

Macular Pigment Measurement by Heterochromatic Flicker Photometry in Older Subjects: The Carotenoids and Age-Related Eye Disease Study

D. Max Snodderly,¹ Julie A. Mares,² Billy R. Wooten,³ Lisa Oxton,² Michael Gruber,² and Tara Ficek,² for the CAREDS Macular Pigment Study Group⁴

PURPOSE. To develop a standardized protocol for measuring macular pigment optical density (MPOD) of experimentally naïve subjects by heterochromatic flicker photometry (HFP).

METHODS. MPOD in eyes of 54 women, age 50 and 79 years (mean, 66), was studied. The spatial profile of MPOD was measured in the right eye, and two spatial points were also measured in the left eye. Forty-eight of these inexperienced subjects completed the protocol on two separate visits. For a subset of the group, the MPOD at two different wavelengths was measured.

RESULTS. The test-retest correlation at 0.5° eccentricity in the right eye was 0.9. On the second visit, more than 90% of the subjects were able to perform the HFP test with results that were consistent with the absorption spectrum of macular pigment. On the first visit, data from the inexperienced subjects deviated more from the expected relationships between the two wavelengths, presumably because they had less skill in performing the task. However, subjects with high or low macular pigment density were distinguished clearly.

CONCLUSIONS. Reliable and meaningful measurements of macular pigment density in older subjects can be made using HFP, with a standardized protocol in the limited time available in large epidemiologic studies. This protocol will be made freely available to other researchers on request. (*Invest Ophthalmol Vis Sci.* 2004;45:531-538) DOI:10.1167/iovs.03-0762

The macular pigment (MP) offers the opportunity to measure noninvasively the total density of xanthophylls in the foveal region. Recent research indicates that the macular xanthophylls lutein and zeaxanthin, which are derived from the diet, may lower risk for age-related macular degeneration and nuclear cataract. However the evidence, at this early stage of investigation, is inconsistent.¹ These issues are currently being studied as part of the Carotenoids in Age-Related Eye Disease Study (CAREDS), an ancillary study of the Women's Health

Initiative. Establishing relationships between MP levels and the onset and progression of eye disease will provide estimates of the degree to which MP may lower the risk for these conditions.

The technique for measuring MP that is most widely used is the psychophysical method of heterochromatic flicker photometry (HFP). Already, more than 2000 people in separate study samples of 300 or fewer subjects have had MP measured by HFP (e.g., Refs. 2-4, and ongoing studies). However, each of these studies was conducted at a single site or by a single examiner, and elements of the protocol may have varied from one site to another. Further insights into the determinants of MP optical density (MPOD) and its relationship to age-related chronic disease would benefit from studies of larger and more diverse samples of people. Use of HFP in large studies with multiple sites and/or examiners requires specification of protocols that result in valid and reliable data. Moreover, if MP is found to protect against age-related eye disease, the protocols must facilitate routine measurements in older subjects.

The widespread use of HFP has been made possible by the fact that the instrumentation needed is relatively inexpensive, and research staff can be trained to operate it in a short period. HFP also does not require pupillary dilation, and it is relatively unaffected by the changes in the ocular media that accompany aging. The method has been validated by demonstrating that measurement of MP density by HFP at different wavelengths reproduces the absorption spectrum of MP measured by various other methods.⁵ Results from HFP also correlate well with measurements made by reflectometry and autofluorescence.⁶

Potentially offsetting these advantages is the limitation that HFP requires the participation of the subject for a significant portion of the time devoted to a typical clinic visit or research study appointment. The subject must adjust the radiance of a light source and make reliable judgments about the disappearance of flicker when the adjustments are made. Experience to date has shown that most subjects, including elderly ones with moderate cataracts,⁷ can perform this task. However, some subjects cannot perform the task, and others make settings that are variable or are inconsistent with known properties of MP. Informal observations suggest that the fraction of subjects who have difficulty with the task increases with age and with more advanced stages of ocular disease. Therefore, it is important to define the characteristics of subjects in whom HFP is a suitable measurement technique and to assess its validity and reliability.

In this article we report on the reliability and practicality of using HFP to measure MP in older women. As part of our investigation of the relationship of MP to the risk of age-related eye disease (the CAREDS study), we have developed a standardized protocol to facilitate measurement of MP at multiple study sites with large populations. This protocol was developed to meet the following criteria:

1. High test-retest reliability.
2. Mean time for completion of 45 minutes or less.
3. Ease of performance for a wide range of subjects.

From the ¹Schepens Eye Research Institute, Boston, Massachusetts; the ²Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, Wisconsin; and the ³Department of Psychology, Brown University, Providence, Rhode Island.

⁴CAREDS Macular Pigment Study Group members are listed in the Appendix.

Supported by National Eye Institute Grant EY13018 and Research to Prevent Blindness.

Submitted for publication July 18, 2003; accepted August 16, 2003.

Disclosure: **D.M. Snodderly** (P); **J.A. Mares**, None; **B.R. Wooten**, Macular Metrics Corporation (I, P, R); **L. Oxton**, None; **M. Gruber**, None; **T. Ficek**, None

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: D. Max Snodderly, Department of Ophthalmology, Medical College of Georgia, Augusta, GA 30912-3402; msnodderly@mcg.edu.

4. Ease of training examiners to perform the day-to-day testing of large numbers of subjects.
5. A measure of the spatial profile of MP in one eye.
6. An internal check for consistency of the data.
7. A limited comparison of MP density in the two eyes.

Because this is the first standardized protocol for MP measurement by HFP, we included more features than may be needed for a minimal assessment. In the Discussion section, we consider the compromises that might be made to shorten the test if time is a critical consideration.

