Published in final edited form as:

Optom Vis Sci. 2011 December ; 88(12): 1496–1506. doi:10.1097/OPX.0b013e3182348d0b.

Vision Test Variability in Retinitis Pigmentosa and Psychosocial Factors

Ava K. Bittner, O.D., Ph.D., FAAO, Mohamed A. Ibrahim, M.D., Jennifer A. Haythornthwaite, Ph.D., Marie Diener-West, Ph.D., and Gislin Dagnelie, Ph.D., FAAO

Johns Hopkins University, Wilmer Eye Institute, 550 N. Broadway, 6th floor, Baltimore, MD 21205

Abstract

Purpose—We explored whether greater amounts of short-term variability in visual acuity (VA), contrast sensitivity (CS), or visual field (VF) in retinitis pigmentosa (RP) was related to disease severity or psychosocial factors.

Methods—We obtained spectral domain-optical coherence tomography in 27 RP subjects and determined variability (SD) of VA, CS and VF during a mean of 16 tests self-administered at home on a personal computer (PC) twice a week. Subjects completed the Positive and Negative Affect Schedules at each PC-test session, and SF-36 general health and Beck Depression Inventory questionnaires on one occasion.

Results—There was a 0.10 log unit increase in VA variability for every 0.58 logMAR increase (worse mean VA) ($p=0.001$). For subjects with reduced foveal thickness, mean VA explained more of the total VA variability than foveal thickness ($R^2=0.72$ and 0.46 , respectively, in simple linear regressions). There was a statistically significant 4.3% increased log VF area variability for every 50% mean log VF area decrease ($p<0.001$); explaining most of the total variability in log VF area variability ($R^2=0.44$). When controlling for mean log VF area, there was a statistically significant increase in log VF area variability for subjects with greater than minimal depressive symptoms ($p=0.015$), with increased mean irritability scores ($p=0.02$), decreased SF-36 physical functioning subscale scores ($p=0.03$), or decreased mean score for feeling active, strong and proud ($p=0.008$) (adjusted $R^2=0.62$). CS variability was low, and not statistically significantly related to mean CS, macular thickness or psychosocial factors.

Conclusions—Increased VA and VF variability was predicted largely by increased RP severity. Greater VF variability occurred in subjects with reduced VF who reported less physical activity or increased negative psychosocial states. These associations should be considered during clinical exams and trials for RP.

Keywords

retinitis pigmentosa; variability; fluctuation; visual field; psychosocial; affect; depression

Introduction

The reliability of psychophysical vision measures is dependent upon the performance of both the patient and examiner, as well as possible shifts in retinal sensitivity. Even with standard protocols and high levels of cooperation from the patient and examiner, test-retest variability can be considerable in some patients with retinal disease. Therefore, investigators

are currently working to validate patient-reported outcome measures of vision disability and quality of life¹⁻³ and imaging methods to assess retinal structural changes *in vivo*⁴⁻⁷ as possible surrogate markers for changes in vision, hoping that they will be more reliable outcome measures for use in future clinical trials. However, despite the potentially high variability that may occur during psychophysical vision testing in retinitis pigmentosa (RP), many clinical trials will likely continue to use visual acuity (VA) and/or visual field (VF) tests as primary outcome measures and a conservative test-retest reliability margin will have to be maintained. This means that the precision to determine whether a true change in vision has occurred may be reduced, unless we can determine if there are certain RP patients who are more susceptible to large fluctuations in vision and may therefore not be the best candidates for such trials.

Few previous studies have examined the amount of variability in visual acuity (VA) specifically in RP patients. First, Berson et al. used a Snellen acuity chart to evaluate intervisit variability in 32 RP subjects and determined a threshold criterion for significant change of VA of 0.102 logMAR⁸. Grover et al. and Fishman et al. found that the test-retest variability of VA between two sessions resulted in a criterion for significant change of seven or more letters (>0.14 logMAR) with the ETDRS charts in RP subjects with VA between 20/25 and 20/200^{9,10}. Kiser et al. also used the ETDRS charts and found that RP subjects with mean VA better than 20/40 had a significantly lower mean one-sided coefficient of repeatability of 0.12 logMAR across four test sessions than the 0.24 logMAR found in 22 RP subjects with mean VA between 20/40 and 20/1000¹¹. Although there was a greater group mean variability in subjects with reduced mean central acuity, there were several RP subjects with reduced VA who were as consistent as those with better VA. Birch et al. reported higher variability for VA among 67 RP patients whose VA was typically better than 20/50, and demonstrated a one-sided 95% CI of 0.21 with the Bailey-Lovie charts¹².

A considerable amount of variability has been previously documented during a study involving Humphrey VF testing in RP subjects versus normally-sighted controls. Seiple et al. found that the vast majority of the SDs for Humphrey thresholds on repeat visits were larger than the mean SD of normally-sighted controls, and about half were above the upper limit of the 99% confidence interval¹³. Another study in RP subjects found that the test-retest repeatability of the Humphrey VF test was nonmonotonically related to the level of sensitivity loss, with the largest variability in test locations with mean sensitivity of 10 dB¹⁴.

Seiple et al. previously commented that it would be valuable if the magnitude of a RP patient's variability could be predicted based on either the clinical findings or the amount of variability exhibited during the first two assessments¹³. In their study of a small group of 8 RP subjects with central acuity better than 20/40, they were unable to demonstrate any significant relationships between the magnitude of the variability of Humphrey visual field tests (average SD across four repeat visits) and subjects' age, Goldmann visual field extents, average Humphrey thresholds, average mfERG amplitudes, or average mfERG implicit times. Therefore, they and others advocate estimating each patient's variability as an individual criterion for a significant change in vision^{11,13,15}.

