Optometry and Vision Science 2014

An Evaluation of Two Candidate Functional Biomarkers for AMD

Miss Claire Mckeague BSc (Hons)[1]

Dr Alison M Binns PhD BSc (Hons)[2]

Dr Tom H Margrain PhD BSc (Hons)[1]

Affiliations:

1 School of Optometry and Vision Sciences, Cardiff University, UK

2 School of Health Sciences, City University London

Correspondence:

Tom H. Margrain

School of Optometry and Vision Sciences

Cardiff University

Maindy Road

Cathays, Cardiff, CF24 4HQ, United Kingdom

E-mail: margrainth@cf.ac.uk

Fax: +44 (0)29 2087 4859

Number of Tables and Figures: 5

Submission Date: 12 December 2013

Abstract

PURPOSE

To evaluate the inter-session repeatability of the Colour Assessment and Diagnosis (CAD) test and a novel 14-Hz flicker test in a population of healthy participants in order to provide benchmark data for their use as functional biomarkers for age-related macular degeneration (AMD).

METHODS

Visual function was assessed using both techniques in 30 healthy adults (mean age 36.3 ± 14.1 years) on 2 separate days. Inter-session repeatability of RG and YB CAD thresholds and 14-Hz flicker thresholds was assessed by determining their coefficient of repeatability (CoR).

RESULTS

The CoR was calculated to be 0.39 CAD units (17.0%) for RG thresholds, 0.43 CAD units (31.1%) for YB thresholds and 0.015 (53.4%) for 14-Hz flicker contrast thresholds. On average, thresholds improved by 4.72% (RG), 6.33% (YB) and 13.3% (14-Hz flicker) between visits 1 and 2, suggesting a small but consistent learning effect. The CoR for all parameters was relatively small compared to the mean thresholds obtained (RG: mean 2.27 ± 4.58, CoR 0.39; YB: mean 1.37 ± 0.55, CoR 0.43; 14-Hz flicker: mean 0.028 ± 0.01, CoR 0.015).

CONCLUSIONS

This study has described the repeatability of the CAD and 14-Hz flicker tests. The data can help clinicians decide if the results from repeated measures are of clinical significance. Despite pre-test training, there was some evidence of a learning effect. Therefore, clinical trials using these techniques should ensure training is sufficient to minimize these effects.

KEY WORDS: repeatability, age-related macular degeneration, biomarkers, color vision, flicker

INTRODUCTION

Age-related macular degeneration (AMD) is a disorder of the central retina that is characterized by progressive dysfunction and death of photoreceptor cells. It is thought that 7.2 million people in the United States suffer from some form of AMD[1], whilst 56% of registrations as sight impaired in the United Kingdom are attributable to the condition[2], and over 50% of people aged over 65 in Europe are believed to have signs of AMD[3]. On a personal level, the condition is associated with an increased risk of falls, depression, and increased difficulty carrying out daily tasks[4-6]. The economic costs are also substantial, for example, AMD is estimated to cost the US economy $30 billion per annum[7]. Whilst anti vascular endothelial growth factor (anti VEGF) treatment provides a means of treating neovascular AMD (nAMD), it is both expensive and invasive. Giving up smoking, adopting healthy diets and consuming antioxidants can reduce AMD progression, but a treatment for early AMD and geographic atrophy is absent. It is not, therefore, surprising that a substantial research effort is being directed towards the development of new treatments for AMD.

Early AMD develops very slowly over time[8] and, therefore, it is not practicable to use end stage disease as an outcome measure for Phase II trials of new interventions. This necessitates the identification of biomarkers which may be used as surrogate outcome measures in clinical trials. The key requirements of these biomarkers are (i) that they must be sensitive to disease progression, and (ii) they must have a high level of inter-session repeatability. The development of tests sensitive to early AMD, and to disease progression is also a necessity in the early diagnosis and monitoring of patients with AMD in a primary care setting.

The standard psychophysical test of visual function used in clinical trials and in optometric practice is visual acuity. However, the high contrast visual acuity test does not meet either criterion for an optimal biomarker [9,10]. Whilst VA is substantially reduced by advanced AMD, during the earlier stages of the disease process it remains relatively unaffected[9]. This may be partially attributable to the relative sparing of the fovea in early stage disease[11], but is also likely to be due to the inherent variability in the test results (the between session coefficient of repeatability is around 0.15

logMAR / 1.5 lines for a standard logMAR test[10]). Hence, recent cross-sectional studies have evaluated a range of alternative functional biomarkers for AMD[12-16].