METHODS

Subjects

Volunteer subjects were recruited from the Ophthalmology Clinic, the Institute on Aging program, and the campus area of the University of Wisconsin. All procedures conformed to the Declaration of Helsinki and were approved by the Institutional Review Board of the University. Fifty-four women aged 50 to 79 years (mean, 66 years) were tested. Data from the following subjects ($n = 6$) were excluded from the analyses of the MP spatial profile and comparisons between the two eyes: One subject (subject 8) was unable to comprehend and execute the task. Another (subject 25) did not make a repeat visit. One subject (subject 52) had a vitreous floater in the right eye that interfered with the test. Three subjects (38, 53, and 54) completed all tests on the right eye, but due to time constraints, could not complete testing of the left eye. In the core group of 48 subjects who completed the protocol on both visits, 15 had one or more self-reported eye diseases, including cataract ($n = 10$), glaucoma ($n = 3$), and age-related macular degeneration ($n = 2$). They were relatively well educated; all but 7 had some post-high-school education, and 34 had graduated from college or technical school.

Principles of the Method

Measurements were made by the psychophysical method of HFP. The subject views a small test field superimposed on a blue background. The test field alternates between a wavelength (blue or blue-green) that is absorbed by the MP and a reference (green to yellow-green) wavelength that is outside the absorption band of MP. When the frequency of alternation is chosen correctly, the test field appears to flicker. The basic assumption underlying the method is that the spectral sensitivities of the visual mechanisms detecting the stimulus in the fovea and in the parafovea are the same (for a summary, see Ref. 8). The blue background and the use of flicker are designed to exclude participation of the rods and the S-cones, so that detection is mediated only by the L and M cones. These cones are thought to be present in relatively constant proportions in the fovea and the parafovea, consistent with the assumption of uniform spectral sensitivity. Nevertheless, differences in visual pigment density of cones in the fovea and parafovea can introduce minor differences in the spectral sensitivity at the two locations. However, effects on the measurement of MP are small (0.05 optical density at 460 nm at 0.5° eccentricity; Smollon WE, et al. *IOVS* 2002;43:ARVO E-Abstract 2952) and for present purposes can be ignored.

When making settings, the subject is instructed to adjust the energy of the bluish test light so that the flicker stops. The amount of bluish light that is required to produce this flicker null provides a measure of the optical density of MP (MPOD) at the retinal location of the test light. For the case of a blue test light at the peak wavelength of the MP spectrum and a reference wavelength outside the absorption band, $MPOD = -\log_{10} (R_f/R_p)$, where R_f is the radiance of the blue light needed for a flicker null at the foveal location being measured, and R_p is the radiance for a flicker null at the reference location in the parafovea, where MPOD is negligible.⁹

Testing Setup

Stimuli were generated by a slightly modified version of the tabletop macular densitometer previously described.⁹ This instrument utilizes light-emitting diodes (LEDs) and electronic controls. For the present study, the following features were added to the device: (1) The aperture and the diffuser forming the test field were mounted on a motorized filter wheel so that the size and shape of the test field could be changed easily by the examiner. (2) Another LED was added to the test source to make available an auxiliary blue-green test stimulus in addition to the blue stimulus near the spectral peak of MP. (3) A small red LED was added to serve as a fixation point to place the test field at the reference locus in the parafoveal retina. (4) An ophthalmic lens was placed in front of the subject's test eye. This lens was designed to focus an emmetropic subject at the plane of the test aperture. Another lens was sometimes used to match the refractive correction of the subject (described later).

Four test field configurations—two solid disks and two annuli—were used to derive a spatial profile of MPOD in the right eye. With solid disks, it has been shown that the measure of MPOD is obtained at the edge of the disk.^{2,5,10} Thus, the subject was instructed to fixate at the center of the following targets: (1) A solid disk of 15 minarc radius; (2) a solid disk of 0.5° arc radius; (3) an annulus with inner radius 50 minarc, outer radius 70 minarc (to derive a measure at 1° eccentricity); and (4) an annulus with inner radius 90 minarc, outer radius 120 minarc (to derive a measure at 1.75° eccentricity). A small black dot was present at the center of the solid disks as a fixation aid. A fixation target of 5 minarc radius was also centered within each annulus.

For the reference measurement, the subject viewed a test target of 1° radius centered at 7° temporal eccentricity while looking at the red fixation LED. A spatial profile was not derived in the left eye; only two test targets were used (at 0.5° and 1° radius), and the reference was the 1° target at 7° nasal eccentricity. In all cases, the test target was superimposed on the center of a blue background field of 3° radius.

The peak wavelength, full bandwidth at half maximum, and luminance of the stimuli were respectively as follows: background, 468 nm, 30.6 nm, 2.6 cd/m²; yellow-green test, 564 nm, 32.5 nm, 1.9 cd/m²; blue test, 460 nm, 22.5 nm bandwidth, luminance adjusted by subject; blue-green test, 488 nm, 33 nm bandwidth, luminance adjusted by subject. Note that these luminance values are for each wavelength band in isolation. In the experiment, the test stimuli were superimposed on the background, so that the luminance of the test target was added to the luminance of the background.

In this article, MPOD calculations are based on the MP spectrum for fresh macaque retina measured by the late Paul Brown.¹¹ The absorption of MP for each test stimulus was computed by integrating over the wavelength band of the relevant LED. MPOD is referenced to the peak wavelength, 460 nm for the blue test, and to 500 nm for the blue-green auxiliary test.

Sequence of Events

Before testing, the subject's current corrected visual acuity was obtained. If the subject's corrected visual acuity was 20/40 or better and her personal correction lenses were neither tinted nor progressive, she wore her own lenses for testing. If the subject's visual acuity was less than 20/40, a pinhole acuity check was performed. If the subject's pinhole acuity was two lines or more better than her current corrected acuity, a complete refraction protocol was performed. Subjects who underwent a complete refraction were then fitted with trial frames and appropriate lenses for testing. Similarly, subjects who arrived with tinted or progressive lenses were tested with clear trial lenses. The use of progressive lenses was avoided to prevent defocus of the more eccentric stimuli.

The examiner familiarized the subject with the apparatus and showed the subject an instructional video that explained the task. The video was designed to provide a uniform introduction to the task for all subjects. Copies of this video may be obtained from the authors; a

downloadable version will be made accessible on the Internet on receipt of an email request.

