Quantitative Phenotyping of Chromatic Dysfunction in Best Macular Dystrophy

Sónia H. Campos, MD; Vasco Forjaz, BSc; Lajos R. Kozak, MD, MSc; Eduardo Silva, MD, PhD; Miguel Castelo-Branco, MD, PhD

Objectives: To quantify chromatic dysfunction in Best disease to reassess the classic categorization of macular chromatic damage and to investigate psychophysical and clinical correlations.

Methods: Color-contrast discrimination was measured using 2 different psychophysical strategies in age-matched control (n=41) and patient (n=34) eyes. The first strategy measured performance along 3 main confusion lines (testing cone function), and the second evaluated discrimination ellipses (modified Cambridge Color Test). The main outcome measures were chromatic discrimination variables (confusion line length, ellipse length, angle, and axis ratio) and visual acuity (VA).

Results: Significant loss of performance was seen in all color axes in our patients, and it increased monotonically with staging, becoming significant in Fishman stages 2 and 3. The classically assumed preferential type I red-green deficit was true only for stage 4. Substantial chromatic dysfunction occurred even with relatively preserved VA despite that negative correlations between all test variables and VA reached statistical significance. Partial correlation analysis showed that protan/deutan loss was related to VA independent of tritan loss. Statistically significant positive correlations were also found between lesion size and chromatic dysfunction.

Conclusions: Chromatic discrimination is often impaired in Best disease, even when VA is still spared. Our quantitative psychophysical approach shows that the classic categorization as a type I red-green deficit is valid only for disease stage 4.


Author Affiliations: Center for Ophthalmology (Drs Campos, Kozak, Silva, and Castelo-Branco and Mr Forjaz) and Department of Biophysics (Dr Castelo-Branco), Institute of Biomedical Research in Light and Image, Faculty of Medicine, University of Coimbra, Coimbra, Portugal; Association for Biomedical Research and Innovation in Light and Image, Coimbra (Mr Forjaz and Dr Kozak); and Institute for Psychology of the Hungarian Academy of Sciences, Budapest (Dr Kozak). Dr Campos and Mr Forjaz contributed equally to this work.

Financial Disclosure: None.

A recent article1 showed that computerized methods of studying color vision provide more information than the more usual evaluation using the Farnsworth-Munsell 100-hue (FM-100) test, which has been shown to be semiquantitative only.2 Farnsworth himself considered that 30% changes in test-retest could occur, and this has been repeatedly confirmed.2 Indeed, a difference in global scores is only significant if it is greater than 50.2 It is important to use novel approaches when quantifying and phenotyping chromatic damage in macular disorders and to reassess traditional clinical classifications using new quantitative criteria.1,4

Most of our current understanding of chromatic dysfunction in macular disorders comes from the seminal contributions of researchers such as Pokorny et al.,3,4 who established a clinical classification of dyschromatopsias that is now widely used. It is believed that most hereditary forms of macular disease, such as Stargardt disease, exhibit the so-called type I red-green deficit. Roth and Lanthony recently reviewed the current consensus on how dyschromatopsia evolves in different stages of Stargardt disease. They emphasized the difficulty in staging chromatic deficits even when using the FM-100 test, and they have mostly relied on colorimetric equations (such as the Raleigh equation) to describe that in the mentioned condition there occurs initially a stage of mild red-green deficit. At a later stage, a blue-yellow deficit is observed, as documented by the Moreland equation. In the final stage, functional achromatopsia may occur. However, colorimetric equations provide only indirect estimates of relative cone dysfunction.

Type I red-green deficits have also been attributed to other macular disorders, such as vitelliform macular dystrophy (VMD) (Best disease).2,5 New quantitative paradigms are, however, necessary to measure cone function in a more direct manner and to investigate whether such deficits are really type I red-green across disease stages. Best disease is an autosomal dominant disorder with variable expressivity5-13 and is characterized by the accumulation of a yellowish lipofuscin-like material within and beneath the retinal pigment epithelium (RPE).14-18 The “egg-yolk” or vitelliform lesion is easily visible on fundus examination and evolves...
through several stages across many years. Lesions may be central or paracentral, single or, rarely, multifocal, with foci at different stages.18

The VMD gene was isolated several years ago,19,20 and many mutations have been identified.21-23 It is likely that bestrophin (the VMD gene product) mutations lead first to alterations in the RPE, although one cannot exclude a direct role in photoreceptors. Bestrophin defines a new family of chloride-channel proteins26 and is in the signal transduction pathway that modulates the light peak of the electrooculogram (EOG).27 In VMD, a possible decrement in Cl– conductance occurs across the basolateral membrane of the RPE. The notion that the primary defect is located in the RPE is also suggested by the considerably reduced light peak to dark-through in clinical EOGs in the presence of normal global electroretinographic (ERG) findings.8,28-32 Multifocal ERG techniques can be useful in further isolating macular deficits in VMD.29-31 The multifocal ERG peak amplitudes of the central and pericentral responses are indeed significantly reduced in patients with VMD.29

