A Novel Method for Quantitative Serial Autofluorescence Analysis in Retinitis Pigmentosa Using Image Characteristics

Jasleen K. Jolly1–3, Siegfried K. Wagner1–3, Jonathan Moules4, Florian Gekeler5, Andrew R. Webster3, Susan M. Downes1,2, and Robert E. MacLaren1–3

1Nuffield Department of Clinical Neurosciences, Oxford Biomedical Research Centre, University of Oxford, Oxford, UK
2Oxford Eye Hospital, John Radcliffe Hospital, Oxford, UK
3Moorfields Eye Hospital-UCL Institute of Ophthalmology NIHR Biomedical Research Centre, London, UK
4LightPear, Oxfordshire, UK
5University Hospital Eye Clinic, Stuttgart, Germany

Correspondence: Robert E. MacLaren, Nuffield Laboratory of Ophthalmology, Levels 5&6 John Radcliffe Hospital West Wing, Headley Way, Headington, Oxford OX3 9DU.
e-mail: enquiries@ndcn.ox.ac.uk

Received: 17 March 2016
Accepted: 13 September 2016
Published: 1 December 2016

Keywords: retinitis pigmentosa; autofluorescence; quantification; retinal degeneration


Purpose: Identifying potential biomarkers for disease progression in retinitis pigmentosa (RP) is highly relevant now that gene therapy and other treatments are in clinical trial. Here we report a novel technique for analysis of short-wavelength autofluorescence (AF) imaging to quantify defined regions of AF in RP patients.

Methods: Fifty-five–degree AF images were acquired from 12 participants with RP over a 12-month period. Of these, five were identified as having a hyperfluorescent annulus. A standard Cartesian coordinate system was superimposed on images with the fovea as the origin and eight bisecting lines traversing the center at 45 degrees to each other. Spatial extraction software was programmed to highlight pixels corresponding to varying degrees of percentile fluorescence such that the parafoveal AF ring was mapped. Distance between the fovea and midpoint of the AF ring was measured. Percentage of low luminance areas was utilized as a measure of atrophy.

Results: The hyperfluorescent ring was most accurately mapped using the 70th percentile of fluorescence. Both the AF ring and peripheral hypofluorescence showed robust repeatability at all time points noted (P = 0.93).

Conclusions: Both a hypofluorescent ring and retinal pigment epithelium atrophy were present on a significant proportion of RP patients and were consistently mapped over a 12-month period. There is potential extrapolation of this methodology to wide-field imaging as well as other retinal dystrophies. This anatomical change may provide a useful anatomical biomarker for assessing treatment end points in RP.

Translational Relevance: Spatial extraction software can be a valuable tool in the assessment of ophthalmic imaging data.

Introduction

The emergence of molecular therapies for inherited retinal degenerations has highlighted the need for imaging modalities that can sensitively identify structural disease progression. One such technique is short-wavelength fundus autofluorescence (AF) using confocal scanning laser ophthalmoscopy (cSLO), a noninvasive method that evaluates retinal disease. Increased AF is believed to indicate regions of elevated fluorophores, primarily lipofuscin, a product of photoreceptor outer-segment disc shedding and phagocytosis,1,2 while decreased AF may signify obstruction of the autofluorescent signal or atrophy of the underlying retinal pigment epithelium (RPE) and/or photoreceptors.3–5

In retinitis pigmentosa (RP), AF has been recognized to identify a parafoveal ring of increased AF in some patients6 and varying degrees of patchy hypofluorescence within the periphery. Classically, a region of preserved normal central AF is surrounded by an annulus of increased AF, which delineates disruption of the photoreceptor inner segment and outer-segment (IS/OS) junction on spectral-domain
optical coherence tomography (SDOCT). Gradual constriction of the hyperfluorescent annulus occurs as the disease progresses, while in the peripheral retina, hypofluorescent lesions of varying size have been shown to act as a surrogate for duration of disease. In view of these findings, areas of both hyperfluorescence and hypofluorescence can be combined and analyzed as a useful disease biomarker.