Numerous studies have found that temporal sensitivity is adversely affected by AMD[14,15,17-24], to a greater extent than the generalized loss which occurs due to normal aging[25]. This is thought to be due to the compromised outer retinal oxygen supply in AMD being unable to meet the increased metabolic demand elicited by flickering stimuli[26,27]. *Flicker frequencies of above 10Hz have been shown to increase the difference in oxygen tension between retinal arterial and venous blood substantially more than lower frequencies[28]. This indicates that the metabolic activity of the retinal tissue is upregulated in response to this high temporal frequency stimulation. Given the recent evidence to suggest that early functional changes in AMD are initiated by chronic retinal ischemia[29], a functional test which causes a greater demand on the retinal oxygen metabolism is more likely to detect the ischemic deficits in early AMD.*

The threshold for flicker detection is a desirable test to use when monitoring functional changes in AMD as it can be performed quickly, is reproducible and diagnostically sensitive[14,22]. For this reason, Dimitrov et al. rated 14-Hz flicker threshold measurement as having the greatest potential clinical value out of a battery of functional tests in the diagnosis and monitoring of AMD[14]. Furthermore, flicker threshold has also been shown to increase gradually with disease progression[15].

An increase in chromatic thresholds, especially in the yellow-blue (YB or tritan) domain, is another functional change which has long been reported to occur in the early stages of AMD[12,30-36]. Tritan color contrast thresholds are abnormal in patients with AMD and minimal lens opacities[36], and they also change significantly over time in patients with early AMD compared with age-matched controls[35]. However, for chromatic sensitivity to be employed as a functional biomarker of AMD, a means of accurately quantifying chromatic thresholds is required, which falls beyond the remit of standard clinical color vision tests. A computer-based technology, known as the Color Assessment and Diagnosis (CAD) test, has been developed which implements dynamic luminance contrast noise, in order to isolate red-green (RG) and YB thresholds[37-39]. This allows a rapid quantification of thresholds along 16 different directions in color space[40]. Using the CAD test, YB thresholds in patients with AMD

have been shown to increase linearly with disease severity[12]. O'Neill-Biba et al. reported evidence of an elevation in threshold, even when the retina appeared normal, in individuals whose fellow eye demonstrated signs of advanced AMD, indicating that impaired color vision may be an early functional indicator of retinal dysfunction in AMD[12]. Barbur et al. evaluated an approach to maximizing the diagnostic sensitivity of the test through the calculation of an index representing chromatic threshold as a function of light level in the low photopic, high mesopic range[41]. This resulted in a reduction in the substantial between subject variability in chromatic thresholds conferred by individual differences in factors such as media opacity, pupil diameter and macular pigment optical density, and removed the effect of age on color vision in healthy individuals. However, to date, no data have been published regarding the between session variability of the CAD test.

It is clear that both the 14-Hz flicker and CAD chromatic sensitivity tests may be useful as functional biomarkers in future clinical trials, fulfilling the first requirement of showing a sensitivity to increased severity of funduscopic changes associated with AMD[12,14,15]. Repeatability data have recently been published for the assessment of cone dark adaptation[42], another potentially important biomarker for early AMD[13-16,33,43-47]. However, there is currently little published data regarding the inter-session repeatability of the flicker and chromatic threshold assessment techniques. This is crucial in determining the minimum change in each parameter which may be considered to be clinically significant – an important issue when powering trials and interpreting outcomes, as well as in the clinical management of patients with early AMD.

The aim of this study was to assess the inter-session repeatability of the color assessment and diagnosis (CAD) test and the 14-Hz flicker test in a population of healthy participants.