The present study aims to evaluate simultaneously the function of L−, M−, and S-cone pathways in Best disease. It has been suggested that it is important to find useful measures of cone function in this condition because patients with VMD who have normal or near-normal Snellen visual acuity (VA) may have abnormal flicker fusion threshold intensities.33 Direct cone-probing measures can also be more easily related to other techniques used to measure macular damage.34-36 What is critical is to find better ways to detect and monitor disease progression in VMD because although EOG is helpful for diagnosis, it does not correlate with other clinical measures, such as VA.32 A recent ERG study30 suggested a relative change in M- and L-cone responses at high temporal frequencies, but because these objective responses are paradoxically normal or supernormal, it is difficult to relate them to psychophysical or clinical measures. We attempted to quantify the pattern of chromatic dysfunction in Best disease and to correlate discrimination measures—obtained along 3 chromatic confusion lines and confirmed by fitting ellipses along 8 confusion vectors—with anatomic, clinical, and electrophysiologic markers of disease progression.

**METHODS**

**PATIENT SELECTION AND CLASSIFICATION**

We included in this study 17 patients (34 tested eyes) whose diagnosis of Best disease was obtained based on the characteristic photographic appearance of the fundus and on changes in EOG findings (an Arden ratio <1.8 was considered abnormal). No patient revealed any other ophthalmologic or systemic conditions. Most patients had a positive family history of dominant inheritance (15 of 17 patients in 5 families). Patients were classified in accordance with the Fishman classification37: a fundus appearance develops from a normal fovea with abnormal EOG findings (stage 0); to minimal macular pigment mottingling and hypopigmentation (stage 1); to a typical egg-yolk vitelliform lesion, usually slightly elevated (stage 2); which then can break through the RPE and accumulate in the subretinal space in a cyst with a fluid level formed that moves with head position changes called pseudohypopyon; and follows various stages of resorption of the vitelliform lesion (stage 3); to resorption plus flat-to-wrinkle or supernormal-appearing scar formation with or without neovascular membranes (stage 4). Patient distributions across stages were as follows: stages 0/1, 6 eyes; stages 2/3, 14 eyes; and stage 4, 14 eyes. Visual acuity was determined using postcycloplegic manifest refraction on Snellen charts, in a masked manner. We used standard operational procedures for fundus photography following the guidelines of the Fundus Photograph Reading Center, Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison.

A population of 21 normal-sighted controls (+1 eyes) was also selected for statistical comparisons. The research followed the tenets of the Declaration of Helsinki, and informed consent was obtained from all participants in strict accordance to the Institute of Biomedical Research in Light and Image—Faculty of Medicine guidelines. Patient and control populations were age matched (mean±SD age, 29.558±14.894 and 28.024±9.940 years, respectively). Analysis of variance (ANOVA) showed no significant age difference between groups (P=.59).

**PSYCHOPHYSICAL METHODS AND ANALYSIS**

We used a slightly modified version of the Cambridge Color Test (Cambridge Research Systems Ltd, Rochester, England) to modulate chromaticity along selected axes in color space. Participants looked monocularly, with the refraction corrector for viewing distance, at a screen with a pattern of disks of varying sizes and luminances with superimposed chromatic contrast defined in a gap in a Landolt-like C-shaped ring (Figure 1) (viewing distance, 1.8 m; gap size, 1.60°; outer diameter, 7.60°; and inner diameter, 3.81°). Tinted contact and spectacle lenses were replaced by trial lenses in a trial frame. The chromaticity of the Landolt C shape was adjusted according to a staircase procedure. The participant had to indicate 1 of 4 possible positions (bottom, top, left, and right) of the gap of the Landolt C by pressing a 4-button response box (time out, −3 seconds). Luminance and size variation of stimulus disks forced the participant to use specific color cues because he or she could not use spatial or luminance cues to infer the embedded shape (luminance noise levels, 8, 10, 12, 14, 16, and 18 candela/m²). A minimum excursion of 0.002 CIE 1976 u′v′ color space units was superimposed on such noise levels to define the chromatic shape.

Quantitative modulation of chromatic contrast was performed in CIE 1976 u′v′ color space. Calibration procedures were performed using software and hardware provided by Cambridge Research Systems Ltd (Minolta colorimeter, VSG 2/5 graphics card, with 15-bit contrast resolution per pixel). Stimuli were displayed on a 21-inch monitor (GDM-F520; Sony, New York, NY), which was gamma corrected.