Ross et al. found that the variability of Goldmann VF area was about 2–3 times greater in RP subjects than in normally-sighted controls, with a mean of 12% between sessions, and up to 50%¹⁶. The authors did not comment on the magnitude of the variability of Goldmann VF area in relation to subjects' mean Goldmann VF extent. Examination of their published data indicates that as the Goldmann VF area decreases, some subjects' eyes demonstrate increased variability between-sessions, while others maintain low variability. This begs the question, why are some but not all prone to increased variability as vision loss increases? It

would be valuable to determine if there are factors that may help predict increased variability.

The current study was designed to explore whether the magnitude of the variability of VA, contrast sensitivity (CS) and VF in RP was related to identifiable factors. Previous studies have found limited support for the hypothesis that RP patients with more advanced vision loss are more likely to experience larger fluctuations in their vision^{11,16}. We therefore aimed to test this hypothesis in a group of RP patients with a broad range of vision loss and duration, and used spectral domain-optical coherence tomography (SD-OCT) to evaluate the severity of RP indicated by the extent of structural loss in macular thickness resulting from a loss of photoreceptors and RPE. Negative psychosocial states, such as anxiety and depression, are prevalent among RP patients¹⁷, who have previously reported that episodic changes in vision are perceived to be often related to stress¹⁸. Therefore, we also included a battery of psychosocial measures that could reveal possible associations with sleep disturbances, daytime sleepiness, perceived stress, positive and negative affect, depression, and general health.

Methods

The protocol for the study was approved by the Institutional Review Board of the Johns Hopkins University School of Medicine and followed the tenets of the Declaration of Helsinki. Informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study.

Subjects

Study participants included 27 individuals diagnosed with RP. The majority of the subjects (n=20; 74%) were recruited through the clinical practices at the Johns Hopkins Wilmer Eye Institute, from low vision optometrists or retinal specialists. The remaining 7 subjects were self-referred after learning of the study through online listings.

Individuals over age 18 were eligible for the study if they had a confirmed diagnosis of RP, with any level of vision, provided they could read reverse contrast, large sized font on the personal computer (PC) to complete the questionnaires and vision tests used in the study. Participants possessed basic computer skills at a minimum, and the study loaned a PC to subjects if they did not have regular access to one. Subjects were recruited with the premise that their ocular status was likely to remain stable throughout a 2- to 3-month period, during which the testing occurred.

Data Collection: Vision Measures

Data collection occurred from December 2008 through April 2010. The subjects enrolled at the Johns Hopkins Wilmer Eye Institute's Lions Vision Center and performed lab-based vision tests during a single session administered by a single examiner (AKB). Best-corrected VA was measured binocularly with the Early Treatment of Diabetic Retinopathy Study (ETDRS; Lighthouse International, New York) charts at 3 meters, or closer if fewer than 10 letters were identified. Best corrected, binocular Pelli-Robson CS (Metropia, Ltd., Essex, U.K.) was assessed at 1 meter. The VF in each eye was measured using the Goldmann VF V4e and III4e test targets. These lab-based measures were obtained to characterize the subjects' level of vision loss using standard test procedures, and were not used to quantify subjects' vision variability.

The study used a series of PC-based vision tests and questionnaires self-administered at home to assess variability of vision, instead of requiring several repeated visits to our lab to complete vision tests. This was done in order to facilitate study participation and feasibly

measure the variability of vision across a large number of test sessions within a short period of time in a reasonably sized group of presumably stable, untreated RP patients. These PC-based vision tests were previously developed at our center and validated in RP subjects while on placebo during a clinical trial¹⁹. The PC-tests included binocular letter VA and CS tested at 150 cm. During each trial a test letter was shown in the center of the screen for 3 seconds or until the subject right-clicked the mouse to shorten the trial, and the subject then had to pick a match among 10 letters (C,D,H,K,N,O,R,S,V,Z). The PC-tests also included a binocular, static VF test at 25 cm, based on the assumption that all patients had a central island. The VF radius was measured along meridians at 15 degree intervals, and the test algorithm flashed a dot along one of 24 axes, straddling the putative scotoma border; these tests are akin to Goldmann visual fields with a static (flashing) rather than kinetic (moving) test dot. All tests used a Bayesian adaptive algorithm to find the threshold in a minimum number of trials: 16 for VA, and 11 for CS and each VF axis.

The PC-test series consisted of 16 sessions completed twice a week over a period of about 2 to 3 months. At the study onset, it was not known whether vision might vary according to the time of the day (morning versus evenings) or day of the week (weekday versus weekend). Therefore, the times at which subjects were instructed to complete the PC-based tests were systematically assigned: four tests were taken in the morning before noon on either Wednesday or Thursday, four tests in the morning on weekend days (Saturday or Sunday), four tests in the evening after 6pm on either Wednesday or Thursday, and four tests in the evening on weekend days.

