentially so the patient could not see any of the text before they started reading. The time taken for the patient to read each sentence was measured along with recording the number of words read incorrectly. The patients were asked to attempt every sentence until they were unable to read any of the text. This enabled calculation and graphical determination of critical print size, reading acuity and maximum reading speed.

Reading acuity (the smallest print size read correctly) was defined precisely as 1.4-(sentences read x0.1) + (number of words read incorrectlyx0.01). Critical print size was the size of print at which reading speed first starts to decline from the maximum value, which was measured as number of words read per minute. These last two measurements were determined by graphical plotting the time taken to read each sentence against sentence print size. A "subjective best fit" curve was applied to the data and the maximum reading speed was read off the graph and critical print size taken as the first point at which the curve starts to descend from the plateau.

All the vision measurements were retaken under identical conditions 4 weeks apart, which is the typical inter-session interval used in the majority of AMD services for patient re-assessment. Any patients who had evidence of disease progression or active neovascular AMD in their study (untreated) eye by clinical assessment, OCT imaging or fluorescein angiography at either visit were excluded from the analysis.

Repeatability of each visual function was demonstrated using Bland–Altman plots [18] for all visual function types (Medcalc v 11.4.4.0) as well as calculation of British Standards Institution repeatability coefficient (British Standards Institution BS5497) [19](statsdirect v2.7.8). Intra-class correlation coefficients (one-way random effects) were also calculated for each visual function, providing coefficients that would allow for direct comparison of the different visual functions. Additionally, we present within-subject standard deviations (Medcalc v 11.4.4.0) which, to allow for scale differences, we normalised by the mean values to give coefficients of variation. The same analysis was repeated after the exclusion of all patients with disciform scarring to determine whether this subgroup of patients had a deleterious impact on the overall reliability statistics

Results

A total of 83 patients met all the entry criteria and were included in the analyses. The average age was 80 years old (range 50–95 yrs). There were 33 males and 50 females. There were 58 patients who had features ≤ early AMD, defined as presence of one or more drusen ≤ 64 microns in size and mild retinal pigment epithelial (RPE) disturbance (including pigment clumping and/or dropout). Six patients had features of moderate AMD defined as presence of at least one drusen of 64–125 microns in size. Four patients had geographic atrophy, and 15 had disciform scarring.

Bland–Altman plots for each of the visual functions measured are shown in Fig. 1. Bland–Altman plots did not reveal any systematic trends either with or without patients with disciform scarring. Scrutiny of Bland–Altman charts did not demonstrate any untoward relationship between agreement and magnitude of any measures, with random scatter above and below zero and no significant bias.