METHODS

Participants

Adults with limited experience in psychophysical experiments were recruited to the study from the staff and students at the School of Optometry and Vision Sciences, Cardiff University. Thirty healthy adults (13 female), aged 22-72 years (mean 36.3 ±

14.1 years) took part in the study. This study was powered to detect within subject standard deviation to within 25% of the true population value[48]. All participants had corrected visual acuity of 20/20 or better (logMAR 0.0) in their test eye, age-normal lens clarity and a normal retinal appearance with no history of any ocular or systemic disease known to affect visual function. All participants had a LOCS score of 0 for all parameters, apart from RE, who had NO2 and NC2 (LOCS III)[49]. As a random sample of the population was desired, subjects were not excluded on the basis of having a color vision defect. The School's Research Ethics Committee approved the study and all procedures were carried out in accordance with the tenets of the Declaration of Helsinki. All participants provided written consent to taking part in the study, having received an information sheet prior to their appointment and having had the opportunity to ask any questions.

Experimental procedure

All participants attended the laboratory on two separate days within a period of two weeks. Screening data were obtained at the beginning of the first session to ensure that eligibility criteria were met. This included patient ocular and medical history, logMAR visual acuity (ETDRS chart), and fundus imaging (Optical Coherence Tomography and fundus photography; Topcon 3D OCT 1000). Lens clarity was assessed using a slit lamp biomicroscope, and graded according to the LOCS III system for nuclear opalescence (NO), nuclear color (NC), cortical opacity (C) and posterior subcapsular opacity (P)[49].

Stimuli for both psychophysical tests were presented on a calibrated, high-resolution 24" widescreen LCD monitor (NEC MultiSync PA241W) with a frame rate of 60Hz, as depicted in Figure 1. The luminance of the monitor was ϒ-corrected[50]. In a dimly illuminated room, participants were positioned 1.4 m away from the monitor, and any required refractive correction, appropriate for the viewing distance, was provided. The test eye was the eye with better visual acuity or, in the case of equal acuity, the right eye was selected. The fellow eye was occluded. The test order was randomized between subjects, but kept the same on both visits for each subject.

<u>14-Hz flicker sensitivity</u>

Flicker thresholds were determined using the well-established Bayesian adaptive psychometric method known as QUEST[51,52]. In this method, the strength of each successive stimulus presentation is set to match the current most probable estimate of threshold. In practice, QUEST was implemented in Matlab (The Math Works Inc.) using routines available within Psychophysics Toolbox to drive a yes / no adaptive staircase[53]. The results from a practice run that included 10 trials were used as the starting point for a final threshold estimate that converged after 40 trials. False positive responses were deemed to be responses that occurred more than 1s after stimulus offset.

Subjects were asked to fixate the center of the screen where the test stimulus, a 4^o foveated Gaussian blob at a temporal frequency of 14Hz, was presented to the fovea for a duration of 2 seconds. The flickering stimulus was generated by modulating a luminance increment following a sinusoidal temporal profile. The mean luminance of the monitor was 51 cd/m^2 and the chromaticity co-ordinates were 0.305, 0.323. To ensure that participants could not anticipate the next presentation, the inter stimulus interval was varied randomly between 4 and 10 seconds. The participants received verbal instructions on how to perform the test before undertaking the familiarization trial. Their task was to press a button on a keypad as soon as they perceived a flickering stimulus in the center of the monitor. If more than one false positive response was made, the practice trial was repeated until they were able to complete the familiarization trial with a maximum of 1 false positive response.

Color Contrast Sensitivity

Color contrast sensitivity was assessed using the CAD test (v2.2.4, City Occupational Ltd). RG and YB color detection thresholds were measured by employing colored stimuli moving against an achromatic background. The background (chromaticity co-ordinates 0.305, 0.323; mean luminance 26 cd/m^2) comprised a checkerboard of 15x15 squares (total 3.3 degrees diameter), which fluctuated randomly in luminance above and below the average background level in order to generate dynamic luminance contrast noise. The check luminance was distributed with equal probability within +/- 55% of background luminance. This noise masked the detection of residual luminance contrast cues in the isoluminant colored stimulus. The color-defined stimulus comprised a checkerboard of 5 × 5 squares (total 1.1° diameter) moving

diagonally across the checkerboard, in one of four directions. The stimulus duration was 600ms. A four-alternative forced choice procedure was used, whereby the participant was required to press a button indicating the direction of movement. Displacement thresholds were measured in 16 directions in color space (6 red, 6 green, 2 blue, 2 yellow), with color directions selected to correspond to the red / green color confusion lines (140 to 175 degrees) and the S-cone isolating axes (58 to 68 degrees). Threshold was determined using a two-down, one-up staircase in which color intensity was reduced by an initial step size of 0.006 CD units until the colored stimulus could not be distinguished from the background by the observer. This staircase procedure was repeated for nine reversals, at each of which the step size was reduced by 0.001 CD units until a final step size of 0.002 CD units was attained. Thresholds were obtained by averaging the chromatic distance in the CIE color space during the last four staircase reversals.