Participants installed and ran the PC-test software under Windows. During the initial study visit, all subjects received careful instruction regarding the PC-tests, including a detailed manual. Also, on-screen reminders were designed before each test to help the subjects with correct setup conditions. To standardize test conditions as much as possible, all subjects were given 150 and 25 cm distance gages, to be attached to the PC monitor, and were instructed to close window blinds or curtain prior to testing to attempt to maintain consistent room illumination across sessions. When the PC-test software was initially installed, each subject was instructed to calibrate their monitor's pixel size using a dollar bill they placed in the upper left corner of the monitor, and they used their mouse to move a horizontal line and vertical line to frame the dollar bill. The gamma function of their monitor was calibrated by comparing an even gray field with a field of fine black and white lines, while they adjusted the grayscale level of the field to match the average of the fine lines at a distance far enough that the lines were not resolvable. The formula we used to linearize the grayscale was: $L(n)/L(255) = (n/255)^\gamma$

Data Collection: Factors potentially related to variability of vision

Four standardized questionnaires to assess psychosocial factors were administered after each PC-based vision test session: the Stanford Sleepiness Scale²⁰, Epworth Sleepiness Scale²¹, Perceived Stress Scale (PSS)²² and the Positive and Negative Affect Schedules (PANAS)²³. The Stanford Sleepiness Scale consisted of a single item 7-point scale to determine sleepiness at the time of test administration, and the Epworth Sleepiness Scale contained 8 items inquiring about subjects' tendency to doze during various activities on the day of the test. The 14-item version of the Perceived Stress Scale (PSS) was used, and the PANAS consisted of 10 items to assess positive affect (interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, active) and 10 items to assess negative affect (distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, afraid), scored individually. The subjects were asked to rate the questionnaire items based on their experiences in the last 24 hours. On a single occasion during the course of the study, participants were mailed or emailed the following questionnaires to complete: Beck

Depression Inventory (BDI)²⁴, Pittsburgh Sleep Quality Index²⁵, and all subscales of the SF-36 questionnaire^{26,27}.

SD-OCT macular thickness measurements were obtained at the time of study enrollment from all subjects using the Heidelberg Spectralis HRA+OCT™ (Heidelberg Engineering, Vista, CA). In addition, control data from 27 age-matched normally-sighted individuals without retinal disease at our institution were used to determine the mean and SD for normal macular thickness using this same instrument and protocol. The acquisition protocol consisted of a series of horizontal raster scans covering 20° centered on the fovea. The raster scans were 240 microns from each other, with each raster scan consisting of a series of A-scans with a transverse resolution of 14 microns and in-depth resolution of 3.9 microns. The resolution of each line was enhanced by repeating and averaging the measurements on each line at least 15 times using the proprietary TruTrack™ image alignment software.

The measurements of retinal thickness represented the full thickness of the retina; from the inner border of Bruch's membrane (BM) to the internal limiting membrane (ILM) encompassing the full thickness of the retinal pigment epithelium (RPE) and photoreceptor complex. Each individual scan was reviewed by an experienced investigator (MAI) and the lines were manually corrected to ensure that the off-algorithm segmentation lines were correctly placed along the inner borders of the BM and ILM. Average macular thickness was calculated for 3 circular areas or rings within 1, 3 and 6mm areas from the fovea. The foveal area was defined by the most central 3.3° diameter circle. Parafoveal thickness was defined as the 3.3°–10° middle diameter ring around the foveal area, and the perifoveal thickness was defined as the 10°–20° outer diameter ring around the parafoveal area. Mean macular thickness was calculated using data from both eyes for comparison with the binocular PC-vision test results.

Data Analyses

We defined variability in vision as the SD of all measures across sessions for VA, CS, and log VF area (log of the viewing area at a 25 cm test distance calculated in cm²). We performed a log transformation of the VF area at each session that was used in the calculations of log VF area variability and mean log VF area, since it has been documented in several longitudinal, natural history studies that log VF declines linearly over time²⁸. We used simple and multiple linear regressions to analyze the data, using Stata/IC Version 10.0 (Stata Corp., College Station, TX). We explored whether there were any qualitative differences in our results when using mean macular thickness across both eyes for each subject versus the macular thickness for either the better or worse eye, but found none and therefore report our findings that used the mean macular thickness. We calculated the mean questionnaire score for each subject across all PC-test sessions by using the repeated measures of psychosocial questionnaires administered by PC (PSS, PANAS, SSS, ESS).

We used a systematic approach to explore the primary aim of the study, which was to identify significant factors that help explain the variability in vision in RP. First, we used simple linear regressions to determine whether age was significantly related to variability in each of the vision measures. We then used Welch's two-sample t-tests with unequal variances to test whether there were any significant differences in the variability of the vision tests according to gender or presence of CME. Next, we used simple and multiple linear regressions to ascertain the amount of variability in vision explained by the mean level of vision across all PC-test sessions and/or macular thickness (measures of RP severity). Last, if a majority of the variability in vision was not explained by the mean vision measure or macular thickness, we further explored whether any of the psychosocial variables either individually or collectively were able to explain any additional variability in vision.

Results

Subjects' duration of daytime vision loss ranged from 1 to 51 years (mean 16; SD 13 years), while their duration of night vision loss ranged from 1 to 71 years (mean 27; SD 18 years). 48% of the subjects were female, and subjects' ages ranged from 20 to 76, with a mean of 51 years. Seven percent of the subjects (n=2) were African-American, 7% were Hispanic and 85% (n=23) were Caucasian.

Across subjects, the mean ETDRS VA was 0.23 logMAR (SD 0.34, range -0.12, 0.98) and mean Pelli-Robson CS was 1.3 logCS (SD 0.6, range 0.05, 2.0). Mean PC-based VA was highly correlated with lab-based ETDRS VA ($r=0.96$) and mean PC-based CS was highly correlated with lab-based Pelli-Robson CS ($r=0.90$). These relationships are depicted in figure 1, which also shows that the simple linear regression slopes are close to unity ($m=1.12$ for VA; $m=0.89$ for CS). Across subjects, the mean Goldmann VF diameter was 54° (SD 53°, range 7°, 149°) and 39° (SD 42°, range 5°, 129°) with the V4e and III4e test targets, respectively, in the eye with the larger VF diameter.

