Quantification of Fundus Autofluorescence to Detect Disease Severity in Nonexudative Age-Related Macular Degeneration

Ira H. Schachar, MD, MSc; Sarwar Zahid, MD, MS; Grant M. Comer, MD, MS; Maxwell Stem, MD; Asa G. Schachar, BS; Stephen J. Saxe, MD; Thomas W. Gardner, MD, MS; Victor M. Elner, MD, PhD; Thiran Jayasundera, MD

IMPORTANCE Accurate assessment of disease burden and determination of disease progression are challenging in dry age-related macular degeneration (AMD). We assessed the utility of quantified fundus autofluorescence (FAF) in the evaluation and follow-up of dry AMD.

OBJECTIVE To develop a method for quantitative FAF image analysis that is capable of stratifying severity of nonexudative AMD.

DESIGN, SETTING, AND PARTICIPANTS A retrospective analysis from 2008 to 2012 at a university eye center of FAF images taken of normal and nonexudative AMD eyes compared the Index of Retinal Autofluorescence (IRA) with retinal specialists’ clinical rankings of FAF images and the Age-Related Eye Disease Study (AREDS) grading scheme of corresponding color fundus photographs.

INTERVENTION Digital files of Heidelberg Spectralis FAF images taken of normal and nonexudative AMD eyes were analyzed. For each image, a unique horizontally oriented FAF signature composed of vertically averaged gray-scale values was generated through the fovea. A pairwise comparison of 2 signatures was performed using a modified difference of squares method, which generated a single quantitative value, the IRA.

MAIN OUTCOMES AND MEASURES The effects of intersession testing, cataract extraction, pupillary dilation, focal plane, and gain settings on the IRA were assessed.

RESULTS The FAF images taken of the same subjects at different times demonstrated low intersession variability of the IRA (intraclass coefficient = 0.75; 95% CI, 0.45-0.92). The IRA was affected by cataract severity, cataract extraction, small pupillary diameters (<5.5 mm), defocusing, and excessive high or low camera gain. The IRA values correlated with both subjective clinical rankings by retinal specialists ($r_s = 0.77$). The IRA was positively correlated with AREDS score ($r_s = 0.73$) and could statistically distinguish AREDS grades 3 and 4 ($P < .001$). Serial imaging demonstrated the utility of the method for identifying clinically meaningful disease progression.

CONCLUSIONS AND RELEVANCE The IRA method applied to FAF digital files can quantify AMD disease severity and may be helpful in identifying AMD progression.

Published online June 20, 2013.

Related article page 1004

Author Affiliations: Department of Ophthalmology and Visual Sciences, Kellogg Eye Center, University of Michigan, Ann Arbor, Michigan.

Corresponding Author: Victor M. Elner, MD, PhD, Kellogg Eye Center, University of Michigan, 1000 Wall St, Ann Arbor, MI 48105 (velner@umich.edu).
Approximately 90% of the 8 million people in the United States with age-related macular degeneration (AMD) are affected by the dry form of the disease. Pathogenesis of AMD involves progressive dysfunction and degeneration of the retinal pigment epithelium (RPE) with the accumulation of lipofuscin and sub-RPE drusen deposition. Dry AMD progression may result in gradual deterioration in visual function, but clinical quantification of the pathological alterations remains challenging.

Fundus autofluorescence (FAF) imaging provides an opportunity for objective quantification of alterations caused by AMD at the level of the RPE. Increased RPE lipofuscin accumulation results in hyperfluorescence, while atrophic areas with geographic atrophy (GA) exhibit hypofluorescence. To date, assessment of FAF images has been subjective and dependent on recognition of qualitative disease alterations. For example, Delori et al reported the FAF appearance of different types of drusen, while Holz et al have described various patterns of fluorescence that may surround areas of GA.

Quantification of FAF photographs by measuring regions of GA hypofluorescence using contour lines and monitoring their progression has been described in patients with late-stage AMD. Bearley et al described the prognostic utility of quantifying the area of hyperfluorescence in areas bordering GA. However, objective quantitative analysis of FAF images has only been described for late-stage AMD. Application of FAF quantification to earlier stages of AMD when alterations are subtle is complicated by variability in fundus image acquisition factors including defocusing, gain setting, photograph bleaching, media opacity, and pupillary dilation.

In a recent publication, Delori et al described compensation for variability in gain due to excitation intensity by using an internal fluorescent reference in a customized scanning laser ophthalmoscope. Unfortunately, these modifications are not applicable to existing databases of FAF images.

To address the need for objective quantification of dry AMD and its progression, we have developed a technique that analyzes and compares fluorescence intensity of macular FAF images to derive a single value for each image, the Index of Retinal Autofluorescence (IRA). This index was used to quantify dry AMD severity and its progression.