Fig. 1 Bland–Altman Plots for visual function measures. a. LogMAR Visual Acuity Reliability (Number of Letters). b. Peli Robson contrast sensitivity reliability (number of letters). c. reading acuity reliability. d. reading speed reliability (words/min). e. critical print size reliability
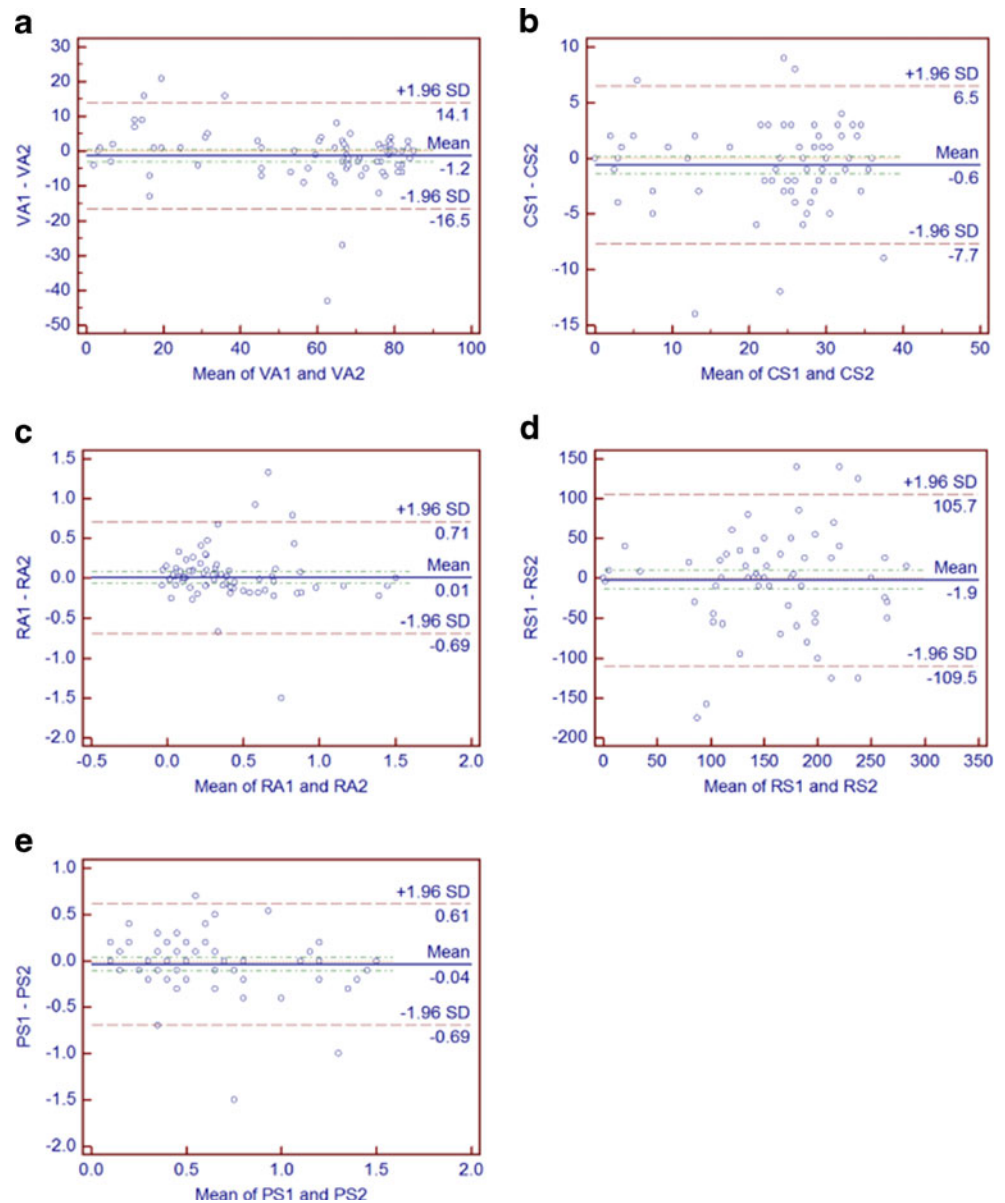

Table 1 offers a summary of BSI repeatability coefficients, intraclass correlation coefficients and coefficients of variation for all the visual performance measures evaluated in this study. Table 2 summarizes the same measures of repeatability after the exclusion of the 15 patients with disciform scarring. Intraclass correlation coefficients were 0.96 (95 % CI 0.93-0.97) for LogMAR visual acuity, 0.93 (95 % CI 0.89-0.95) for Peli Robson contrast sensitivity, 0.75 (95 % CI 0.63–0.83) for reading acuity, 0.79 (95 % CI 0.69–0.86) for reading speed and 0.74 (95 % CI 0.63-0.83) for critical print size. After the exclusion of patients with disciform scarring, corresponding coefficients became 0.91 (0.87–0.95), 0.84 (0.76–0.90), 0.69 (0.55–0.80), 0.68 (0.53–0.79) and 0.68 (0.53–0.79) respectively. Coefficients of variation were 9.4 % for distance visual acuity, 10.7 % for contrast sensitivity, rising to 48.4 % for reading acuity, 28.3 % for

reading speed and 31.8 % for critical print size. After the exclusion of patients with disciform scarring, these values became 7.6 %, 8.7 %, 56.4 %, 25.1 % and 35.9 %, respectively.

Discussion

We found the coefficient of repeatability for distance LogMAR ETDRS acuity to be 14.9 letters, slightly higher than the 12 letters demonstrated in a paper by Patel et al. [11] using a similar study design to the one presented herein. Protocols for visual acuity testing were similar in the two studies, and this slight discrepancy might be explained by random differences in study population pathology between the two treatment groups. The coefficient of repeatability for

Table 1 Summary of repeatability statistics (all patients)

Visual function measure	BSI repeatability coefficient	Intraclass correlation coefficient (95 % CI*)	Coefficient of variation
LogMAR visual acuity	14.86 letters	0.96 (0.93–0.97)	9.4 %
Peli Robson contrast sensitivity	7.24 letters	0.93 (0.89–0.95)	10.7 %
Reading acuity	0.72	0.75 (0.63–0.83)	48.4 %
Reading speed	110.4 words/min	0.79 (0.69–0.86)	28.4 %
Critical print size	0.67	0.74 (0.63–0.83)	31.8 %