Exactly one third of the RP subjects (n=9) had cystoid macular edema (CME) detectable with OCT in one or both eyes. Table 1 lists the mean, SD and range of macular thickness values across RP subjects' eyes with CME, RP subjects without CME, and normal controls without RP. Using t-tests, we found that on average, subjects with CME did not exhibit statistically significantly larger amounts of variability (SD) in VA, CS or log VF area when compared to those without CME. Table 2 displays the amount of variability in VA, CS and VF when comparing subjects with and without CME, by gender, and age. Using t-tests and simple linear regressions, we found that on average, variability in VA, CS or log VF area did not differ statistically significantly by gender or for every 1 year increase in age, respectively. Using repeated measures data analyses (generalized estimating equations), across subjects we found that there were no statistically significant differences in the VA, CS or VF measures according to the time of day (morning versus evening) or day of the week (Wednesday/Thursday versus weekend). Therefore, it was not necessary to adjust for time, age, gender or CME during the following analyses of variability in vision.

Using simple linear regressions, we explored whether overall increased negative psychosocial states (i.e. mean PSS, positive PANAS, negative PANAS or BDI scores) were qualitatively or statistically significantly associated with RP severity (i.e. reduced lab-based binocular ETDRS VA or smaller Goldmann VF diameter with the V4e test target), and found that no significant relationships emerged. We also examined whether increased RP severity was qualitatively or statistically significantly associated with poor reported general health status on the SF-36. These RP subjects did not appear to consider their level of vision when rating their health.

VA Variability

Using a simple linear regression, we found that on average there was a statistically significant 0.10 log unit increase in VA variability (SD) for every 0.58 logMAR increase (worsening) in mean VA (95% CI: 0.08, 0.12; $p=0.001$). Across all subjects, 80% of the total variability in VA variability was explained by the mean VA. The relationship between VA variability and mean VA is shown in the scatterplot in figure 2A. The regression lines in figure 2 were fit to all of the data regardless of CME status. Figure 2B demonstrates the non-linear, monotonically decreasing relationship between VA variability and mean foveal thickness across subjects. The drawn linear regression line shows that VA variability appears to increase for the 15 subjects with mean foveal thickness $<265\mu\text{m}$, indicated by the arrow in figure 2B, which was the mean foveal thickness for normal controls without RP. For the 15 subjects with mean foveal thickness $<265\mu\text{m}$, on average there was a 0.02 log

unit increase in VA variability for every 10 μ m reduction in mean foveal thickness (95% CI: 0.006, 0.03; $p=0.006$). However, among these subjects, mean VA still explained more of the total variability in VA variability ($R^2=0.72$) than mean foveal thickness ($R^2=0.46$) in simple linear regression models. Foveal thickness was no longer statistically significantly related to VA variability ($p=0.27$), when also considering mean VA ($p=0.003$) in a multiple linear regression model for the subjects with foveal thickness $<265\mu$ m.

VF Variability

Figure 3A shows the relationship between mean PC-based VF area and the Goldmann VF diameter measured with the V4e test target in the eye with the larger VF, while figure 3B displays the linear relationship between log VF area variability and mean log VF area. The regression line in figure 3B was fit to all of the data regardless of CME status. For every 50% decrease in mean log VF area, on average there was a statistically significant 4.3% increase in log VF area variability (SD) (95% CI: 2.2, 6.3%; $p<0.001$), based on a simple linear regression. A moderate amount of the total variability in log VF area variability (44%) was explained by the mean log VF area alone.

The relationships between log VF area variability and mean foveal and perifoveal thickness are shown in figure 3C. For every 100 μ m decrease in mean foveal thickness, on average there was a statistically significant increase in log VF area variability by 5.7% (95% CI: 0.7, 10.9%; $p=0.026$), and for every 100 μ m decrease in mean perifoveal thickness, on average there was a statistically significant increase in log VF area variability by 8.6% (95% CI: 0.3, 17.6%; $p=0.04$), based on simple linear regression models. Minimal amounts of the total variability in log VF area variability, only 19% and 16%, were explained by the mean foveal and perifoveal thickness alone, respectively. Moreover, including foveal or perifoveal thickness in a multiple linear regression model was not able to explain any more of the total variability in log VF area variability than mean log VF area alone.

We explored whether any of the psychosocial questionnaire responses helped to predict any more of the total variability in log VF area variability beyond what mean log VF area explained as a single predictor (44%). We found statistically significant relationships between log VF area variability and the SF-36 physical functioning subscale score, mean positive affect score, and subjects with greater than minimal depressive symptoms (BDI score ≥ 10 ; $n=9$), after adjusting for mean log VF area. Table 3 displays the crude (using simple linear regressions) and adjusted average amounts of increased log VF area variability, according to changes in each psychosocial variable, as well as the results adjusted for mean log VF area and all other variables in table 3. Fifty to 53% of the total variability in log VF area variability was explained when either the SF-36 physical functioning subscale score, the mean positive affect score or greater than minimal depressive symptoms was considered along with the mean log VF area. Figures 4A and 4B show the relationships for log VF area variability versus greater than minimal depressive symptoms, and SF-36 physical functioning, respectively.

Since decreased mean positive affect score was statistically significantly associated with decreased log VF area variability after adjusting for mean log VF area ($p=0.04$), we performed an exploratory analysis to examine which of the ten positive affect items were statistically significantly related to log VF area variability. Feeling active ($p=0.025$), strong ($p=0.016$) and proud ($p=0.009$) were each statistically significant, while enthusiastic ($p=0.06$), attentive ($p=0.09$), alert ($p=0.09$), determined ($p=0.09$) and inspired ($p=0.09$) approached statistical significance, after adjusting for mean log VF area. Figure 4C shows the association between log VF area variability and the mean score for “active,” “strong” and “proud.” The relationship is stronger for subjects with mean VF area $<200\text{cm}^2$ than for those with a larger mean VF area.