Studies exploring the implications of increased AF in retinal dystrophies have demonstrated that hyperfluorescent demarcation arcs correlate with functional parameters. In patients with electrophysiologically confirmed RP, pattern electroretinography (PERG) P50 amplitude is strongly associated with the radius of the hyperfluorescent annulus among patients with RP and normal visual acuity. Psychophysical testing has shown a strong association between both automatic and kinetic perimetry with the parafoveal ring radius, although this was only pertinent in those without any atrophic disease within the vascular arcades, suggesting that AF evaluation may be more useful in early disease. Similar results have been found with microperimetry. Such correlation suggests that AF may provide a useful additional endpoint in illustrating disease progression.

Most published work regarding quantitative AF utilizes luminance data, but special adaptations and processing are required. Delori et al. described a powerful technique where highly repeatable measures of AF can be achieved through installation of an internal fluorescent reference to obviate concerns over cSLO gain sensitivity and laser variation. Alternatively, area measurements are utilized, such as in Best vitelliform macular dystrophy or Stargardt disease; however, these are not appropriate for RP as this may omit part of the disease process.

It is not always possible to capture images in a routine clinical setting in a way that is compatible with the quantitative analysis techniques reported. Moreover, the internal fluorescent reference is not commercially available. We included AF analysis in the observational arm of a clinical trial assessing the impact of transcorneal electrical stimulation in participants with RP. As part of that study, we were able to rigorously assess AF changes in the control eyes of participants at regular time points. This provided insight into the baseline characteristics of AF in RP and most importantly on the reproducibility of serial AF analysis in individual patients over a period of time. Here, we report a novel method for serially evaluating AF in patients with RP, which utilizes luminance values in order to interpret distance data and is therefore less influenced by raw values, which might vary from session to session. This method does not require the installation of additional equipment to the cSLO. Such a method could inform patient prognosis through revealing the rate of AF loss as well as provide a useful structural trial end point in subsequent studies.

Methods

Participants with confirmed RP were recruited into an open-label observational trial assessing a novel therapy for RP (NCT01847365) at the Oxford University Eye Hospital and Moorfields Eye Hospital. Participants underwent clinical assessment and AF imaging at the baseline visit and each subsequent 3 months for a total of 1 year. Autofluorescence images of the control eye of a subset of five participants were analyzed to inform the development of this methodology.

The study was approved by the National Health Service (NHS) Research Ethics Committee and received institutional approval from the University of Oxford Clinical Trial Research and Governance Unit (12/SW/0293). The study protocol was carried out in accordance with the Declaration of Helsinki (2008) and good clinical practice. All clinical trial participants provided informed consent prior to enrolment.

Pupil mydriasis was achieved through administration of 2.5% phenylephrine and 1% tropicamide. Autofluorescence images with a 55-degree field were acquired using a confocal scanning laser ophthalmoscope (Heidelberg Spectralis; Heidelberg-Engineering, Heidelberg, Germany) with an excitation filter of 488 nm in automatic real-time mode, using a minimum average of 25 scans. A standard corneal keratometry (K) measurement of 7.70 was used for each patient.

Images were exported as TIFF files in an anonymized format. Geographic Information Systems (GIS) Extract Transform Load software (FME, Safe Software Inc., Surrey, BC, Canada) was then used to evaluate the AF images to allow automated analysis to create a number of outputs. The images were first georeferenced (“georeferencing” is the GIS term for aligning images consistently with a known base) by ensuring the fovea on each AF image was given coordinates of 0,0. The image was then scaled to place the center of the optic cup at a constant x coordinate of 200 units (equivalent to 300-μm distance). Following this, the image was then rotated around the central coordinates (0,0) to give the optic
cup center marking a y coordinate of 0, thus giving the optic cup final coordinates of 200,0 consistently in all images. Although it is appreciated that the fovea is inferior to the optic disc, this distance can vary with refractive status, and the direct linear distance at the 200,0 position provides consistent disc–fovea alignment for all images. As the rotation involves interpolation, the red and yellow marks were removed from the images prior to this to stop data contamination. Following the rotation, all image pixels were converted into a square polygon, termed a cell, which contained the hyperfluorescence value for the pixel from which it was derived (Supplementary Fig. 1).