The participant's task was to press one of four buttons on a keypad to indicate the perceived direction of motion of the colored stimulus. Each stimulus presentation was followed by an audible 'bleep' to indicate when to respond. A response was required, even if the participant was uncertain of the direction of movement. Any trial could be presented for a second time at the participant's request. A familiarization trial lasting approximately 1 minute was performed prior to commencing the main trial. 100% correct response was required in the learning test to ensure that the subject understood the requirements of the test. The 'definitive' CAD program was then implemented and RG and YB thresholds were measured over a of 12 to 15 minute period.

Statistical Analysis

Flicker thresholds were transformed into Weber contrast values by dividing pedestal luminance ($I - I_b$) by the average luminance (I_b). The repeatability of the color and flicker thresholds was assessed using established statistical techniques[55]. The coefficient of repeatability (CoR) was calculated by multiplying the standard deviation of the differences between the two visits by 1.96. Confidence intervals for the CoR were calculated according to the method described by Bland and Altman[55].

RESULTS

Chromatic sensitivity and flicker thresholds were successfully obtained from all 30 participants on two separate days. Data from the 2 visits were generally collected on successive days but always within two weeks. None of the participants required additional practice sessions for either test, which minimized potential inter-individual differences in any learning effect. An example of the flicker data obtained on both visits from a typical participant (AB) is shown in Figure 2. In each plot, the solid horizontal line represents the final threshold and the dashed horizontal lines denote the 95% confidence intervals. Sample CAD results from the same observer are shown in Figure 3.

Only 1 subject (TM) with a congenital protanopic deficiency had a RG CAD threshold outside of the age-corrected statistically determined normal limits[56]. Similarly, only 1 subject (RE) had YB thresholds outside of the normal range. The lens opacities of this 72-year old participant had been graded as NO2 and NC2, so this YB defect is most likely due to the early stages of nuclear cataract. Both of these participants, whilst falling outside of the published limits of normality[56], showed repeatable results.

The difference in RG thresholds recorded at the first and second visits is plotted as a function of the mean RG threshold for all 30 participants in the Bland and Altman plots shown in Figure 4a, whereas Figure 4b shows the Bland Altman plot for RG thresholds with the protanopic individual's data point removed to aid visualization of the spread of the other data. Similar plots for all 30 individuals are shown for YB and 14-Hz flicker thresholds in Figure 4 c and d.

In each graph, the solid horizontal line depicts the bias, i.e. the mean difference between the two visits, and the dashed horizontal lines represent the 95% limits of agreement, i.e. the mean difference ± the coefficient of repeatability (CoR). These plots describe the between session repeatability for all 3 measures. There was no evidence of a systematic change in repeatability with increasing thresholds (i.e. no heteroscadicity). The bias line crosses the y-axis slightly above 0 in all cases. Relative to visit 1, thresholds improved by 4.72%, 6.33% and 13.3% for RG, YB and 14-Hz Flicker respectively, indicating the presence of a possible small learning effect.

The mean RG, YB chromatic thresholds and 14-Hz flicker thresholds for visits one and two are shown in Table 1, along with the CoR for each test. The expression of the CoR as a percentage of the group averaged test result (at visits 1 and 2) allows a direct comparison of the repeatability of parameters with different units. Although the RG thresholds were more repeatable than the YB thresholds, the difference in the CoR was not significant (95% confidence intervals did not overlap). There was also no significant difference in repeatability between the YB CAD thresholds and 14-Hz flicker. However, the CoR for the RG CAD thresholds was significantly better than that of the 14-Hz flicker (see Table 1). Scatter plots showing the effect of age on the between visit variability are shown in Figure 5. There was no evidence of any systematic effect of age on variability for any parameter.