Testing was begun with the right eye. First, the subject's critical flicker frequency was measured with the 0.5° target in the fovea (CFF_{fov}) and the 1° target in the parafovea (CFF_{pfov}). For these measurements, the blue test LED was turned off, and the subject saw only the green test LED flickering on the blue background. The flicker frequency was initially set to 7 Hz, and the subject was instructed to increase the frequency by turning a knob until the flicker stopped. For these stimulus conditions, CFF values were 17.6 ± 3.2 in the fovea and 14.9 ± 3.4 in the parafovea. This task served two purposes. It taught the subjects to recognize the presence and absence of flicker, and it provided a starting point, as first proposed by Rosen (Rosen RB, personal communication, 2002) for choosing the optimal flicker frequency for the HFP.

An algorithm was established to help the examiner determine the best flicker frequency for HFP. The algorithm was implemented with the aid of a worksheet provided for the examiner. For the foveal targets, a practice test was conducted with the 0.5° target, with a flicker rate 7 Hz lower than the CFF_{fov} . If the subject was able to adjust the radiance of the blue light to find a null zone of no flicker or to clearly minimize the flicker, this frequency was used for all the foveal targets. If the subject could not minimize the flicker or find a null zone, the flicker frequency was increased until a flicker minimum could be identified. Conversely, if the subject reported a large null zone so that a minimum was not precisely defined, the flicker frequency was decreased until a flicker minimum could be identified. A lower bound of 8 Hz was established to avoid violating the assumptions about the underlying visual mechanisms described earlier. The flicker frequency finally established in this way was used for all the foveal targets.

If the subject performed well on the foveal practice test with the predicted best frequency of $CFF_{fov}-7$, the parafoveal HFP measurements were begun with a flicker rate 9 Hz lower than CFF_{pfov} . If the flicker frequency established for the foveal targets differed from the predicted value, the parafoveal test frequency was adjusted accordingly. For example, if the best foveal frequency was $CFF_{fov}-8$, instead of $CFF_{fov}-7$ then the starting value for the parafoveal practice test was set at $CFF_{pfov}-10$ Hz rather than $CFF_{pfov}-9$. If the subject still could not minimize the flicker or find a null zone, the flicker frequency was increased until a flicker minimum could be identified. Conversely, if the subject reported a large null zone so that a minimum was not precisely defined, the flicker frequency was decreased until a clear minimum could be identified. A lower bound of 5 Hz was mandated (the parafoveal bound was lower than the foveal ones). Flicker frequencies actually used for HFP were 11.5 ± 2.5 in the fovea and 7.3 ± 2.6 in the parafovea. Note that only two different flicker frequencies were used, one for all foveal targets, and one for the parafoveal reference target.

The exact formula for selecting the frequency for HFP based on the CFF values may be adjusted slightly in the future. For example, our preliminary results from more than a thousand subjects tested subsequently suggest that the initial frequency for HFP with foveal stimuli could be set at 6 Hz rather than 7 Hz below the CFF (Snodderly DM, unpublished data, 2003). Similarly, the initial frequency for the parafoveal reference stimulus could probably be set 0.5 to 1 Hz higher. Because the optimal flicker frequency is a function of many stimulus parameters, different starting points will probably be needed for studies using other stimulus conditions.

The subject made five separate determinations for each target. To maintain uniformity during testing, the examiner followed a protocol narrative of instructions to the subject for each target. Throughout testing, subjects were reminded of the task and told where to maintain the line of sight.

Measurements were made first in the right eye with the blue test stimulus with all targets to estimate the spatial profile of MPOD at 15 minutes and 0.5°, 1°, and 2° eccentricity. Then measurements were made with the blue-green auxiliary stimulus with the 0.5° target in the fovea and with the parafoveal reference target. Finally, the subject was

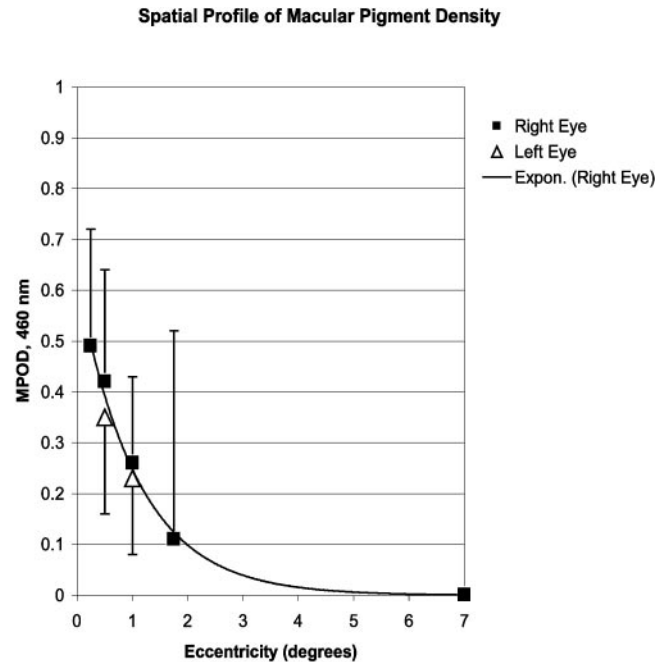


FIGURE 1. Mean MPOD measurements at 460 nm as a function of eccentricity. Data are from measurements made during two visits of 48 women using the standard HFP protocol. Error bars, SD. An exponential function has been fitted to the data for the right eye ($y = 0.63e^{-0.92x}$).

repositioned to test the left eye with the blue test stimulus for the foveal targets of 0.5° and 1° radius and the parafoveal reference target. Fewer targets were tested in the left eye to limit the total time required of the subject.

The entire set of measurements was repeated about a week later (5.8 ± 4.4 days) to assess test-retest reliability.

Training of the Examiner

Before testing study subjects, the examiner-in-training completed an instructional sequence. The trainee first served as a subject to have MP measured by the standard protocol and received instruction in the calibration and use of the densitometer. Then the trainee measured the MP of a certified examiner or one of the authors. After the training session, the trainee was required to test and retest at least three subjects and submit the data to the Coordinating Center. If the test-retest reliability and the variance of the measurements were within the normal range, the trainee was certified as an examiner to perform testing on study subjects.

Examiners followed a protocol narrative of instructions to the subject and recorded data on standard forms. Copies of the narrative and the worksheets may be obtained from the authors; downloadable versions will be made accessible on the Internet on receipt of an email request.