The mean negative affect score (PANAS) and perceived stress scale (PSS) were not significantly related to increased log VF area variability after adjusting for mean log VF area ($p=0.33$ and 0.26 , respectively). However, an exploratory analysis indicated that two items from these scales were statistically significantly associated with log VF area variability: the mean score for “irritable” from the PANAS and “feeling anger because of things that happened that were outside your control” from the PSS, shown in table 3. Fifty-one to 55% of the total variability in log VF area variability was explained when either the mean score for any of these statistically significant items was considered along with the mean log VF area. Mean log VF area, SF-36 physical functioning, and the mean scores for “proud” and “irritable” were able to account for 62% of the total variability in log VF area variability.

Table 4 displays the Pearson correlation coefficients for the vision and psychosocial variables. Mean PC-based VA, CS and VF were weakly correlated with most of the psychosocial variables, except for a moderate positive correlation found for CS and SF-36 physical functioning, and moderate negative correlations between VF area and feeling active and proud. The BDI depressive symptoms score was moderately correlated with the other psychosocial variables, and most correlated with feeling strong. The SF-36 physical functioning subscale was only weakly positively correlated with feeling active, strong and proud, and moderately negatively correlated with feeling irritable or anger because of things that happened outside your control. Feeling active, strong and proud were all highly positively correlated. Feeling irritable was highly positively correlated with anger due to uncontrollable situations, and moderately correlated with the other psychosocial variables. Therefore, it is not surprising that when we adjust for all other psychosocial variables in a multiple linear regression model (far right column in table 3), they become statistically non-significant.

CS Variability

Figure 5 shows the relationship between CS variability and mean CS across subjects, and the regression line was fit to all of the data regardless of CME status. Unlike VA and VF, we did not find an increase in the magnitude of CS variability, as CS or either of the other vision measures was reduced. Overall, CS variability was low with a mean SD of 0.09 log units, and was not statistically significantly related to mean CS ($p=0.32$), foveal thickness ($p=0.36$), or psychosocial factors.

Parafoveal thickness was not statistically significantly related to the variability of VA ($p=0.09$), CS ($p=0.55$), or log VF area ($p=0.13$). Sleepiness and sleep disturbances were not statistically significantly related to the variability of any of the vision measures.

Discussion

In the present study, we have shown that a substantial amount of VA variability was predicted by mean VA alone, while in comparison, accounting for physical functioning, reduced positive affect or increased negative psychosocial factors, in addition to mean VF, helped to explain VF variability. The strongest predictor of increased magnitude of VA or VF variability was increased RP severity, indicated by a measured reduction in mean VA or VF. These results support the hypothesis that more advanced vision loss is associated with less reliable vision. As photoreceptors and RPE are gradually lost due to disease progression in RP, the few remaining photoreceptors may not respond as consistently to light stimulation and visual input. As a consequence of retinal function loss, increased variability may also occur with less consistent fixation as macular areas outside the fovea are used.

Increased negative psychosocial states may occur in any RP patient, regardless of their level of vision loss, and may be related to reductions in their ability to successfully cope or

manage stressors. Our data also suggest that decreased physical functioning and positive affect are associated with increased VF variability. Previous research has suggested that positive affective states are associated with a broadening of attentional processes, which are aspects of cognition that control information processing, thought and behavior²⁹, and are also important elements of reliable VF results. Increased negative psychosocial states, such as irritability or anger, may result in inefficient attentional allocation, such that attention is shifted away from VF stimuli towards internal monitoring of thoughts and feelings. Increased depressive symptoms may lead to a more effortful and reduced capacity to process and react to information during a VF test, since remaining resources are directed toward depression-related thoughts other than the VF task, thus potentially affecting reliability. Therefore, interventions to maintain or improve RP patients' mental, emotional and physical health may reduce VF variability and should be explored in the future.

We performed a comprehensive search of previous publications examining VA and VF variability in RP patients, reviewed in the introduction. The present study's finding of an increased magnitude of VA and VF variability as mean vision measures were reduced was not revealed by some of the previous publications; this may be due to the narrower or skewed range of vision levels in the other study populations. The current results show some agreement with another analysis completed in this group of RP patients, which found that during PC-test sessions when the VF was reduced, subjects were more likely to report a corresponding reduction in their general health or increased stress on the day of the test³⁰. That analysis differed from the ones presented here since we used multilevel, repeated measures models to examine day-to-day changes in psychosocial variables in relation to PC-test sessions when subjects experienced a significant deviation in vision from their mean.

The current findings also replicate the previous finding that the magnitude of CS variability did not increase according to reductions in mean CS¹¹. CS appears to be relatively insensitive to day-to-day potential changes in testing conditions, but is not likely insensitive to disease state in moderate to advanced RP since our subjects demonstrated approximately equal reductions in VA and CS, and subjects with CS <1.0 logCS also had reduced VA >0.2 logMAR. CS measured with large letter targets may be more sensitive to changes over larger areas of retina, and may depend on the pattern of RP progression or genetics, as well as the type of treatment or its target in the retina, since CS is processed by inner retinal or cortical areas. Low CS variability across subjects regardless of CS impairment indicates that this may be a reliable, useful measure of central visual function changes in some clinical trials for RP.