DISCUSSION

In order to monitor the progression of AMD and determine the efficacy of novel therapies, functional biomarkers must be identified that are reliable, repeatable and clinically applicable. This will allow candidate treatments to be assessed with maximum efficiency by minimizing the sample size and follow-up duration required to achieve a useful end point. The development of functional tests sensitive to subtle changes in AMD status is also important in the clinical diagnosis and management of the condition in clinical settings. Visual acuity, despite the common acknowledgement that it is a poor assay of early AMD, is still the standard functional vision test amongst both clinicians and researchers. It is therefore necessary that new functional tests are developed that are as quick to perform and have the same ease of use as VA, but with improved sensitivity to disease progression and better inter-session repeatability. Two such tests that have been shown to be sensitive to disease severity in AMD are the 14-Hz flicker and CAD chromatic sensitivity test[14,15]. The flicker test employs a stimulus which is bigger (4 degrees diameter) than the stimulus presented in the CAD test (1.1 degree diameter). However, the CAD stimulus moves out from a central fixation position to a location extending to 2.3 degrees into the parafovea. Hence both stimuli are assessing a region of the macula extending to around 2 degrees from fixation. This targets the parafoveal region in which functional deficits have been identified early in the AMD disease process[43].

The coefficient of repeatability (CoR) is an important statistical technique due to its potential to describe the smallest change that can be deemed clinically significant [55]. This is helpful in identifying those individuals who have shown a "clinically significant decline" in performance, and can therefore be used to determine the optimal sample size for a trial, i.e. it can be powered to detect a certain percentage who show this level of functional decline. The most repeatable test was found to be the RG CAD threshold test. This performed significantly better than the 14-Hz flicker test which produced the least repeatable results.

The Bland Altman plots all showed a mean difference between visits that was slightly above zero, suggesting a small learning effect for both the 14-Hz flicker test and the CAD parameters. This was confirmed by a post hoc paired samples *t*-test ($p < 0.05$) for all tests. This learning effect may have been limited by the familiarization trials which were carried out for the two techniques before both visits. If more than 1 false positive occurred on the 14-Hz flicker practice trial lasting 1 minute or if the subject did not score 100% in the CAD practice trial which also took 1 minute to complete, they were made to repeat it until they achieved the required standard and were deemed competent in task performance. However, the familiarization trials were clearly not sufficient to saturate learning.

A limitation of the study is that different repeatability values will need to be established if the tests are applied under different experimental conditions. A change in stimulus size, eccentricity, temporal frequency, retinal illuminance, or a change in the psychophysical procedure used, are all likely to affect the measured variability of the techniques. For example, in their recent evaluation of the effect of retinal illuminance on chromatic thresholds, Barbur et al. hypothesized that the assessment of color vision at mesopic levels may increase the diagnostic sensitivity of the test, through the exacerbation of the effect of disease-related hypoxia [41]. Their 'healthy retina index' (HR_{index}) is a measure of the effect of retinal illuminance on chromatic thresholds. Additional repeatability data will be required to evaluate the clinical interpretation of mesopic chromatic thresholds and the HR_{index}. Inter-session repeatability is also likely to be influenced by the characteristics of the patient population. Hence, a further potential limitation of the repeatability data reported in this study is that the participant cohort was recruited from a University environment,

and may not be generalizable to the population of patients with age-related macular degeneration. However, the age-range of participants extended to 72 years, and none were experienced observers in psychophysical experiments, so they may be considered to be broadly representative of naïve participants in a clinical environment. Furthermore, we found no evidence of an effect of age on the between session variability, suggesting that the findings of this study will be broadly applicable across age groups.

One limitation of the yes / no adaptive staircase procedure used in the flicker sensitivity test is that results are dependent on stimulus strength and an individual's response criterion i.e. their willingness to guess. Response criterion can vary between and within individuals. We attempted to minimize within subject changes in response criteria by providing identical instructions at each visit. However, we cannot rule out the possibility that the systematic difference between visits (i.e. the bias) was due to a change in response criterion. Many things, including instructions, can induce the observer to raise or lower his or her criterion, causing threshold to shift up or down. This unknown internal criterion of the observer typically differs among observers and may vary across populations and over time. The four alternative forced choice paradigm employed by the CAD test negates the effect of inter-individual differences in the response criterion.