*95 % CI=95 % Confidence Interval

contrast sensitivity was 7.2 letters, a result very similar to a previous report of 7 letters [12]. Intra-class correlation coefficients for these two particular functions in our study were similar (log MAR acuity 0.96, contrast sensitivity 0.93). However, reliability of the reading measures appeared less on Bland–Altman charts compared to distance VA and CS and intraclass correlation coefficients confirmed this relationship (reading acuity 0.75, reading speed, 0.79, critical print size 0.74). The source of increased variability of reading measures may be the patient's ability to use cognitive processes to mask visual disability by guessing correctly on the basis of context rather than visual information [20]. Factors such as eccentric fixation in patients with extensive scarring could have led to poorer reliability scores and so the statistical analyses were repeated after exclusion of the 15 patients with disciform scarring to assess whether repeatability was improved without these patients. Repeatability coefficients showed negligible differences after this exclusion, for all measures other than reading speed which actually worsened, most likely due to a floor effect of reading measurement (0.72 for all eyes vs. 0.59 after exclusion of disciform scarring).

There are two principle limitations in application of these results to interpretation of patients' visual function scores. Firstly, we specifically present data for reliability of visual measures for patients with no active neovascular AMD in the study eye but active neovascular AMD in the fellow eye which required regular review appointments and treatments with intravitreal anti-VEGF agents. Our conclusions are therefore only valid for similar clinical scenarios; different presentations of AMD may be associated with different intrinsic variability of visual function testing. For example, it has been observed that the impact of macular lesions on VA

in patients with neovascular AMD depends on whether the study eye is the better or the worse-seeing eye; the correlation between structural macular changes and VA is stronger for the worse-seeing eye and it has been suggested that visual loss in one eye may enhance the ability of the fellow eye to reach full functional potential [21]. Various attempts have been made for determining the repeatability of visual function measures under different clinical scenarios, rendering them not directly comparable to the present work. In the study by Kiser et al. [22] reliability of VA and CS in eyes falling within the definition of legal blindness, irrespective of underlying cause, was, not surprisingly, slightly lower than the one reported in the present study. In an older publication, repeatability of VA measurements was reported to be higher, though calculations were based on pooled data from both eyes of patients suffering in their majority, though not exclusively, from AMD [23]. Clinicians should therefore be cautious about extrapolating the data from the present study to alternative clinical scenarios. Secondly, our study assessed and compared reliability of vision measures when each measure was individually optimised to the highest practical standards. They were assessed by trained optometrists in a research setting to standardised research protocols designed to maximise reliability. For example, as well as standardisation of illumination and new refraction before assessments, two readings for visual acuity and for contrast sensitivity were taken and then averaged to produce a final value, improving reliability. Reading tests took longer and were not repeated but involved considerable assessor vigilance in ensuring even illumination, patient explanation and timing accuracy. It is important to note therefore, that assessments done by less qualified personnel, in less rigorous a manner and in busier environments may be less reliable.

Table 2 Summary of repeatability statistics (patients with disciform scarring excluded)

Visual Function measure	BSI repeatability coefficient	Intraclass correlation coefficient (95 % CI*)	Coefficient of variation
LogMAR visual acuity	14.00 letters	0.91 (0.87–0.95)	7.6 %
Peli Robson contrast sensitivity	6.76 letters	0.84 (0.76–0.90)	8.7 %
Reading acuity	0.59	0.69 (0.55–0.80)	56.4 %
Reading speed	113.4 words/min	0.68 (0.53–0.79)	25.1 %
Critical print size	0.63	0.68 (0.53–0.79)	35.9 %

*95 % CI=95 % Confidence Interval

Despite limitations, the data presented here provides a useful complement to existing information on vision test reliability in AMD, particularly by adding crucial new information allowing comparison between measures. Knowledge of reliability of different functional measures is important for assessing, in a clinical setting, whether changes in vision of fellow eyes of patients attending for intravitreal anti-VEGF injections should trigger further attention such as OCT assessment. Similarly, patients with unilateral advanced AMD should be advised to monitor their own vision in the eye that is not being treated for symptoms of new pathology and knowledge of the reliability of the measurements used is crucial for setting thresholds for management action. The data from this study demonstrates that distance visual acuity and contrast sensitivity testing are more reliable than reading measures when testing fellow eyes of patients under treatment for neovascular AMD, and indicates the level of change that could be regarded as clinically significant over the 4-week period between patient reviews as commonly used by AMD services.