This study enrolled RP patients with a wide range of ages above 18 and duration of vision loss, was well balanced in terms of gender, and covered a broad range of measurable vision. The subjects in this study self-selected to participate after learning of the study objectives. It would be difficult to determine if and how they are different from RP subjects who volunteer for treatment clinical trials. Interestingly, most of the subjects who reported reduced SF-36 physical functioning had smaller mean VFs, reduced mean CS, but not necessarily increased age. Perhaps some of the reported reductions in physical activity captured subjects' difficulty with orientation and mobility due to visual impairment. It has been previously reported that reductions in CS and VF are both related to reductions in self-reported mobility³¹, walking speed³², and mobility performance³³ in RP.

Self-administered PC-based vision tests are being increasingly developed and used for vision testing^{34–37}. They represent a feasible method to obtain frequent, repeated measures outside of the lab or clinic. PC-tests are especially valuable since they help facilitate participation of subjects who do not reside locally to the study center, drive, or have access to public transportation. For this study evaluating the natural day-to-day variability of

vision, it was important to be able to obtain frequent measurements in the subjects' homes, and thus capture their typical living circumstances and stressors. One might argue that imprecision in vision measurement may be inherent with self-administration. This appears to be unlikely, however, since one of the three tests (CS) was very reliable across sessions; moreover, these PC-tests have been previously documented to be as reliable as lab-based vision tests in RP¹⁹, and other computer-based vision test systems have also been reported as reliable³⁸.

Conclusions

Greater VA and VF variability was predicted largely by greater RP severity, defined by a reduction in mean VA or VF across test sessions. Greater VF variability was also explained in part by reduced physical functioning or mean positive affect, or increased depressive symptoms or negative psychosocial states. For reasons other than variability in vision, RP subjects reporting poor physical functioning, less positive moods or greater negative psychosocial states, such as depression, anger and irritability, may not be the most ideal candidates for early phase safety and efficacy trials. Clinicians and researchers who encounter significant amounts of VF variability among RP patients should consider and account for these possible factors. Individuals with more advanced RP who are still relying on their vision may experience reduced quality of life, which may not only be due to their restricted mean level of vision, but also to the uncertainty of large episodic fluctuations in vision. Therefore, it is vital to develop strategies to maintain or improve not only RP patients' retinal function, but also their overall mental, emotional and physical well-being.

Acknowledgments

The authors wish to thank Liancheng Yang for the programming for the PC-based tests and questionnaires. Funding support: NIH K23EY018356 to AKB.

REFERENCES

1. Massof RW. A clinically meaningful theory of outcome measures in rehabilitation medicine. *J Appl Meas.* 2010; 11(3):253–270. [PubMed: 20847474]
2. Massof RW. Moving toward scientific measurements of quality of life. *Ophthalmic Epidemiol.* 2008 Jul–Aug; 15(4):209–211. [PubMed: 18780252]
3. Massof RW, Ahmadian L, Grover LL, et al. The Activity Inventory: an adaptive visual function questionnaire. *Optom Vis Sci.* 2007 Aug; 84(8):763–774. [PubMed: 17700339]
4. Charbel Issa P, Troeger E, Finger R, Holz FG, Wilke R, Scholl HP. Structure-function correlation of the human central retina. *PLoS One.* 2010 Sep 22; 5(9):e12864. [PubMed: 20877651]
5. Apushkin MA, Fishman GA, Alexander KR, Shahidi M. Retinal thickness and visual thresholds measured in patients with retinitis pigmentosa. *Retina.* 2007 Mar; 27(3):349–357. [PubMed: 17460591]
6. Rangaswamy NV, Patel HM, Locke KG, Hood DC, Birch DG. A comparison of visual field sensitivity to photoreceptor thickness in retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 2010 Aug; 51(8):4213–4219. [PubMed: 20220048]
7. Jacobson SG, Roman AJ, Aleman TS, et al. Normal central retinal function and structure preserved in retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 2010 Feb; 51(2):1079–1085. [PubMed: 19797198]
8. Berson EL, Sandberg MA, Rosner B, Birch DG, Hanson AH. Natural course of retinitis pigmentosa over a three-year interval. *Am J Ophthalmol.* 1985 Mar 15; 99(3):240–251. [PubMed: 3976802]
9. Grover S, Fishman GA, Gilbert LD, Anderson RJ. Reproducibility of visual acuity measurements in patients with retinitis pigmentosa. *Retina.* 1997; 17(1):33–37. [PubMed: 9051840]