The published limits of normality for the CAD test are based on data collected from 250 color normal participants[56]. The majority of participants in this study produced thresholds which fell within these limits, apart from one protanope (TM), and one older participant with significant nuclear lens opacities (RE). Excluding these 2 participants, the mean (SD) RG thresholds for visit 1 were 1.45 (0.29) and for YB 1.34 (0.37). The RG threshold is very similar to that reported by O'Neill-Biba et al[12] but the YB value is somewhat lower than that reported previously 1.6 (0.15). Control participants in the O'Neill-Biba study were on average 20 years older than those studied here and increasing lens opacification may therefore, explain the difference. Barbur et al reported that chromatic thresholds, uncorrected for differences in media absorption and pupil diameter, increase significantly with increasing age in the healthy population[41].

In summary, this study has described the inter session repeatability of two tests that may be used in the diagnosis and monitoring of AMD. Both color vision and flicker sensitivity tests have been shown to have excellent diagnostic capacity[12,14,15,17-24,30-36]. The results of this study will help clinicians to determine if changes observed over time are due to measurement imprecision or disease progression, provided that the experimental conditions and psychophysical procedures are kept constant. The observation that a small but significant learning effect exists highlights the need for control groups in clinical trials of new AMD therapies. These and other candidate biomarkers must now be validated in longitudinal studies to confirm their prognostic and predictive capabilities.

ACKNOWLEDGMENTS

This study was funded by a research grant from the College of Optometrists, United Kingdom.

REFERENCES

1. Klein R, Chou C-F, Klein BEK, Zhang X, Meuer SM, Saaddine JB. Prevalence of age-related macular degeneration in the US population. Arch Ophthalmol 2011;129:75–80.

2. Bunce C, Xing W, Wormald R. Causes of blind and partial sight certifications in England and Wales: April 2007-March 2008. Eye 2010;24:1692–1699.

3. Augood CA, Vingerling JR, de Jong PTVM, et al. Prevalence of age-related maculopathy in older Europeans: the European Eye Study (EUREYE). Arch Ophthalmol 2006;124:529–535.

4. Brody BL, Gamst AC, Williams RA, et al. Depression, visual acuity, comorbidity, and disability associated with age-related macular degeneration. Ophthalmology 2001;108:1893–9001.

5. Ivers RQ, Cumming RG, Mitchell P, Simpson JM, Peduto AJ. Visual risk factors for hip fracture in older people. J Am Geriatr Soc 2003;51:356–363.

6. Lamoureux EL, Hassell JB, Keeffe JE. The determinants of participation in activities of daily living in people with impaired vision. Am J Ophthalmol 2004;137:265–270.

7. Brown GC, Brown MM, Sharma S, et al. The burden of age-related macular degeneration: a value-based medicine analysis. Trans Am Ophthalmol Soc 2005;103:173–86

8. Davis MD, Gangnon RE, Lee L-Y, et al. The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS Report No. 17. Arch Ophthalmol 2005;123:1484–1498.

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

10. Siderov J, Tiu AL. Variability of measurements of visual acuity in a large eye clinic. Acta Ophthalmol Scand 1999;77:673–676.

11. Sunness JS, Rubin GS, Zuckerbrod A, Applegate CA. Foveal-Sparing Scotomas in Advanced Dry Age-Related Macular Degeneration. J Vis Impair Blind 2008;102:600–610.

12. O'Neill-Biba M, Sivaprasad S, Rodriguez-Carmona M, Wolf JE, Barbur JL. Loss of chromatic sensitivity in AMD and diabetes: a comparative study. Ophthalmic Physiol Optic 2010;30:705–716.

13. Gaffney AJ, Binns AM, Margrain TH. The topography of cone mediated dark adaptation in age-related maculopathy. Optom Vis Sci 2011;88:1080–1087.

14. Dimitrov PN, Robman LD, Varsamidis M, et al. Visual function tests as potential biomarkers in age-related macular degeneration. Invest Ophthalmol Vis Sci 2011;52:9457–9469.

15. Dimitrov PN, Robman LD, Varsamidis M, et al. Relationship between clinical macular changes and retinal function in age-related macular degeneration Invest Ophthalmol Vis Sci 2012;53:5213–5220.

16. Gaffney AJ, Binns AM, Margrain TH. The effect of pre-adapting light intensity on dark adaptation in early age-related macular degeneration. Doc Ophthalmol 2013;127:191-199

17. Brown B, Lovie-Kitchin J.E. Temporal function in age-related maculopathy. Clinical and Experimental Optometry 1987;70:112-116.