Methods

The study was approved by the University of Michigan institutional review board. All FAF images were obtained by trained photographers using the Heidelberg Spectralis autofluorescence digital imager (Heidelberg Engineering) (excitation, 488 nm; emission, 500 nm–700 nm). All images were obtained from the University of Michigan electronic medical record system. Subjects with optic nerve or retinal diseases other than dry AMD were excluded. All subjects were white of European descent.

FAF Analysis

A 3°, 768 × 768-pixel, 8-bit FAF image centered on the fovea. These images were exported directly from the Heidelberg Spectralis as XML files. Each XML file contained 8-bit gray-scale intensity (gsi) values on a 256-point scale from 0 (absent fluorescence, black) to 255 (maximum fluorescence, white). All files were analyzed using an automated analysis program. A FAF signature was generated for each 8-bit image by calculating vertically averaged gray-scale values extracted from a 400 (horizontal) × 50 (vertical)-pixel area centered on the fovea (Figure 1).

Comparison of 2 superimposed FAF signatures was performed while eliminating image acquisition variables as follows: the IRA was calculated as described earlier where $a(p)$
and $b(p)$ are vertically averaged gray-scale values at point $p$ on their respective FAF signatures; value is divided by a constant $k$ for scaling; and $x$ and $y$ represent curve shift along the $x$- and $y$-axes, respectively. In FAF images containing GA, the minimum averaged gray-scale value in the FAF signature was set at or below the minimum pixel value in the comparison FAF signature.

The IRA is a comparative index that simultaneously quantifies changes in both hyperfluorescence and hypofluorescence between 2 FAF signatures (Figure 2). The IRA can be used to compare FAF signatures from a subject with AMD with a control subject to stratify disease severity, or it can be used to compare FAF signatures from a single subject with serial imaging to reflect changes in disease over time.

Subject and Image Selection

1. Ten subjects (60-85 years of age) with normal vision, no significant media opacity, and no evidence of dry AMD or other retinal disease on clinical examination were recruited as control subjects. Color fundus photography and FAF images of these subjects were used to confirm the absence of retinal pathology. An FAF signature for each eye of each subject was created from the FAF image. By averaging the ipsilateral FAF signatures of the 10 subjects, right eye and left eye composite FAF control signatures were generated.

2. Fifteen subjects with dry AMD were enrolled for serial FAF imaging (3 sessions) over a period of 1 month. To assess IRA reproducibility for each subject, images from different sessions were compared to determine IRA intersession variability. The 15 groups of subject FAF images were then compared with the appropriate composite ipsilateral control FAF signature to determine intraclass correlation.

3. Seven subjects undergoing cataract surgery with implantation of an Alcon SN60WF intraocular lens were imaged before and 1 week after surgery to determine the effect of cataract extraction on IRA reproducibility.

4. To determine the effect of pupillary dilation on IRA, serial FAF images were taken of a single normal subject over the course of 30 minutes as the pupil dilated. To determine the effect of camera gain and plane of focus, FAF images were taken of a single normal subject at different gains, as measured in total sensitivity (range, 44-107), and different planes of focus.

5. Color images and FAF images of 1 eye from each of 39 subjects with dry AMD were analyzed. For each eye, the color photograph was graded by a masked, trained reader and confirmed by a retina specialist (T.J.), using the Age-Related Eye Disease Study (AREDS) 4-point grading scheme. The AREDS grade was then compared with the IRA value of the same eye. A Kruskal-Wallis test was used to assess for significant differences of IRA values across the AREDS grades of the photographs. A 2-tailed Mann-Whitney $U$ test was used for intergroup comparisons with a Bonferroni correction ($\alpha = .01$). A Spearman rank correlation coefficient was used to assess the association between AREDS grade and IRA value.

6. To determine whether IRA values correlate with expert subject grading of FAF images, FAF images of 55 eyes with dry AMD were analyzed by severity by 3 retinal specialists (T.J., T.W.G., and G.M.C.). Because there are no established guidelines for FAF grading in dry AMD, the specialists were instructed to rank the severity of FAF alterations in each image by observing the areas of hyperfluorescence and hypofluorescence in the macula as would routinely be done in the clinical counseling of patients with the disease. A Spearman rank correlation coefficient was used to assess the association between retinal specialist ranking and IRA value.

Results

Control Subjects

The 10 control subjects exhibited similar FAF signature characteristics. A deep central depression was located at the fovea with a sharp rise in gray-scale values in the perifoveal region. Moving into the parafoveal region, the FAF signature exhibited a more tapered rise in gray-scale values, which con-
continued in the nasal and temporal macula. These conserved features allowed for averaging of the 10 control FAF signatures to generate a composite control FAF signature that exhibited these key characteristics (Figure 2B [average control curve]).