10. Fishman GA, Gilbert LD, Anderson RJ, Marmor MF, Weleber RG, Viana MA. Effect of methazolamide on chronic macular edema in patients with retinitis pigmentosa. *Ophthalmology*. 1994 Apr; 101(4):687–693. [PubMed: 8152764]
11. Kiser AK, Mladenovich D, Eshraghi F, Bourdeau D, Dagnelie G. Reliability and Consistency of Visual Acuity and Contrast Sensitivity Measures in Advanced Eye Disease. *Optom Vis Sci*. 2005; 82:946–954. [PubMed: 16317369]
12. Birch DG, Anderson JL, Fish GE. Yearly rates of rod and cone functional loss in retinitis pigmentosa and cone-rod dystrophy. *Ophthalmology*. 1999 Feb; 106(2):258–268. [PubMed: 9951474]
13. Seiple W, Clemens CJ, Greenstein VC, Carr RE, Holopigian K. Test-retest reliability of the multifocal electroretinogram and humphrey visual fields in patients with retinitis pigmentosa. *Doc Ophthalmol*. 2004 Nov; 109(3):255–272. [PubMed: 15957611]
14. Kim LS, McAnany JJ, Alexander KR, Fishman GA. Intersession repeatability of humphrey perimetry measurements in patients with retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2007 Oct; 48(10):4720–4724. [PubMed: 17898296]
15. Brown B, Lovie-Kitchin J. Repeated visual acuity measurement: establishing the patient's own criterion for change. *Optom Vis Sci*. 1993 Jan; 70(1):45–53. [PubMed: 8430008]
16. Ross DF, Fishman GA, Gilbert LD, Anderson RJ. Variability of Visual Field Measurements in Normal Subjects and Patients with Retinitis Pigmentosa. *Arch Ophthalmol*. 1984 Jul; 102(7):1004–1010. [PubMed: 6743076]
17. Strougo Z, Badoux A, Duchanel D. Psycho-affective Problems Associated with Retinitis Pigmentosa. *J Fr Ophtalmol*. 1997; 20(2):111–116. [PubMed: 9099269]
18. Heckenlively JR, Yoser SL, Friedman LH, Oversier JJ. Clinical Findings and Common Symptoms in Retinitis Pigmentosa. *Am J Ophthalmol*. 1988 May; 105:504–511. [PubMed: 3259404]
19. Bahrami H, Melia M, Dagnelie G. Lutein Supplementation in Retinitis Pigmentosa: PC-Based Vision Assessment in a Randomized Double-Masked Placebo-Controlled Clinical Trial. *BMC Ophthalmol*. 2006 Jun 7.6(1):23. [PubMed: 16759390]
20. Hoddes E, Dement WC, Zarcone V. The Development and Use of the Stanford Sleepiness Scale. *Psychophysiology*. 1972; 9:150.
21. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991; 14(6):540–545. [PubMed: 1798888]
22. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. *J Health Soc Behav*. 1983; 24(4):385–396. [PubMed: 6668417]
23. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol*. 1988 Jun; 54(6):1063–1070. [PubMed: 3397865]
24. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An Inventory for Measuring Depression. *Arch Gen Psychiatry*. 1961 Jun; 4:561–571. [PubMed: 13688369]
25. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a New Instrument for Psychiatric Practice and Research. *Psychiatry Res*. 1989 May; 28(2):193–213. [PubMed: 2748771]
26. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual Framework and Item Selection. *Med Care*. 1992; 30:473–483. [PubMed: 1593914]
27. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993; 31:247–263. [PubMed: 8450681]
28. Massof RW, Dagnelie G, Benzschawel T, Palmer RW, Finkelstein D. First-order dynamics of visual field loss in retinitis pigmentosa. *Clin Vis Sci*. 1990; 5:1–26.
29. Rowe G, Hirsh JB, Anderson AK. Positive affect increases the breadth of attentional selection. *Proc Natl Acad Sci U S A*. 2007 Jan 2; 104(1):383–388. [PubMed: 17182749]
30. Bittner AK, Haythornthwaite JA, Diener-West M, Dagnelie G. Measured and Reported Worse-than-Usual Vision in Retinitis Pigmentosa are Related to Perceived Stress, Light Exposure or General Health. 2010 Submitted to Retina in November.

31. Turano KA, Geruschat DR, Stahl JW, Massof RW. Perceived visual ability for independent mobility in persons with retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 1999 Apr; 40(5):865–877. [PubMed: 10102283]
32. Geruschat DR, Turano KA, Stahl JW. Traditional measures of mobility performance and retinitis pigmentosa. *Optom Vis Sci*. 1998 Jul; 75(7):525–537. [PubMed: 9703042]
33. Haymes S, Guest D, Heyes A, Johnston A. Mobility of people with retinitis pigmentosa as a function of vision and psychological variables. *Optom Vis Sci*. 1996 Oct; 73(10):621–637. [PubMed: 8916133]
34. Shah N, Laidlaw DA, Brown G, Robson C. Effect of letter separation on computerised visual acuity measurements: comparison with the gold standard Early Treatment Diabetic Retinopathy Study (ETDRS) chart. *Ophthalmic Physiol Opt*. 2010 Mar; 30(2):200–203. [PubMed: 20444125]
35. Schlenker MB, Christakis TJ, Braga-Mele RM. Comparing a traditional single optotype visual acuity test with a computer-based visual acuity test for childhood amblyopia vision screening: a pilot study. *Can J Ophthalmol*. 2010 Aug; 45(4):368–374. [PubMed: 20648081]
36. Beck RW, Moke PS, Turpin AH, et al. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol*. 2003 Feb; 135(2):194–205. [PubMed: 12566024]
37. Arditi A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol Vis Sci*. 1993 Jan; 34(1):120–129. [PubMed: 8425819]
38. Paysse EA, Camejo L, Hussein MA, Coats DK. Parent-administered visual acuity testing: is it reliable and can it improve office efficiency? *J AAPOS*. 2004 Aug; 8(4):332–337. [PubMed: 15314593]

Figure 1.

Scatterplot demonstrating the relationship between mean PC-based VA across sessions and lab-based ETDRS VA (dark triangles), and the relationship between mean PC-based CS across sessions and lab-based Pelli-Robson CS (open triangles). Each point represents a subject, and the drawn lines represent the linear regression lines.

Figure 2.

Scatterplots demonstrating the relationship between (A) variability of VA and mean VA, (B) variability of VA and mean foveal thickness. The arrow represents the mean foveal thickness for normal controls. Each point represents a subject, depicting whether they had CME (open triangle) or did not have CME (dark triangle) in at least one eye. The drawn lines represent the linear regression lines.

Figure 3.