Four standardized bisecting lines were created, one each for vertical and horizontal and the two diagonals, to create an eight-point star, with the center of the point being 0,0, where it intersected the fovea (Supplementary Fig. 1B). The lines were of a defined length and were used for each image to ensure consistency of the area being analyzed across patients and over time. Using a spatial filter, each bisecting line was spatially compared to the cells, and any cells that intersected it (overlapped, touched, contained, crossed, or within) were kept in the sequence they appeared from start to end along the bisecting line. The maximum hyperfluorescence value (max) across all cells from each image was then extracted and used as the reference.

In order to improve the consistency of analysis of the ring, a series of image refinements were applied to each image. First, a visualization output was created that demonstrated the extent of the hyperfluorescent ring. The cells were then colored depending on whether they were within a varying percentage of the max value (set to green) or not (set to red), that is, within that percentage of the maximal luminance value in the image. The green cells were also buffered to make them larger and more obvious before all the cells were output on top of the georeferenced image. The percentage was adjusted between 60% of the maximum luminance to 90% of the maximal luminance in 5% increments, and all input images were processed for each percentage. The resulting images were examined by clinicians to decide which percentage best captured the extent of the hyperfluorescent ring across the cohort, the final decision being 70%, as it equated to a threshold where six of eight meridians were captured in all participant images (Fig. 1). The pattern of change over time for each percentile is shown in Figure 2.

With the 70% chosen as the cutoff percentage, the FME workspace was expanded. Another output was created that started by removing all cells that had a hyperfluorescence value of less than 70% of the max value in order to leave only those points representing the hyperfluorescent ring. Of the cells that passed this test, the distance was calculated in micrometers from the cell to the fovea. As the fovea splits each bisecting line in half following the georeferencing process described above, this resulted in distances from eight transects to the middle, each separated by 45 degrees and centered on the fovea. Along each half line, the nearest and farthest cell distances were extracted and the midpoint between them calculated by dividing the difference between them by two. This
number represents the distance from the fovea to the hypothetical midpoint of the hyperfluorescent ring at the fovea. By summing together the two midpoint distances for each half of their respective line, a single value was determined that represented the diameter of the midsection of the hyperfluorescent ring on that bisecting line.

Finally, the georeferenced image was taken and clipped to a circle, centered on the fovea that was slightly wider than the optic cup (Supplementary Fig. 2). This was to ensure the same area of the image and identical features were being analyzed across the cohort regardless of the centration of the original image as shown in Supplementary Figure 2. The luminance value of each pixel was measured, and the percentage of pixels with a value within the lowest 10\% was returned. This percentage, highlighted in green, was taken to indicate the degree of atrophy and the optic disc. The optic disc was assumed not to change significantly over this time period, so any change detected in serial analysis can be attributed mainly to progression of atrophy.

These processes were conducted on a series of images from six eyes with three different sensitivity measures for each eye. This was conducted to assess the robustness of the measurements in images of varying quality to reflect a routine clinical scenario. Repeated measures ANOVA was used to investigate statistical significance.

**Results**

The AF images of the control eye for three male and two female participants, aged 18 to 56, with confirmed rod–cone dystrophy over a 12-month period were included (Table). Participant A, B, and C had well-defined parafoveal annuli of increased AF and varying degrees of decreased patchy AF in the periphery. Participant D had advanced rod–cone dystrophy with significant speckled foveal hyperfluorescence. Participant E had a diffuse region of increased AF parafoveally with scalloped peripheral areas of decreased AF (Fig. 3).