RESULTS

Spatial Profile of MP Density

The means and SDs (error bars) of data from 48 subjects who completed the protocol on two separate visits are displayed in Figure 1. The mean spatial profile for the right eye is approximately exponential in form and is similar to profiles measured previously.¹⁰ Data for two spatial points in the left eye are also plotted. Our subject group had a mean MPOD at 0.5° in the right eye of 0.42 ± 0.22 , which is somewhat higher than mean

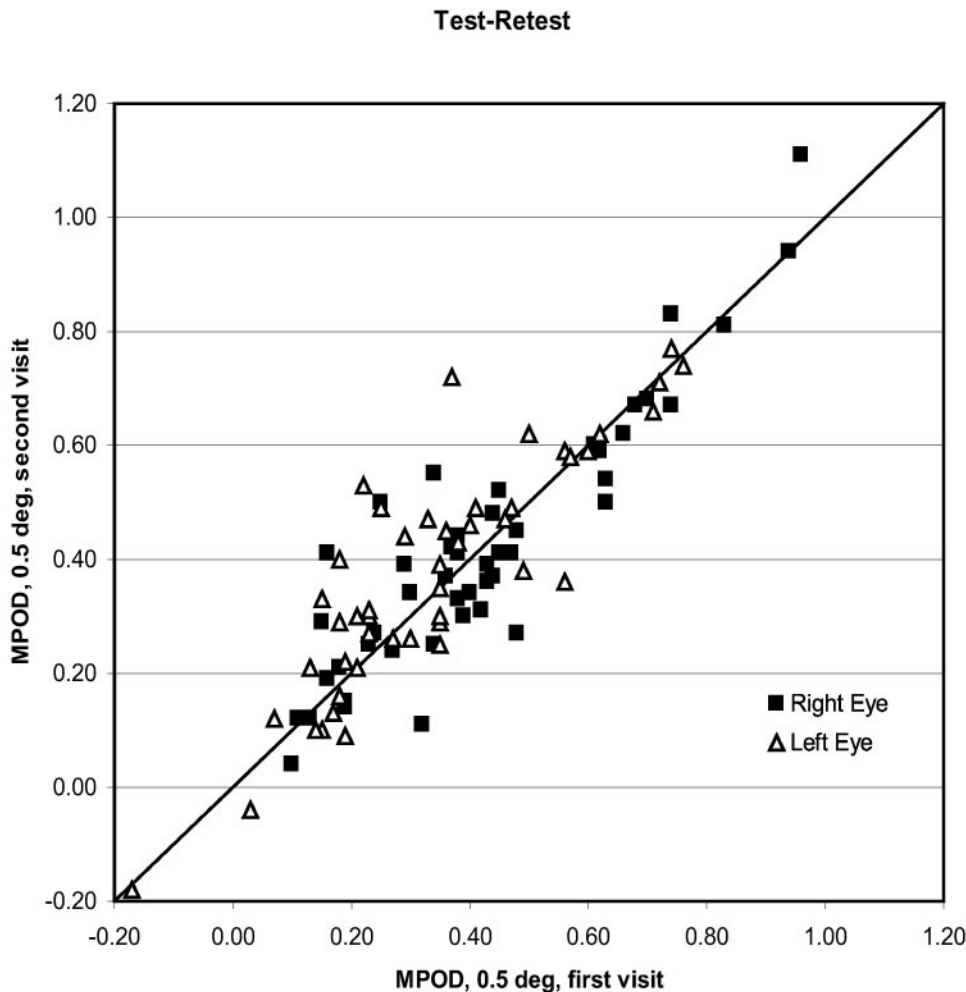


FIGURE 2. MPOD at 460 nm measured at 0.5° eccentricity on the first and the second visit in 48 women. The line of equality is drawn for reference.

ODs measured by HFP at other study sites in the United States and Europe (0.29,¹² 0.26,² and 0.25.³ Note that the MPODs in the latter two studies have been corrected for effects of bandwidth of the stimuli, which were not considered in the original publications).

The measurement made at 0.5° eccentricity in the right eye had the lowest ratio of within- to between-subject variation (0.09 at the first visit). Within- to between-person variations for the 0.25°, 1.00° and 1.75° targets were 0.25, 0.20, and 0.25, respectively on the first visit. A measure of area under the MP profile between 0° and 1.75° correlated highly with the density at 0.5° ($r = 0.98$) and therefore was not investigated further.

Test-Retest Reliability

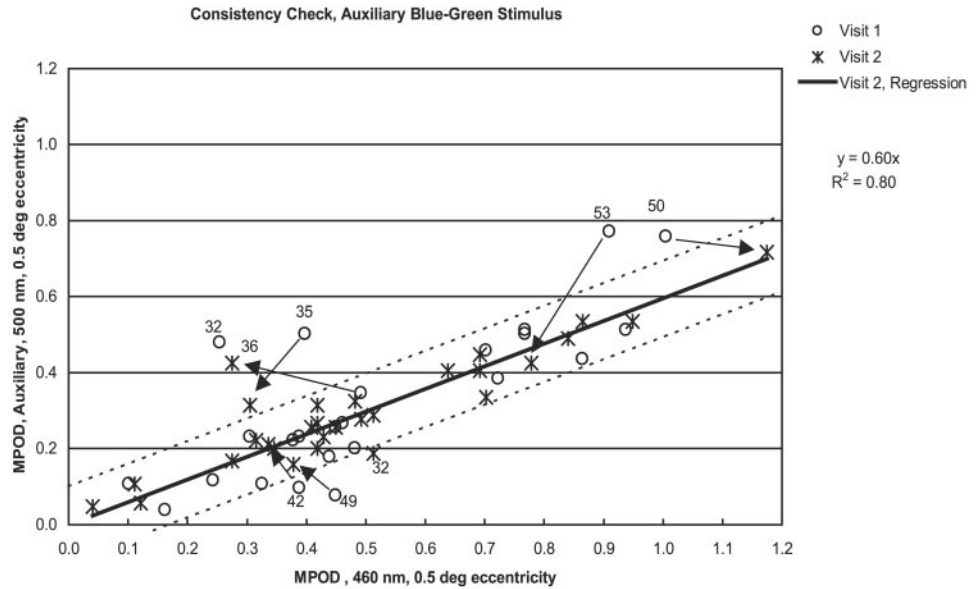
When subjects were retested using the same protocol about a week later, the mean MPODs at 0.5° eccentricity were identical (MPOD = 0.42; Fig. 2). The test-retest correlations (Pearson r) for measurements in the right eye were 0.86, 0.90, 0.86, and 0.68 for measurements at 0.25°, 0.50°, 1.00°, and 1.75° from the foveal center, respectively. In the left eye, mean MPOD at 0.5° eccentricity was slightly lower on the first visit (0.33) than on the second (0.37, $P = 0.02$ by matched-pairs t -test) but the test-retest correlation was still high (0.86 and 0.75 at 0.5° and 1.0°, respectively). These test-retest results of older women are similar to earlier results attained by a skilled examiner with young subjects of college age ($r = 0.73$ – 0.94).^{13,14} Although it was an implausible outcome, one subject had negative values for MPOD in the left eye that were