Intersession Variability
Of the 45 FAF images (3 serial images for each of 15 subjects), 4 FAF images of 4 subjects with poor resolution and 1 exhibiting excessively high gain were excluded. Another 3 FAF images from 1 subject with excessive media opacity were excluded. The remaining FAF images of 14 subjects had a median intersession IRA of 8.1 gšt² ± 3.9. The intraclass correlation coefficient of the 14 subjects was 0.75 (95% CI, 0.45-0.92). These results identified the expected degree of intersession variation in the IRA in the absence of AMD disease progression.

Effect of Cataract Extraction
To determine the effect of cataract extraction on the IRA, 7 subjects underwent imaging before and 1 week after cataract extraction. Four subjects had moderate nuclear sclerosis of the lens while 2 had severe nuclear sclerosis and 1 had severe nuclear sclerotic and posterior subcapsular cataract. Cataract extraction with Alcon SN60WF intraocular lens insertion had a variable effect on the FAF signature. In the 4 subjects with moderate nuclear sclerotic cataracts, modest postoperative increases in median IRA were observed (IRA = 18.4 gšt²). All 3 subjects with severe nuclear sclerotic cataracts had images of poor resolution and could not be evaluated. Changes in lens status (significant cataract progression or cataract extraction) can cause elevations in IRA, which can mimic disease progression.

Effect of Pupillary Diameter
Pupil diameter has been reported to affect FAF images. Serial photographs of the same control subject during dilation revealed that once the pupil diameter reached 5.5 mm, serial photographs had a median IRA of 2.6 gšt² ± 0.9, which was similar to that obtained with intersession testing. However, serial photographs at pupillary diameters less than 5.5 mm exhibited much greater image-to-image variability, with a median IRA of 19.2 gšt² ± 7.9. This variability was further increased in the undilated state. Image consistency is best achieved with pupil diameters at or more than 5.5 mm.

Effect of Focal Plane/Image Contrast
Adequate focus to obtain images showing resolution of vessels measuring 15 to 25 μm (A-C) and gain settings with total sensitivities ranging from 67 (F) to 100 (G) did not significantly affect IRA reproducibility. Defocusing (D) and excessive gain (E and H) adversely affected IRA reproducibility.

Figure 3. Effect of Defocusing and Gain Setting on Index of Retinal Autofluorescence (IRA)

Focusing adequate to obtain retinal vessels measuring 15 to 25 μm (A-C) and gain settings with total sensitivities ranging from 67 (F) to 100 (G) did not significantly affect IRA reproducibility. Defocusing (D) and excessive gain (E and H) adversely affected IRA reproducibility.
Detecting Disease Progression

Three subjects underwent repeated FAF imaging over periods ranging from 9 months to 36 months. One subject, who showed no clinical progression based on color fundus photographs, also exhibited virtually no change in IRA values over a 9-month period. A second subject, who showed mild clinical progression with slow expansion of RPE atrophy, exhibited IRA values that modestly increased in both eyes over a 4-year period (Figure 4B). A third subject, who experienced rapid clinical deterioration manifested as large increases in RPE atrophy in only 9 months, exhibited substantial increases in the IRA values (Figure 4C).

Correlation of IRA With Retinal Specialist Grading

Fifty-five FAF images were ordered according to their IRA values, and the same images were ordered by subjective assessment of disease severity by 3 retinal specialists. Linear correlation of disease severity among retinal specialists was observed ($r_s = 0.92$ and $0.93$) (pairwise comparisons). A similar linear correlation was found between subjective ordering of fundus autofluorescence images and ordering using their respective IRA values ($r_s = 0.77$, Spearman rank coefficient) (Figure 5).

Correlation of IRA With AREDS Color Photography Grading and FAF Grading in Dry AMD

When comparing IRA values with AREDS grades based on standardized color fundus photographs, IRA values were positively correlated with AREDS grading ($r_s = 0.73$): median IRA for controls ($n = 10$) was $9.8 \, \text{gsi}^2$, for grade 2 ($n = 8$) was $22 \, \text{gsi}^2$, and...
Different than controls (pairwise comparisons, all Spearman rank coefficient). All AREDS grades are individually significantly different than controls (pairwise comparisons, all \( P < .01 \)). The largest increase in median IRA occurs between AREDS grades 3 and 4 (\( P < .001 \)).

Discussion

We report a novel technique of image analysis for quantification of conventional FAF images. The IRA quantifies heterogeneity of autofluorescence, as opposed to absolute intensity differences, as a key feature of AMD progression. The technique permits assessment of disease severity in any single FAF image or progression of disease by comparing sequential images of the same individual.