Sensitivity was set automatically in order to optimize the image captured. To assess the extent to which the sensitivity value impacts the measurements described from AF images, a subset of participants underwent serial testing with three different sensitivity values: 80\% to 89\%, termed dim; 90\% to 99\%, termed mid; and 100\% to 107\%, termed bright. The results are shown in Supplementary Table 1. Repeated measures ANOVA showed a significant difference between AF images captured at dim versus mid conditions ($P = 0.03$), while those taken at mid...
sensitivity were not significantly different from bright images \((P = 0.32)\). Similar results were found when highlighting areas below the 10th percentile of brightness—dim versus mid sensitivity neared statistical significance \((P = 0.09)\), while those between mid and bright conditions did not \((P = 0.21)\).

Autofluorescence images over a 1-year period underwent serial analysis at baseline, 6 months, and 12 months for change in AF. Figure 2A–E, displays serial analysis at each of the cutoff points. The 70% cutoff displays the most stable results in all the participants and displays small error bars when the results are combined from the participants (Fig. 2F). Standard deviation of mean radius was 486 μm at 70% versus the highest error at 65% (536 μm).

Changes in regions of AF below the 10th percentile were evaluated in the same five participants and shown in Figure 5. Area of the image below the 10th percentile increased in five participants, A, B, C, D, and E, by 0.23%, 3.05%, 1.82%, 1.56%, and 17.7% respectively.

**Discussion**

With potential treatments such as gene therapy, now entering clinical trials for inherited retinal dystrophies, there has been an impetus to develop robust methods of analyzing disease progression. Autofluorescence may provide additional insight into the structural changes occurring following an intervention and act as an additional clinical trial end point. Our proposed technique uses a pixel-by-pixel percentile...
grading system to identify desired regions above or below a defined threshold. Georeferencing of the image and establishment of the fovea as the central coordinate in a standard Cartesian plane allows comparison within the same patient between different images while obviating major concerns with variability in image magnification. In our study, the 70th percentile (i.e., all points within 70% luminance of the brightest point in the image) was chosen as the threshold that best highlighted the hyperfluorescent annulus when considering the presence of false positives and false negatives. Despite tailoring the choice of percentile in this study, there were unavoidable points of artefactual fluorescence, and conversely, regions of the AF ring could be missed as the level of AF varies along the circumference. However, the percentile can be adjusted between subjects to optimize capture of the AF ring provided that the same level is used within images of the same individual. Note that as the percentile threshold was increased, the annulus radius decreased. We hypothesize that this may reflect the centripetal progression of the hyperfluorescent annulus where the region of greatest hyperfluorescence lies along the inner aspect. We have also shown the measurement method is robust to changes in the sensitivity setting when capturing images for all values ≥90. This can be used to guide future photography. This methodology leads to reproducible measures between visits, allowing for quantification of disease change.

It has been shown that AF can vary among healthy individuals depending on race, sex, and age.16 This method relies on percentile grading rather than absolute luminance values, mitigating the impact of the general background AF of the subject. Our study benefited from a number of consistencies—device sensitivity did not vary significantly, and images were acquired by the same two experienced operators throughout the study. Subjects had good fixation and were familiar with the device. Exclusion criteria were also such that participants had no other significant retinal pathology besides RP and no substantial media opacity. Undoubtedly, the presence of cataract or corneal pathology would impair the quality of AF while epiretinal membrane and cystoid macular edema might confound pixel capture.

Mapping of the AF ring was achieved through four major meridians of equal length traversing the fovea and at 45 degrees to each other, forming a regular octagram star polygon. Thus, the predominant direction of disease progression could be mapped such that superior ring constriction may alert one to the deterioration of the inferior visual field. In our cohort, participant A had greatest reduction in annulus radius in the temporal region, whereas the superonasal aspect was greatest in participant B. There are a plethora of options around meridian position and number: Cases of sectoral RP could be described with further capture lines in the appropriate sector. Similarly, the method could be extrapolated to form longer meridians in wide-field imaging. In theory, the progression of mid-peripheral granular hypofluorescence could also be measured.