precisely reproduced on the second visit. This observation underscores the point that repeatability of the results does not address the question of their validity, an issue that we consider later.

Right Eye–Left Eye Comparison

MPOD at 0.5° eccentricity correlated highly in the two eyes on both the first test ($r = 0.79$) and the retest ($r = 0.80$). However, the mean OD in the left eye was consistently lower ($P < 0.001$ for visit 1 and $P < 0.05$ for visit 2 by matched-pairs t -test). This difference in MPOD between the two eyes was unexpected, because the fellow eyes of young subjects are usually well matched.¹⁴ Although some very experienced subjects have reliable differences between fellow eyes,¹⁵ we are not aware of any population studies that have shown MPOD to be systematically higher in the right eye. One possibility is that age-related changes may have reduced the interocular symmetry in our older sample of subjects. Women who had 30% or greater differences between right and left eyes for both the test and retest ($n = 11$) were older (71.0 vs. 63.5 years, $P = 0.02$) and were more likely to have visual acuity of 20/40 or worse in the left eye (45% vs. 11%, $P = 0.02$). It is also possible that the data were biased by having measured the right eye first; we have not investigated this potential order effect. However, the asymmetry was small (0.05 OD on the second visit) and it is likely to be a sampling bias (see the Discussion section).

FIGURE 3. MPOD measured at 0.5° eccentricity with an auxiliary blue-green stimulus referred to 500 nm compared with MPOD measured with the blue test stimulus referred to 460 nm. The least-squares regression line through the data points for the second visit (*solid line*) was forced to an intercept of 0 and has a slope of 0.60. The predicted value based on the absorbance spectrum of MP is 0.56. *Dashed lines*: ± 0.1 OD from the regression line. The ID numbers of subjects with data most divergent from the regression line are shown. *Arrows* connect the values from the first and the second visits of these subjects. *Arrow* for subject 32 was omitted to avoid clutter.



Validity of the Measurements

Because HFP requires the understanding and participation of the subjects, it is important to be confident that the subject performs the task correctly. As a check for internal consistency of the data, we repeated the measurements at 0.5° eccentricity in the right eye with an auxiliary blue-green wavelength band for 24 subjects on both the first and the second visits, with 5 additional subjects tested only on the second visit. If the subjects performed the task correctly, measurements with the auxiliary blue-green test should give MPODs (referred to 500 nm) proportional to the MPOD measured with the standard blue test (referred to 460 nm). Note that in both cases, corrections are made for the bandwidth of the stimuli. Consistent with the expected proportionality, Figure 3 shows that the data can be fit by a least-squares regression line with a slope of 0.60 (intercept set to 0). For this figure, we calculated the regression for data from the second visit of the 29 subjects, which had the higher correlation coefficient. The slope of the regression line is very close to the value of 0.56 that we calculated for these two test stimuli, using the MP absorbance data of Paul Brown, assuming a mean lens density spectrum¹⁶ appropriate for this age group (mean, 66 years). (Lens density had a slight effect on the blue test-auxiliary test ratio because we were not using monochromatic sources. See the Methods section.)

Measurements that fall more than a criterion value from the regression line are potentially unreliable and experimenters could target them for possible exclusion in specific analyses. In Figure 3, we have drawn dashed lines at ± 0.1 OD from the regression line to help identify subjects whose data deviate from the expected relationship. Data from individual subjects that fell outside this range are labeled with the subjects' ID numbers. Arrows connect the values from the first and the second visits, except for subject 32, whose arrow was omitted to avoid confusion.

The results suggest that there is a practice effect. Data from all the outlying cases but one moved closer to the regression line on the second visit. This improvement is not an artifact of choosing the second visit to calculate the regression. If all data are included, the slope of the regression line is only slightly different, (0.61 instead of 0.60), and the practice effect remains clear. A detailed analysis of the effects of practice will be presented in a subsequent publication (Leung IY-F et al., personal communication, 2003).

As part of the protocol, the examiner was encouraged to record observations of difficulties or unusual occurrences during testing. If the examiner did not comment, we assumed the testing was routine. We summarize in Table 1 the comments of the examiner regarding the performance of the six subjects whose data from the first visit could be considered outliers from the main relationship. All subjects whose data deviated most from the usual relationship were noted by the examiner to have difficulty in executing the task. Thus, the impressions of the examiner that the test was unusually difficult for specific subjects were experimentally confirmed by the use of the auxiliary wavelength band.

When deciding how to analyze the data, there are several considerations. Data from the second visit are consistent with a valid measurement of MP, including typical variability in the settings for at least 27 ($\geq 93\%$) of 29 subjects. The exceptions occurred in subject 36, and if very stringent selection criteria are applied, perhaps in subject 35. On the first visit, more subjects deviated farther from the expected relationship, presumably because of less skillful execution of the task. However, as shown in Table 1, the difference in MPOD between the two visits was still generally small compared with the differences between subjects. Thus, one can expect more variability in the data from the first visit, but in the current study it appears that the errors that were introduced still allowed ordering of subjects (except for subject 36) according to their MPODs. We return to issues of data selection and analysis in the Discussion section.