Instead of comparing FAF images on a pixel by pixel basis, the method produces an FAF signature consisting of a series of 400 points, each of which represents an average of 50 vertically oriented pixels, to permit assessment of horizontal FAF variability across a band through the fovea. This single band alone was intentionally chosen to avoid introduction of heterogeneity due to the retinal vasculature. To conclude that the IRA derived from the comparison of 2 FAF signatures reflects actual disease status requires a number of key assumptions. First, because AMD is a progressive degenerative condition, any change in the shape of the FAF signature, whether due to hypofluorescence or hyperfluorescence, represents disease progression. Importantly, inconsistencies in image acquisition will shift FAF signature baseline intensity but do not significantly affect FAF signature shape. In contrast, disease severity or progression alters FAF signature shape and can be quantified by the IRA. These assumptions provide a direct theoretical basis by which FAF images may be compared quantitatively, while avoiding the need for identifying reference points to calculate and compare absolute FAF intensity. However, comparison with age-appropriate, ethnically similar controls is necessary to assess disease severity. When comparing individuals, macular pigments may vary even in the absence of disease. Therefore, such controls are necessary to limit this source of variation. The FAF signatures of these 10 age-appropriate control subjects were averaged to obtain a control FAF signature with which the individual FAF signatures of the AMD subjects were compared. We assumed that this comparison would yield a disease-associated IRA that compared favorably with clinical assessment of AMD disease severity.

We found a number of practical considerations that might adversely affect the quality of FAF images and prevent effective quantitative analysis. Pupil diameters less than 5.5 mm were found to induce variations in IRA greater than that expected by baseline intersession testing.\(^{18}\) Severe cataract resulted in degradation of FAF images, preventing their evaluation, while cataract extraction had a variable effect on the underlying FAF signature dependent on both the type and density of the cataract, which may affect incident and emitted fluorescence. Thus, IRA values of subjects exhibiting cataract progression or undergoing cataract extraction must be interpreted with caution. Photographic parameters, such as large ranges in camera gain or significant alterations in the plane of focus, can also affect the IRA (Figure 3). All of these limitations indicate that the IRA is dependent on adequate quality and consistent FAF imaging.

We also found that our analysis may be used for longitudinal quantification of changes in FAF images in subjects with AMD. During serial imaging, IRA values larger than what would be accounted for by intersession testing (namely \( 8.1 \text{ gsi}^2 \pm 3.9 \)) were considered to represent disease-associated increases. Thus, we found that the subject who had no progression of AMD had comparable IRA values in the same eye over a 9-month period. In contrast, a subject with slowly progressive GA exhibited increased IRA during sequential testing over a 2-year interval. Finally, a subject with rapidly progressive GA showed large progressive IRA increases (Figure 4). In all 3 cases, the IRA values compared well with clinical assessment of their disease progression.

To assess the utility of this new method of FAF quantification, a comparison with established AMD grading criteria is necessary. Unfortunately, there are no widely accepted criteria for grading alterations in hypofluorescence and hyperfluorescence of FAF images.\(^6,9\) The AREDS grading schemes\(^2,17\) are based on color fundus photography and are heavily influenced by drusen, which are largely obscured in FAF images by pigment in the overlying RPE, and existing semiautomated techniques of FAF quantification, such as the convex hull method, are only applicable to subjects with GA.\(^14,15\) We found the IRA to be robust when compared with subjective retinal specialist grading of FAF images for hypofluorescent and hyperfluorescent changes. This indicates that our method accurately quantifies the expert qualitative assessments on which the clinical severity in FAF imaging is based. Furthermore, we found a strong positive correlation between IRA values and AREDS grade. In this small cohort, we were able to detect sta-
tistically significant differences from controls and among AREDS grades. The applicability of our technique to existing FAF data sets is underscored by the fact that the images in this cohort, taken by different photographers, were analyzed retrospectively. For GA, future studies comparing our fully automated measure with existing semiautomated measures may reinforce the validity of all these methods for quantifying advanced AMD.\(^{14,15}\)

Our method to quantify FAF images attempts to address limitations that are inherent in clinical imaging of patients. Thus, it incorporates what would be considered limitations in study design, namely image acquisition standardization. Such standardization would include same imager, same photographer, precise focus, consistent gain, identical pupil size, and same lens status. However, such standardization would also render the technique less applicable to clinical practice and FAF research data sets from various clinical settings. Thus, we embrace the theoretical limitations of our study to obtain a practical tool to quantify FAF changes in AMD. This novel technique or a modification thereof may set new standards for identifying and quantifying AMD progression at an earlier stage and better assess the clinical effect of interventions aimed at slowing the progression of this blinding disease.

**REFERENCES**