Psychophysical thresholds were obtained through 3 parallel, randomly interleaved staircases, from the Trivector version of the test, which assessed simultaneously the 3 cone confusion axes in color space. This ensures unbiased measurement of thresholds across different chromatic mechanisms. To determine discrimination ellipses, 8 confusion line vectors were measured in an interleaved random manner, with independent staircases running at a neutral background (neutral point coordinates: 0.1977, 0.4689 u′v′; minimum excursion: 0.002 CIE 1976 u′v′ color space units in this space; protan confusion (copunctal) point: 0.678, 0.501 u′v′; deutian confusion point: −1.217, 0.782 u′v′; tritan confusion point: 0.257, 0.0 u′v′; maximum excursion for the Trivector version: 0.1100 [Figure 2A]). Axis angles for ellipses: see Figure 1).

On each axis of assessment of relative damage, the separation between background/target chromaticities is initially large and is decreased/increased after each correct/incorrect response. Each staircase terminated after 11 reversals, and the mean of the last 7
reversals was taken as the threshold estimate. A small subset of random catch trials was included, and they provided the participant clear cases when he or she was near threshold.

The ellipse-fitting method used1 produces ellipses that are centered on the field point and that are obtained by minimizing the sum of squares of the log distances between the ellipse and the fitted point. We cross-validated this method using a custom procedure, which consisted of an equi-angled spline interpolation of the data points around the field point with the determination of the longest diameter of this spline curve and subsequent comparisons of the standard deviations of the data points parallel and perpendicular to this axis. The following quantitative results were analyzed: confusion line length, ellipse length, and axis ratio. Further statistical analyses included factorial and repeated-measures ANOVA, with post hoc Fisher protected least significant difference correction, and multiple linear regressions.

**RESULTS**

Determination of chromatic discrimination ellipses allows for detailed, unbiased assessment of chromatic function because in our procedure, thresholds are obtained simultaneously along multiple, evenly spaced axes in color space. Representative examples of deterioration of chromatic function during different stages of Best disease, as assessed by discrimination ellipses obtained from individual participants, are illustrated in Figure 1. These examples illustrate a steady elongation of the major axis of discrimination ellipses across the stages. It is also noticeable that the angle (orientation) of the ellipses, which is an indicator of the major axis of chromatic dysfunction, jitters markedly in patients in stages 1 and 2, and a predominant tilt toward the red-green axis becomes obvious only for patients in stages 3 and 4.

Because none of the axes used for determination of discrimination ellipses coincided with cone confusion lines (protan, deutan, and tritan, which are more direct measures of cone function), we also measured chromatic thresholds along these lines, which are related to short-, medium-, and long-wavelength selective cones, respectively. Comparison of chromatic loss along the 3 main confusion lines revealed a pattern of substantial functional impairment in Best disease (repeated-measures ANOVA, with a significant group effect: $P < .001$) (Figure 2B). All axes were significantly longer in the patient population than in the age-matched control population ($P < .001$ for the protan, deutan, and tritan axes, Fisher protected least significant difference). The effects remained significant when the analysis was conducted separately for patients’ right eyes ($P < .001$, ANOVA group effect) and left eyes ($P = .001$; even for individual eyes, comparisons split by axes remained in general with $P < .01$). The degree of protan and deutan loss was, on average, more prominent than the degree of tritan loss, but this effect was not significant on post hoc analyses, suggesting that the classic notion of a predominant red-green deficit has to be revised or, alternatively, that other factors, such as clinical staging, have to be considered (Figure 1).

The specificity of damage can be better judged by ana-
lyzing the length, axis ratio, and orientation (angle) of chromatic discrimination ellipses. We found a significant increase in mean ellipse length in the patient population (\(P < .001\), ANOVA group effect) (Figure 3A), which was accompanied by a significant enlargement of the major ellipse axis compared with the minor axis (axis ratio: \(P < .001\)) (Figure 3B). Stage had a significant effect on the deterioration of both measures (\(P < .001\) for both analyses) (Figure 4). Post hoc analysis of such stage effect (Fisher protected least significant difference) was significant not only between stages 1 and 4 but also between controls and stages 2/3 and between stages 2/3 and 4. Regarding the axis ratio, comparisons were significant between all groups and stage 4. The increased axis ratio suggests that although all axes are affected, there is some pattern of preferential damage. Such a finding does not necessarily mean that preferential damage occurs along specific axes in the early stages of Best disease. Indeed, although deviation of ellipse angles across stages shows a significant group effect (\(P = .01\) by ANOVA) (Figure 1), this effect could potentially be explained by late-stage patients. This notion is supported by the observation that the average peak at 87.13° that is measured in the distribution of the control group becomes significantly deviated only in the stage 4 group (Fisher protected least significant difference: \(P = .04\) for comparisons between stages 2/3 and 4; \(P = .03\) for comparisons between controls and stage 4). Taken together, these analyses suggest that the specificity of damage may tilt in multiple directions of color space until stage 3 and that the horizontal orientation becomes prominent only in stage 4, suggesting that protan and deutan losses are more important only in the late stages of this macular disorder.