Control for Experimenter Bias

Because the examiner interacts extensively with the subject, one must always be careful to minimize effects of experimenter bias. However, a mistake made by the examiner enabled us to confirm that experimenter bias had little effect on the results. We had intended to test all subjects with the auxiliary blue-green stimulus, which required changing a switch on the densitometer when changing from the usual blue stimulus. For the first 31 subjects, the examiner failed to change this switch when she conducted the presumed auxiliary test. As a result, the examiner was expecting different results on the second test even though the two tests were identical. Figure 4 shows that the examiners' bias (or expectation) had no discernible effect on the outcome. The slope of the regression line through the data points is 0.96, indistin-

TABLE 1. Examiner's Comments about Subjects with Outlying Data

Subject	Visit	Examiner's Comments	Δ MPOD
32	1	Found test very difficult. May not have done test correctly.	0.26
	2	No comment.	
35	1	Highly variable when testing.	-0.09
	2	Highly variable when testing.	
36	1	Had difficulty following directions of test.	-0.22
	2	Had difficulty following directions of test.	
42	1	No comment.	-0.05
	2	Seemed to have clearer understanding of the task.	
49	1	No comment.	-0.07
	2	No comment.	
50	1	Fatigued near end of session.	0.17
	2	No comment.	
53	1	Understood test, but too slow, fatigued at end.	-0.13
	2	No comment.	

The subject ID numbers are associated with individual data points in Figure 3. The difference between the MPOD at 460 nm measured in the right eye at 0.5° eccentricity on the first and the second visit is listed for each subject.

guishable from the line of equality, but very different from the examiner's expectations.

Time Constraints

For 18 of the subjects, the examiner kept detailed records of the time necessary to complete the testing. The median time was 45 minutes, with a minimum time of 29 minutes and a maximum of 60 minutes. A significant part of the time was

devoted to explaining the task to the subjects and to allowing them to practice making settings and recognizing the flicker null zone. Realistically, one should plan that even the minimum measurement—say the 0.5° point in one eye of a naïve elderly subject—would require 20 to 30 minutes on the first visit. Subsequent visits might require less time if they occurred within a relatively short interval of a few weeks. Three of our subjects who were unable to complete the entire protocol on the first visit were able to do so on the second visit because they could perform the task more quickly.

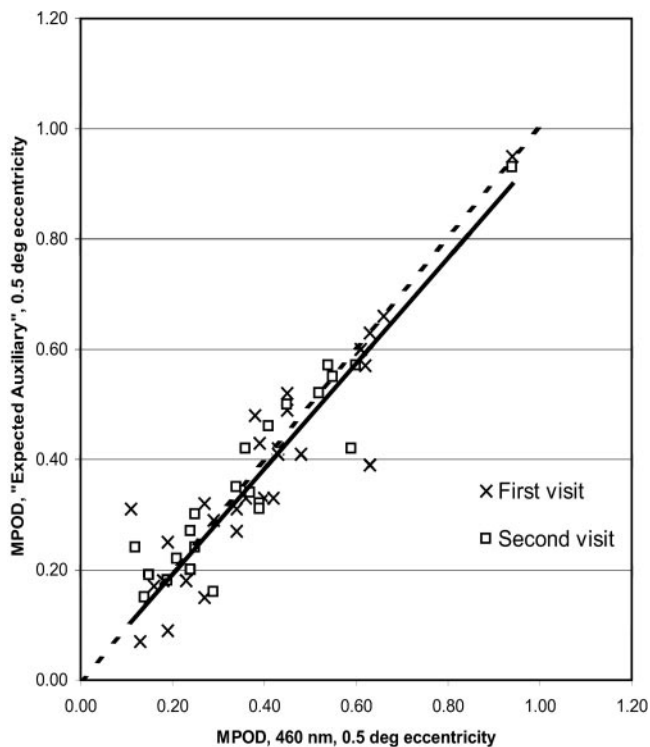


FIGURE 4. Within-session repeats of MPOD measurements. The experimenter intended to use the auxiliary test stimulus as in Figure 3, but because of an error, repeated the blue test in the same experimental session. MPOD from the first test is plotted on the horizontal axis, and MPOD from the second, "expected auxiliary" test is plotted on the vertical axis. The least-squares regression line (solid line) with intercept forced to 0 has a slope of 0.96, indistinguishable from the line of equality (dashed line). Pearson $r = 0.92$. These results show that the expectations of the examiner had no discernible impact on the data.

DISCUSSION

The Spatial Profile of MP

Although the mean data for the MP profile are well fit by an exponential function, individual subjects had spatial profiles that differed from this pattern in the present and previous studies.¹⁰ These differences in spatial profiles are the subject of active investigation (Delori et al., manuscript in preparation). One motivation for obtaining the spatial profile was to be able to estimate total MP¹⁷ rather than just the value at a single point. In our main study (CAREDS), we will test whether the spatial profile provides more information than the MPOD at a single point about relationships between xanthophylls in the retina and xanthophylls in the diet or blood, as well as other potential determinants of MP density.

Reliability

The results of the present study show that HFP produces reliable measurements of MPOD in inexperienced older women when a carefully structured, standardized protocol is used. Using the measurement that best discriminates between MPODs across people (made at 0.5° from the foveal center), the test-retest reliability was high (>0.9 .) This consistency attests to the robustness of the protocol for use in a general population including older subjects. The reliability of the data also indicate that HFP can be used without the necessity for repeat visits, which makes it suitable for studies with large numbers of subjects.

Delori et al.⁶ have compared the repeatability of different in vivo methods for measuring MPOD by using the ratio of the mean absolute difference between repeat measures to the mean MPOD of all measurements, expressed as a percentage. For our 48 subjects, this value for between-session measurements is 17% in the right eye and 22% in the left eye. These

reproducibility figures compare favorably with the same index for reflection densitometry (19%–22%) but the variability is higher than MPOD estimates obtained by autofluorescence with specialized instrumentation (9%–11%).