Conflict of Interests None of the authors declare any proprietary interests

References

1. Arditi A, Cagenello R (1993) On the statistical reliability of letter-chart visual acuity measurements. Invest Ophthalmol Vis Sci 34:120–129

2. Klein R, Wang Q, Klein BE, Moss SE, Meuer SM (1995) The relationship of age-related maculopathy, cataract, and glaucoma to visual acuity. Invest Ophthalmol Vis Sci 36:182–191

3. Lalwani GA, Rosenfeld PJ, Fung AE, Dubovy SR, Michels S, Feuer W, Davis JL, Flynn HW Jr, Esquiabro M (2009) A variable-dosing regimen with intravitreal ranibizumab for neovascular age-related macular degeneration: year 2 of the PrONTO Study. Am J Ophthalmol 148:43–58

4. Neelam K, Nolan J, Chakravarthy U, Beatty S (2009) Psychophysical function in age-related maculopathy. Surv Ophthalmol 54:167–210

5. Rosser DA, Cousens SN, Murdoch IE, Fitzke FW, Laidlaw DA (2003) How sensitive to clinical change are ETDRS logMAR visual acuity measurements? Invest Ophthalmol Vis Sci 44:3278–3281

6. Midena E, Angeli CD, Blarzino MC, Valenti M, Segato T (1997) Macular function impairment in eyes with early age-related macular degeneration. Invest Ophthalmol Vis Sci 38:469–475

7. Rubin GS, Turano K (1994) Low vision reading with sequential word presentation. Vision Res 34:1723–1733

8. Sunness JS, Applegate CA, Haselwood D, Rubin GS (1996) Fixation patterns and reading rates in eyes with central scotomas from advanced atrophic age-related macular degeneration and Stargardt disease. Ophthalmology 103:1458–1466

9. Patel PJ, Chen FK, Da Cruz L, Rubin GS, Tufail A, ABC Trial Study Group (2011) Contrast sensitivity outcomes in the ABC Trial: a randomized trial of bevacizumab for neovascular age-related macular degeneration. Invest Ophthalmol Vis Sci 52:3089–3093

10. Patel PJ, Bunce C, Tufail A (2008) A randomised, double-masked phase III/IV study of the efficacy and safety of Avastin(R) (Bevacizumab) intravitreal injections compared to standard therapy in subjects with choroidal neovascularisation secondary to age-related macular degeneration: clinical trial design. Trials 9:56

11. Patel PJ, Chen FK, Rubin GS, Tufail A (2008) Intersession repeatability of visual acuity scores in age-related macular degeneration. Invest Ophthalmol Vis Sci 49:4347–4352

12. Patel PJ, Chen FK, Rubin GS, Tufail A (2009) Intersession repeatability of contrast sensitivity scores in age-related macular degeneration. Invest Ophthalmol Vis Sci 50:2621–2625

13. Patel PJ, Chen FK, Da Cruz L, Rubin GS, Tufail A (2011) Test-retest variability of reading performance metrics using MNREAD in patients with age-related macular degeneration. Invest Ophthalmol Vis Sci 52:3854–3859

14. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, Roberts C, Shoukri M, Streiner DL (2011) Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 64:96–106

15. Bird AC, Bressler NM, Bressler SB, Chisholm IH, Coscas G, Davis MD, de Jong PT, Klaver CC, Klein BE, Klein R (1995) An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. Surv Ophthalmol 39(5):367–374

16. Carkeet A (2001) Modeling logMAR visual acuity scores: effects of termination rules and alternative forced-choice options. Optom Vis Sci 78:529–538

17. Elliott DB, Sanderson K, Conkey A (1990) The reliability of the Pelli-Robson contrast sensitivity chart. Ophthalmic Physiol Opt 10:21–24

18. Bland JM, Altman DG (1986) Statistical method for assessing agreement between two methods of clinical measurement. Lancet i:307–310

19. British Standards Institution (1979) Precision of test methods, part 1: guide for the determination of repeatability and reproducibility for a standard test method. BS 5497, Part 1. BSI, London

20. Tulving E, Gold C (1963) Stimulus information and contextual information as determinants of tachistoscopic recognition of words. J Exp Psychol 66:319–327

21. Doris N, Hart PM, Chakravarthy U, McCleland J, Stevenson M, Hudson C, Jackson J (2001) Relation between macular morphology and visual function in patients with choroidal neovascularisation of age-related macular degeneration. Br J Ophthalmol 85:184–188

22. Kiser AK, Mladenovich D, Eshraghi F, Bourdeau D, Dagnelie G (2005) Reliability and consistency of visual acuity and contrast sensitivity measures in advanced eye disease. Optom Vis Sci 82(11):946–954

23. Blackhurst DW, Maguire MG (1989) Reproducibility of refraction and visual acuity measurement under a standard protocol. The Macular Photocoagulation Study Group. Retina 9(3):163–169