Although we favor ellipse measurements to quantify relative patterns of damage across stages, the same analyses for protan, deutan, and tritan measures reveal a clearcut monotonic deterioration of chromatic function from early to late disease stages (\(P < .001\), by ANOVA). The post hoc analysis showed that the tritan measure was the only one being significantly elevated between controls and stages 2/3 (\(P = .02\)), unlike protan and deutan measures, which became significant only for comparisons that included stage 4 (\(P < .001\)). In other words, deficits are definitely not type I red-green in the early disease stages.

We examined the correlation between our psychophysical measures and clinical criteria. Strong and significant nega-
The results of the present study provide new insights into the nature of the functional deficits in the neurosensory retina in VMD. In particular, they challenge the view that chromatic deficits in VMD are type I red-green (as is the current view in Stargardt disease as well). This is true only for stage 4, and tritan deficits actually become statistically significant earlier than protan and deutan deficits, whereas the latter become more prominent only when the lesion size has increased or foveal involvement has emerged.

Computerized methods provide more information to study color vision than traditional tests, such as the FM-100 test, whose semiquantitative scores have weak reproducibility. Alternative approaches may, therefore, achieve better quantification of chromatic damage in macular disorders and may help validate traditional classifications, which rely mostly on the FM-100 test and colorimetric equations. This is emphasized in a study by Regan et al (1994), which provides a direct comparison between the FM-100 test and a computerized approach similar to ours.
Psychophysical evidence is important in the context of VMD because ERG responses are often normal or supernormal, which renders interpretation difficult. Indeed, a recent ERG study\textsuperscript{30} suggested that the M-cone response was normal and that the L-cone response was supernormal and phase delayed. It is, however, difficult to conclude from this study whether one cone population is more affected than the other because it is unclear whether a supernormal response implies more or less damage of photoreceptor function.

Our study, which uses a highly sensitive quantitative method, shows that substantial damage occurs in all cone pathways in Best disease, even in a subset of patients with relatively preserved VA. These results are compatible with the now almost consensual postulate that the primary disturbance occurs in the RPE.\textsuperscript{39}

Visual acuity has often been used as a clinical gold standard psychophysical measure, but options are needed to define better cutoff values between populations of controls and patients. All of our quantifiable variables showed statistically significant correlations with VA, but, most important, they proved to be more reliable in quantifying relative damage and in predicting disease progression. Furthermore, they correlated significantly with staging and the size of the retinal lesion.

In summary, our quantitative phenotyping strategy proved to be adequate in defining cone dysfunction in different stages of Best disease. We suggest that this type of investigation, complemented with novel physiologic approaches to study cone function,\textsuperscript{40,41} should be applied in relatives of patients with Best disease who show no apparent phenotype.

Submitted for Publication: December 15, 2003; final revision received July 15, 2004; accepted October 18, 2004.

Correspondence: Miguel Castelo-Branco, MD, PhD, Center for Ophthalmology, Institute of Biomedical Research in Light and Image, Faculty of Medicine, University of Coimbra, Az. de Sta Comba, Celas 3000-354, Coimbra, Portugal (mcbranco@ibili.uc.pt).

Funding/Support: This work was supported by grant POCTI/NSE/35823/2000/FEDER from Fundação para a Ciência e Tecnologia, grant POE 3.1/B1/03/109/2002 from Programa Operacional de Economia (Mr Forjaz), Programa Gulbenkian de Estímulo à Investigação (Drs Kozak and Castelo-Branco), and Infarmed (Dr Kozak), Lisbon, Portugal.

Acknowledgment: We thank Aldina Reis, MSc, and Rita Vaz, BSc, for help in performing electrophysiologic measurements and Mafalda Mendes, BSc, and Fátima Silva, MSc, for help with some experiments.

REFERENCES

6. Chermaine V, Roquezo G, Anglès A, Pierrerrat CO, Orzalesi N, Pirasat A. M-cone spon-

16. Castelo-Branco M, Faria P, Forjaz V, Kozak LR, Azevedo H. Simultaneous com-

17. Marmorstein AD, Marmorstein LY, Rayborn M, Wang X, Hollyfield JG, Pe-


©2005 American Medical Association. All rights reserved.