Another index that has been used is the coefficient of repeatability,¹⁸ which is 1.96 times the standard deviation of the differences between repeat measurements. For the right eye, the coefficient of variability of our between-session MPOD measurements was 0.19, and for the left eye, 0.21. This compares favorably with results of Berendschot et al.¹⁸ derived from within-session reflectance measurements with a scanning laser ophthalmoscope (0.17) and a specialized spectral reflectance analyzer (0.27). Furthermore, one should expect that these reflectance methods will produce more variable results when between-session measurements are made, which is the more important case.

It is difficult to compare reliability of HFP with the Raman method for measuring MP,^{19,20} because data have only been reported on the within-session repeatability of two subjects, choosing the best three of five measurements. The criterion for choice of the three accepted measurements was not defined. No data on between-session repeatability have been presented.

Validity of HFP Measurements

The repeatability of MP measurements and the validity of the measurements are two different issues. The more difficult criterion for MP measurements to meet is proof of validity. In our case, we have shown that the data are consistent with the known spectrum of MP, which is an important requirement. However, the utility of this criterion is limited by the need for a measurable density of pigment and by the variability of subjects' responses. When the subject has a low MPOD, the measurement of density at wavelengths other than the peak have a large uncertainty because of the inherent variability in the measurements. This limitation is shared by all methods. The performance of the subject is, however, central to HFP measurements to a degree not essential for most physical methods. For large-scale epidemiologic studies, a single measure of MPOD is a feasible design, and the performance of inexperienced subjects will be a significant source of variability of the data, but there is no reason to expect bias. In longitudinal studies, in which data for individual subjects are gathered repeatedly, there is a clear advantage to obtaining a second measurement before beginning any intervention, so that the practice effect can be exploited.

Ultimately, the conditions for optimizing validity must be proven by a consensus among methods, and at present there is no standard paradigm to follow. We have shown that there is a high correlation between physical methods of measuring MP and the results of HFP.⁶ There is little reason to doubt that HFP can rank subjects in a meaningful order according to their MPODs. Nevertheless, issues of scaling will remain unresolved until a better understanding of the factors affecting all the methods is reached. In the meantime, researchers must realize that the values obtained by different *in vivo* methods, even if highly correlated, probably do not represent the same absolute quantities of MP.

Interocular Data

We observed slightly higher mean values of MPOD in the right eye. However, the interocular asymmetry was small (0.05 OD at 0.5° eccentricity on the second visit) and it is likely to be a sampling bias. Our preliminary results from several hundred subjects tested subsequently do not show a significant right-left asymmetry (Mares JA, unpublished data, 2003).

For comparison with other personal characteristics, such as diet or ethnic background, there may be a slight advantage in

taking the mean of the MPOD values of the two eyes. The interocular mean value has a slightly lower ratio of within- to between-subject coefficient of variation than the value in the right eye alone (23% vs. 24%). However, the small gain in ability to distinguish between levels of MP across individuals may not justify the additional time to complete tests in both eyes.

Time Constraints

The median time for completing our protocol was 45 minutes. A significant part of the time was devoted to explaining the task to the subject and to allowing her to practice making settings and recognizing the flicker null zone. This part of the protocol may have contributed to the good overall reliability and ultimately may have reduced the time of the test by making the task easier for the subject. To retain this segment, one should plan for 20 to 30 minutes on the first visit of inexperienced older subjects for the measurement of MPOD at even a single retinal locus. For younger subjects, or for practiced subjects, the testing time for this minimum measurement might be shortened to as little as 15 minutes. In our case, we obtained data on the spatial profile and interocular differences that required longer testing times. The utility of these additional data for identifying determinants of MP density will be evaluated in the CAREDS study.

Factors Influencing the Success of the Subject

The ease of performing the HFP task is strongly influenced by the frequency with which the stimulus alternates between the test and the reference wavelengths. Our protocol is designed to give the subjects practice in recognizing when there is a flicker null and to optimize the conditions for the subject. Although the flicker frequency can be fine-tuned to obtain a very narrow null zone, it is not necessary to spend a great deal of effort on this adjustment. The width of the null zone does not appear to be critical, as long as the subject can comfortably identify the center of it.

Some abnormalities of the ocular media can prevent subjects from being able to perform the task. One subject had a vitreous floater in the right eye that interfered with the test, but she was able to complete the main components of the protocol using the left eye. Consistent with previous experience,⁷ self-reported cataracts did not appear to cause serious difficulty.

Finally, the instructions to the subject must be clear, simple, and repeated often. For this reason, training of the examiner in techniques to explain procedures and to motivate the subject is likely to be important in obtaining reliable measurements. Subjects in our study received standardized instructions that were prepared on the basis of extensive preliminary experimentation. For example, as trivial as it may seem, it is important to remind subjects to blink frequently. This helps to avoid fading of stimuli during prolonged fixation, as well as to prevent drying and irritation of the cornea.

Comparison of HFP with Physical Methods for MPOD Measurements

The *in vivo* methods for measuring MPOD have different strengths and weaknesses. The three physical methods (lipofuscin fluorescence, reflection densitometry, and Raman spectroscopy) do not require the subject to make a judgment or to respond to a stimulus. However, they require unpleasant light levels, pupillary dilation, and precise alignment. In some older subjects, the pupil will not dilate to the diameters required by these techniques. The current implementation of Raman spectroscopy also requires the subject to self-align and hold alignment without the aid of a restraint device. For some subjects, this may be a difficult task. In addition, results from reflecto-

metry and Raman spectroscopy are potentially confounded by degradation of the ocular media accompanying aging and disease. Among the methods discussed herein, the Raman technique is particularly vulnerable to changes in the ocular media,²¹ because it does not use a parafoveal reference locus.¹⁹

In comparison, HFP as implemented in our system,⁹ does not require pupillary dilation, is relatively insensitive to changes in the ocular media, and subject alignment is not critical. A bite bar is not required. Any lateral position within ± 3 cm gives the same MPODs. Furthermore, the instrumentation is relatively inexpensive and easy to operate with a modest amount of training of the examiner.

Although it is more obvious that significant time is required for MPOD measurements by HFP, the time needed for physical methods can also be appreciable. Pupillary dilation, preparation of a bite bar, and subject alignment may require time during the subject's visit, and recovery time between repeat exposures is necessary for the Raman method. The time required from the subject should be evaluated in the context of all the procedures that are performed.