18. Eisner A, Klein ML, Zilis JD, Watkins MD. Visual function and the subsequent development of exudative age-related macular degeneration. Invest Ophthalmol Vis Sci 1992;33:3091-3102.

19. Mayer MJ, Spiegler SJ, Ward B, Glucs A, Kim CB. Foveal flicker sensitivity discriminates ARM-risk from healthy eyes. Invest Ophthalmol Vis Sci 1992;33:3143–3149.

20. Mayer MJ, Spiegler SJ, Ward B, Glucs A, Kim CB. Mid-frequency loss of foveal flicker sensitivity in early stages of age-related maculopathy. Invest Ophthalmol Vis Sci 1992;33:3136–3142.

21. Phipps JA, Guymer RH, Vingrys AJ. Loss of Cone Function in Age-Related Maculopathy. Invest Ophthalmol Vis Sci 2003;44:2277–2283.

22. Phipps JA, Dang TM, Vingrys AJ, Guymer RH. Flicker perimetry losses in age-related macular degeneration. Invest Ophthalmol Vis Sci 2004;45:3355–3360.

23. Luu CD, Dimitrov PN, Robman L, et al. Relationship between Clinical Macular Changes and Retinal Function in Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci 2012;53:5213-5220

24. Luu CD, Dimitrov PN, Wu ZC, et al. Static and Flicker Perimetry in Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci 2013;54:3560-3568

25. Kim CB, Mayer MJ. Foveal flicker sensitivity in healthy aging eyes. II. Cross-sectional aging trends from 18 through 77 years of age. J Opt Soc Am A Opt Image Sci Vis 1994;11:1958–1969.

26. Kiryu J, Asrani S, Shahidi M, Mori M, Zeimer R. Local response of the primate retinal microcirculation to increased metabolic demand induced by flicker. Invest

Ophthalmol Vis Sci 1995;36:1240–1246.

27. Riva CE, Falsini B, Logean E. Flicker-evoked responses of human optic nerve head blood flow: luminance versus chromatic modulation. Invest Ophthalmol Vis Sci 2001;42:756–762.

28. Shakoor A, Blair NP, Mori M, Shahidi M. Chorioretinal vascular oxygen tension changes in response to light flicker. Invest Ophthalmol Vis Sci 2006;47:4962-4965

29. Feigl B, Brown B, Lovie-Kitchin J, Swann P. Functional loss in early age-related maculopathy: the ischaemia postreceptoral hypothesis. Eye 2007;21:689-696

30. Bowman KJ. The effect of illuminance on color discrimination in senile macular degeneration. Mod Probl Ophthalmol 1978;19:71–76.

31. Collins MJ. Pre-age related maculopathy and the desaturated D-15 color vision test. Clin Exp Optom 1986;69:223–227.

32. Applegate RA, Adams AJ, Cavender JC, Zisman F. Early color vision changes in age-related maculopathy. Appl Opt 1987;26:1458.

33. Eisner A, Stoumbos VD, Klein ML, Fleming S. Relations Between Fundus Appearance and Function. Invest Ophthalmol Vis Sci 1991;32:8–20.

34. Frennesson C, Nilsson UL, Nilsson S. Color contrast sensitivity in patients with soft drusen, an early stage of ARM. Doc Ophthalmol 1995;90:377-386

35. Holz FG, Gross-Jendroska M, Eckstein A, Hogg CR, Arden GB, Bird AC. Color contrast sensitivity in patients with age-related Bruch's membrane changes. Ger J Ophthalmol 1995;4:336–341.

36. Arden GB, Wolf JE. Color vision testing as an aid to diagnosis and management of age related maculopathy. Br J Ophthalmol 2004;88:1180–1185.

37. Birch J, Barbur JL, Harlow AJ. New method based on random luminance masking for measuring isochromatic zones using high resolution color displays. Ophthalmic Physiol Opt 1992;12:133–136.

38. Barbur JL, Harlow AJ, Plant GT. Insights into the different exploits of color in the

visual cortex. Proc R Soc Lond 1994;258:327–334.

39. Barbur JL. “Double-blindsight” revealed through the processing of color and luminance contrast defined motion signals. Prog Brain Res 2004;144:243–259.