Figure 5. Changes in area of the AF image under the 10th percentile, chosen as a surrogate for measurement of RPE and photoreceptor atrophy. Participants A–D exhibited an increase in peripheral hypofluorescence. The advanced disease of participant E led to highly fluctuant results suggesting that the method is not as reliable in patients with late disease.
The potential of this technique can expand beyond rod–cone dystrophies. Holz et al.17,18 have shown that the boundaries of geographic atrophy secondary to age-related macular degeneration, which demonstrate increased AF, are more likely to progress. Setting an appropriate percentile threshold could quantify the advancement of the lesion.

Our study assessed multiple AF imaging time-points in detail over a 12-month period in the control arm of an RP interventional trial. Twelve months is a relatively short period in the progression of RP, thus allowing us to explore the repeatability of the method. This was shown effectively over each 6-month period. Although our primary objective was to show reproducibility of the methodology, our study mirrored previous reports in demonstrating that serial AF analysis in RP identifies a progressive non–statistically significant constriction of the hyperfluorescent parafoveal ring toward the end of the observation period.19–21

Despite the short duration of our study, the mean annular radius constricted in three subjects. Change in the diameter of the hyperfluorescent ring has important implications as numerous groups have reported on the functional correlation of its constriction. Structurally, the peak AF within the ring is known to correspond to fragmentation of the IS/OS region on SDOCT and demarcates central preservation of this junction and loss peripherally.7,8 Correlates with the PERG, multifocal electroretinography, and psychophysical testing are also well described.5,10,12 Autofluorescence cannot replace these other measures but provides an additional objective adjunct in assessing disease progression and potential response to therapeutic intervention.

Retinal pigment epithelium atrophy, as illustrated by regions of decreased AF, was quantified by setting a percentile grading to 10% of the image’s overall fluorescence. Again, this number can be adjusted as required. For instance, a lower threshold could be set to isolate the central homogenous region of decreased AF from the peripheral granular decreased AF that accompanies certain Stargardt disease phenotypes. Our patients all demonstrated reproducible degrees of atrophy throughout the 12-month period, with marked progression in participant E between 0 and 6 months.

There are a number of limitations with our methodology. There could theoretically be substantial bias if the detector gain were altered, although the impact would be greater on absolute luminance values than a percentile grading system. Mapping of the AF ring is facilitated by a clearly delineated change in AF, which is not present in all cases of RP. Furthermore, as the ring approaches the fovea in more advanced disease, it becomes less defined, which could complicate its capture.

Our study describes a straightforward method for serially analyzing AF images in subjects with RP by utilizing a percentile grading system. It can be applied retrospectively to AF images in a clinical setting and adjusted accordingly for quantifying the hyperfluorescent annulus and peripheral hypofluorescence characteristic of many patients with RP. The technique is most informative in subjects with earlier disease as demarcation arcs retain greater definition. Further work will elucidate how quantifiable progression of AF patterns correlates with or predicts subsequent functional deterioration. Moreover, there is potential for application of this technique in other retinal dystrophies and other imaging modalities.

### Acknowledgments

With thanks to Jon Brett, medical photographer, for providing the additional serial image sets for further analysis.

Supported by the Oxford Biomedical Research Centre based at Oxford University Hospitals NHS Trust, the University of Oxford, and Okuvision, with research infrastructure support from the National Institute for Health Research (NIHR) as a Clinical Research Network Portfolio-adopted study.

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. The sponsor and funding organization had no role in the design or conduct of this research.

Disclosure: J.K. Jolly, R; S.K. Wagner, None; J. Moules, S; F. Gekeler, Okuvision GmbH (P); A.R.
References