Scatterplots demonstrating the relationship between (A) mean PC-based VF area and the Goldmann VF diameter measured with the V4e test target in the better eye, (B) variability of log VF area and mean log VF area, and (C) variability of log VF area and mean foveal and perifoveal thickness. Each point represents a subject; figure C depicts whether they had CME (open triangle) or did not have CME (dark triangle) in at least one eye. The drawn lines in figures B and C represent linear regression lines.

Figure 4.

Box plot (A) for variability of log VF area by BDI depressive symptoms greater than minimal (≥ 10). The bottom and top of the box are the 25th and 75th percentile (lower and upper quartiles, respectively), and the band near the middle of the box is the 50th percentile (the median). An outlier is indicated by a dot above the upper quartile by more than 1.5 times the interquartile range. Scatterplots demonstrating the relationship between variability of log VF area and (B) SF-36 physical functioning score, and (C) the mean score for active, strong and proud across sessions. Each point represents a subject, depicting whether their VF area was <200cm² (dark triangles) or >200cm² (open triangles). The drawn lines in B and C represent linear regression lines.

Figure 5. Scatterplot demonstrating the relationship between variability of CS and mean CS. Each point represents a subject, depicting whether they had CME (open triangle) or did not have CME (dark triangle) in at least one eye. The drawn line represents a linear regression line.

Macular thickness values (μm) by subfield for RP subjects with CME, RP subjects without CME, and normal controls (n=27).

Table 1

Thickness	Foveal Thickness		Parafoveal Thickness		Perifoveal	
	mean (range)	SD	mean (range)	SD	mean (range)	SD
Subject groups						
RP subjects with CME	305 (174–580)	124	327 (208–448)	66	311 (245–405)	45
RP subjects without CME	236 (156–331)	53	262 (212–351)	43	243 (176–323)	40
Normal controls without RP	265 (219–320)	22	333 (283–382)	17	296 (251–354)	20

Table 2

Difference in mean SD* of VA, CS and VF tests when comparing subjects with and without CME, by gender, and age (n=27).

Clinical & Demographic Variables	VA logMAR (95% CI) P value	CS logCS (95% CI) P value	VF (cm ²) log area (95% CI) P value
CME vs. No CME †	0.001 (−0.06, 0.06) 0.97	−0.005 (−0.04, 0.03) 0.77	0.01 (−0.10, 0.13) 0.82
Female vs. male gender †	−0.006 (−0.07, 0.06) 0.85	0.009 (−0.03, 0.05) 0.62	−0.04 (−0.13, 0.05) 0.32
Age ‡	0.001 (−0.0007, 0.003) 0.21	−0.0005 (−0.002, 0.0007) 0.38	0.002 (−0.001, 0.005) 0.26

* SD of PC-test measures across all sessions

† Using Welch's two-sample t-tests with unequal variances

‡ Using simple linear regression models; Per every 1 year increase

Table 3

Crude and adjusted mean SD* of log visual field area according to changes in each variable (n=27).

Psychosocial Variables	Crude (SLR) % (95% CI) P value	Adjusted for mean log VF area % (95% CI) P value	Adjusted for all variables % (95% CI) P value
SF-36 Physical Functioning [†]	3.0% (0.6, 5.4%) 0.016 [†]	2.1% (0.2, 4.0%) 0.03 [†]	2.0% (-0.6, 4.6%) 0.12
More than minimal depressive symptoms [§]	11.1% (2.8, 19.4%) 0.01 [†]	8.3% (1.8, 14.9%) 0.015 [†]	1.7% (-12.0, 8.7%) 0.74
Mean Score for Active, Strong, Proud ^{//}	0.9% (-1.2, 3.0%) 0.40	2.0% (0.6, 3.5%) 0.008 [†]	2.1% (-0.2, 4.5%) 0.07
Mean Score for Irritable ^{//}	7.5% (-2.2, 18.2%) 0.13	8.5% (1.5, 16.1%) 0.02 [†]	7.1% (-3.7, 19.1%) 0.19
Anger due to uncontrollable situations ^{//}	5.0% (-3.0, 13.6%) 0.22	6.8% (1.0, 12.9%) 0.024 [†]	5.0% (-5.8, 17.0%) 0.36

* SD of measures across all sessions

[†] Statistically significant (p<0.05)

[‡] Per every 10 point decrease in questionnaire score

[§] BDI score ≥10

^{//} Per every 1 point score decrease

Table 4

Pearson Correlation Coefficients (r) for Vision and Psychosocial Variables (n=27)

	<u>logMAR VA</u>	<u>logCS</u>	<u>Log VF Area</u>	<u>BDI</u>	<u>Active</u>	<u>Strong</u>	<u>Proud</u>	<u>Irritable</u>	<u>SF-36</u>
logCS	-0.67*								
Log VF Area	-0.37	0.48							
BDI depressive symptoms	0.03	-0.16	0.10						
Active	0.19	-0.13	-0.37	-0.58					
Strong	-0.06	0.11	-0.12	-0.70*	0.83*				
Proud	0.09	-0.07	-0.36	-0.54	0.73*	0.87*			
Irritable	0.05	-0.18	0.06	0.51	-0.50	-0.58	-0.38		
SF-36 Physical Functioning	-0.08	0.42	0.24	-0.30	0.03	0.13	0.08	-0.31	
Anger due to uncontrollable situations	0.15	-0.27	0.12	0.51	-0.70*	-0.70*	-0.52	0.81*	-0.41

* Bolded coefficients with an asterisk are highly statistically significant (p<0.001)

Bolded coefficients without an asterisk are statistically significant (p<0.05)