40. Barbur J, Rodriguez-Carmona M, Evans S, Milburn NJ. Minimum color vision requirements for professional flight crew, part III: recommendations for new color vision standards. FAA, 2009. Available at: http://www.caa.co.uk/docs/33/200904.pdf. Accessed November 6, 2013.

41. Barbur, JL, Konstantakopoulou E. Changes in color vision with decreasing light level: separating the effects of normal aging from disease. J Opt Soc Am A Opt Image Sci Vis 2012 29 (2):27-35.

42. Gaffney AJ, Binns AM, Margrain TH. Measurement of cone dark adaptation: a comparison of four psychophysical methods. Doc Ophthalmol 2014;128:33–41

43. Owsley C, Jackson GR, White M, Feist R, Edwards D. Delays in rod-mediated dark adaptation in early age- related maculopathy. Ophthalmology 2001;108:1196–1202

44. Binns AM, Margrain TH (2007) Evaluating retinal function in age-related maculopathy with the ERG photostress test. Invest Ophthalmol Vis Sci 2007;48:2806–2813

45. Owsley C, McGwin G Jr, Jackson GR, Kallies K, Clark M. Cone- and rod-mediated dark adaptation impairment in age-related maculopathy. Ophthalmology 2007;114:1728–1735

46. Brown B, Kitchin JL. Dark adaptation and the acuity/luminance response in senile macular degeneration (SMD). Am J Opt Physiol Opt 1983. 60:645–650

47. Eisner A, Fleming SA, Klein ML. Mauldin WM. Sensitivities in older eyes with good acuity: eyes whose fellow eye has exudative AMD.Invest Ophthalmol Vis Sci 1987;28:1832–1837

48. Bland M. How can I decide the sample size for a repeatability study? University of York, 2010. Available at: http://www-users.york.ac.uk/~mb55/meas/sizerep.htm.

Accessed November 6, 2013.

49. Chylack LJ, Wolfe JK, Friend J, et al. Quantitating cataract and nuclear brunescence, the Harvard and LOCS systems. Optom Vis Sci 1993;70:886–895.

50. Metha AB, Vingrys AJ, Badcock DR. Calibration of a color monitor for visual psychophysics. Behav Res Meth Instrum Comput 1993;25:371–383.

51. Watson AB, Pelli DG. Quest: A Bayesian adaptive psychometric method. Percept Psychophys 1983;33:113–120.

52. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 1997;10:437–442.

53. Brainard DH. The psychophysics toolbox. Spat Vis 1997;10:433–436.

54. King-Smith PE, Grigsby SS, Vingrys AJ, Benes SC, Supowit A. Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. Vision Res 1994;34:885–912.

55. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.

56. Barbur JL, Rodriguez-Carmona M. Establishing the statistical limits of “normal” chromatic sensitivity. CIE Proceedings Expert Symposium "75 Years of the CIE Standard Colorimetric Observer" (CIE x030) 2006.

FIGURE LEGENDS

Figure 1. Images showing the appearance of the moving colored stimulus used in the CAD test [1] (left panel) and the flickering stimulus (right panel) used to determine flicker thresholds.

Figure 2. 14-Hz Flicker data for participant AB at visit 1 (left panel) and visit 2 (right panel), shown with the threshold in decibels. The dashed lines represent the 95% confidence intervals, with the solid line depicting the final threshold.

Figure 3. CAD data for participant AB at visit 1 (left panel) and visit 2 (right panel). The dotted black ellipse is based on the median RG and YB thresholds from 250 observers, with the grey shaded area representing the 95% limits of variability of these observers. The green, red and blue bands display the deuteranopic, protanopic and tritanopic confusion lines, respectively. The colored symbols show the data measured for participant AB.

Figure 4. Bland Altman plots for RG chromatic thresholds (A), RG chromatic thresholds excluding participant TM (B), YB chromatic thresholds (C) and 14-Hz flicker thresholds (D). The difference between the measurements from visit 1 to visit 2 is plotted as a function of the mean value for all 30 participants, and is shown with the bias (solid line) and 95% limits of agreement (dashed lines).

Figure 5. Scatter plots demonstrating the relationship between age and between visit threshold variation for RG chromatic thresholds (A), YB chromatic thresholds (B) and 14-Hz flicker thresholds (C). Note the lack of a systematic relationship with age for any parameter.