Because of the requirements for subject participation, there are specific groups of people for whom HFP is not a suitable method. These include people with severe visual impairment and those who are too physically or mentally infirm to respond to the testing. However, our results, and preliminary studies under way suggest that most inexperienced subjects, including elderly ones, can execute the task and give data consistent with a valid measure of MPOD. Nevertheless, we note that our subjects were relatively well educated, and the time necessary for completing the task as well as the reliability may differ for groups with less education. As more studies are conducted, it should be possible to identify subject groups that are appropriate for HFP measurements and those that are not. Comparable criteria for identification of appropriate subjects, such as maximum pupillary aperture and the state of the ocular media, should be established in a corresponding manner for physical methods of measuring MPOD.

CONCLUSIONS

Using a standardized protocol, HFP provided reliable estimates of MPOD in older women consistent with the absorbance spectrum of MP. The largest amount of information about the MP density across subjects was provided by measurements made at 0.5° from the foveal center. Reduction of the protocol to this single measure could limit the time required from the subject to 20 to 30 minutes. Ongoing research using the full protocol will provide insights into the added value of making measurements in both eyes and describing the spatial distribution of MP.

Acknowledgments

The authors thank Richard Rosen and Joanne Curran-Celentano for helpful suggestions in developing the protocol and Randy Hammond for comments on the manuscript.

References

- Mares-Perlman JA, Millen AE, Ficek TL, Hankinson SE. The body of evidence to support a protective role for lutein and zeaxanthin in delaying chronic disease: overview. *J Nutr.* 2002;132:518S-524.
- Hammond BR Jr, Caruso-Avery M. Macular pigment optical density in a Southwestern sample. *Invest Ophthalmol Vis Sci.* 2000;41:1492-1498.
- Ciulla TA, Curran-Celentano J, Cooper DA, et al. Macular pigment optical density in a midwestern sample. *Ophthalmology.* 2001;108:730-737.
- Hammond BR, Ciulla TA, Snodderly DM. Macular pigment density is reduced in obese subjects. *Invest Ophthalmol Vis Sci.* 2002;43:47-50.
- Werner JS, Donnelly SK, Kliegl R. Aging and human macular pigment density. *Vision Res.* 1987;27:257-268.
- Delori FC, Goger DG, Hammond BR, Snodderly DM, Burns SA. Macular pigment density measured by autofluorescence spectrometry; comparison with reflectometry and heterochromatic flicker photometry. *J Opt Soc Am A Opt Image Sci Vis.* 2001;18:1212-1230.
- Ciulla TA, Hammond BR Jr, Yung CW, Pratt LM. Macular pigment optical density before and after cataract extraction. *Invest Ophthalmol Vis Sci.* 2001;42:1338-1342.
- Snodderly DM, Hammond BR. *In vivo* psychophysical assessment of nutritional and environmental influences on human ocular tissues: lens and macular pigment. In: Taylor A, ed. *Nutritional and Environmental Influences on Vision.* Boca Raton, FL: CRC Press; 1999:251-273.
- Wooten BR, Hammond BR, Land RI, Snodderly DM. A practical method for measuring macular pigment optical density. *Invest Ophthalmol Vis Sci.* 1999;40:2481-2489.
- Hammond BR, Wooten BR, Snodderly DM. Individual variations in the spatial profile of human macular pigment. *J Opt Soc Am A.* 1997;14:1187-1196.
- Snodderly DM, Brown PK, Delori FC, Auran JD. The macular pigment. I. Absorbance spectra, localization, and discrimination from other yellow pigments in primate retinas. *Invest Ophthalmol Vis Sci.* 1984;25:660-673.
- Beatty S, Murray IJ, Henson DB, Carden D, Koh H-H, Boulton ME. Macular pigment and risk for age-related macular degeneration in subjects from a Northern European population. *Invest Ophthalmol Vis Sci.* 2001;42:439-446.
- Hammond BR Jr, Fuld K, Curran-Celentano J. Macular pigment density in monozygotic twins. *Invest Ophthalmol Vis Sci.* 1995;36:2531-2541.
- Hammond BR Jr, Fuld K. Interocular differences in macular pigment density. *Invest Ophthalmol Vis Sci.* 1992;33:350-355.
- Landrum JT, Bone RA, Joa H, Kilburn MD, Moore LL, Sprague KE. A one year study of the macular pigment: the effect of 140 days of a lutein supplement. *Exp Eye Res.* 1997;65:57-62.
- Pokorny J, Smith VC, Lutze M. Aging of the human lens. *Appl Opt.* 1987;26:1437-1440.
- Handelman GJ, Snodderly DM, Krinsky NI, Russett MD, Adler AJ. Biological control of primate macular pigment: biochemical and densitometric studies. *Invest Ophthalmol Vis Sci.* 1991;32:257-267.
- Berendschot TTJM, Goldbohm RA, Klopping WAA, van de Kraats J, van Norel J, Van Norren D. Influence of lutein supplementation on macular pigment, assessed with two objective techniques. *Invest Ophthalmol Vis Sci.* 2000;41:3322-3326.
- Gellermann W, Ermakov IV, Ermakova MR, McClane RW, Zhao DY, Bernstein PS. *In vivo* resonant Raman measurement of macular carotenoid pigments in the young and the aging human retina. *J Opt Soc Am A Opt Image Sci Vis.* 2002;19:1172-1186.
- Bernstein PS, Zhao DY, Wintch SW, Ermakov IV, McClane RW, Gellermann W. Resonance Raman measurement of macular carotenoids in normal subjects and in age-related macular degeneration patients. *Ophthalmology.* 2002;109:1780-1787.
- Wooten BR, Hammond BR. Assessment of the Raman method of measuring human macular pigment (E-Letter). *Invest Ophthalmol Vis Sci.* 2003.

APPENDIX

The CAREDS Macular Pigment Study Group

Barbara Blodi, MD; Alvin Eisner, PhD; Karen Gehrs, MD; Randy Kardon, MD, PhD; Michael L. Klein, MD; Julie A. Mares, PhD; D. Max Snodderly, PhD; James Ver Hoeve, PhD; and Billy R. Wooten, PhD.