Identifying “Preperimetric” Glaucoma in Standard Automated Perimetry Visual Fields

Ryo Asaoka,1 Aiko Iwase,2 Kazunori Hirasawa,1,3 Hiroshi Murata,1 and Makoto Araie4

1Department of Ophthalmology, The University of Tokyo Graduate School of Medicine, Tokyo, Japan
2Tajimi Iwase Eye Clinic, Tajimi, Japan
3Department of Ophthalmology, Graduate School of Medical Science, Kitasato University, Kanagawa, Japan
4Kanto Central Hospital of the Mutual Aid Association of Public School Teachers, Tokyo, Japan

Purpose. To compare the visual fields (VFs) of preperimetric open angle glaucoma (OAG) patients (preperimetric glaucoma VFs, PPGVFs) with the VFs of healthy eyes, and to discriminate these two groups by using the Random Forests machine-learning method.

Methods. All VFs before a first diagnosis of manifest glaucoma (Anderson-Patella’s criteria) were classified as PPGVFs. Series of VFs were obtained with the Humphrey Field Analyzer 30-2 program from 171 PPGVFs from 53 eyes in 51 OAG or OAG suspect patients and 108 healthy eyes of 87 normal subjects. The area under the receiver operating characteristic curve (AROC) in discriminating between PPGVFs and healthy VFs was calculated by using the Random Forests method, with 52 total deviation (TD) values, mean deviation (MD), and pattern standard deviation (PSD) as predictors.

Results. There was a significant difference in MD between healthy VFs and PPGVFs (−0.03 ± 1.11 and −0.91 ± 1.56 dB [mean ± standard deviation], respectively; P < 0.001, linear mixed model) and in PSD (1.56 ± 0.33 and 1.97 ± 0.43 dB, respectively; P < 0.001). A significant difference was observed in the TD values between healthy VFs and PPGVFs at 25 (P < 0.001) of 52 test points (linear mixed model). The AROC obtained by using the Random Forests method was 79.0% (95% confidence interval, 73.5%–84.5%).

Conclusions. Differences exist between healthy VFs and VFs of preperimetric glaucoma eyes, which go on to develop manifest glaucoma; these two groups of VFs could be well distinguished by using the Random Forests classifier.

Keywords: glaucoma, preperimetric stage, Random Forests method, visual field

Prompt diagnosis of glaucoma is critical to reduce the risk of visual impairment, as glaucomatous visual field (VF) damage is irreparable. There is evidence that glaucomatous axonal loss precedes the development of VF defects. In fact, it is now possible to detect early glaucomatous structural changes before the deterioration of patients’ VFs becomes measurable, owing to the development of imaging devices, such as optical coherence tomography and scanning laser polarimetry. Furthermore, several studies have suggested that some forms of perimetry, such as frequency-doubling perimetry (FDT), may detect glaucomatous VF changes earlier than conventional standard automated perimetry (SAP).

In clinical practice, we often encounter eyes with structural changes suggestive of glaucoma, but with apparently normal VF test results. The term “preperimetric glaucoma” (PPG) has been used to describe eyes with a glaucomatous optic disc and/or fundus appearance and an apparently normal VF. Several methods, such as abnormal result from the glaucoma hemifield test (GHT) or a significant pattern standard deviation (PSD) value, are used to define glaucomatous VFs but Anderson-Patella’s criteria remain one of the most frequently applied methods. Briefly, the criteria are as follows: (1) a cluster of ≥3 points in the pattern deviation plot in a single hemifield (superior/inferior) with P < 0.05, one of which must have been P < 0.01, excluding the outermost test point of Humphrey Field Analyzer 30-2 program (HFA 30-2); (2) GHT result outside of normal limits; or (3) abnormal PSD with P < 0.05. Thus, the criteria are designed to be sensitive to focal VF deterioration as well as global VF changes, which is advantageous to detect early glaucomatous VF damage. Preperimetric glaucoma VFs (PPGVFs) can be defined as VFs taken before diagnosis of manifest glaucoma by Anderson-Patella’s criteria.

In this study, a machine-learning classifier known as the Random Forests method was applied to differentiate PPGVFs from healthy VFs. In short, the Random Forests consists of many decision trees (each tree developed by bootstrapping of data and predictor variables) and outputs a classification based on all the decision trees. This algorithm has been used in many research fields, such as gene selection and cancer classification. Furthermore, its potential for diagnosing glaucoma has been reported. The merit of using this method is that it can cope with intercorrelation of multiple explanatory variables, which makes it especially useful in analyzing VF test results where intercorrelation of VF sensitivities among test points is well known. Additionally, the Random Forests method...
first VF of each eye was excluded,37,38 PPGVFs were identified
589 OAG or OAG suspect patients were identified. After the
Patients aged 20 years or older and eyes with refractive error
surgeries except for uneventful intraocular lens implantation.
cataract except for clinically insignificant senile cataract,
angle glaucoma was diagnosed when the following findings
between April 1997 and October 2006. Inclusion criteria
the consecutive VF test).22 As a result, 171 PPGVFs from 53
 subjects were identified. The study included no abnormal eye-related findings except for clinically
significant ONH changes fulfilling Anderson-Patella’s criteria22 on two consecutive occasions; and
compatible with the glaucomatous ONH changes fulfilling
edge at the ONH margin greater than a major retinal vessel,
diverging in an arcuate or wedge shape confirmed by a panel
of glaucoma specialists (HM and RA) after inspection of stereo-
and in refractive error (unpaired t-test for numerical data and \( \chi^2 \) test for categorical data).
interprets multiple explanatory variables in combination, deciphering important patterns,36 which may be advantageous in trying to detect early glaucomatous VF changes that have not yet fulfilled Anderson-Patella’s criteria.22

METHODS

Subjects
Preperimetric glaucoma VFs were retrospectively acquired from clinical records at the University of Tokyo Hospital (Tokyo, Japan). First, all open angle glaucoma (OAG) or OAG suspect patients who had VF measurements at least 15 times between April 1997 and April 2012 were identified. Open angle glaucoma was diagnosed when the following findings were present: (1) presence of typical glaucomatous changes in the optic nerve head (ONH), such as a rim notch with a rim width \( \leq 0.1 \) disc diameters or a vertical cup-to-disc ratio of \( >0.7 \), and/or a retinal nerve fiber layer (RNFL) defect with its edge at the ONH margin greater than a major retinal vessel, diverging in an arcuate or wedge shape confirmed by a panel of glaucoma specialists (HM and RA) after inspection of stereofundus photographs; (2) presence of glaucomatous VF defects compatible with the glaucomatous ONH changes fulfilling Anderson-Patella’s criteria22 on two consecutive occasions; and (3) absence of other systemic or ocular disorders, including cataract except for clinically insignificant senile cataract, shallow peripheral anterior chamber that could affect the ONH and VF, and history of intraocular surgeries or refractive surgeries except for uneventful intraocular lens implantation. Patients aged 20 years or older and eyes with refractive error \( \geq -6.0 \) D and \( <6.0 \) D, and also visual acuity equal to or better than 30/30 were included. As a result, VFs from 790 eyes of 589 OAG or OAG suspect patients were identified. After the first VF of each eye was excluded,37,38 PPGVFs were identified as all VFs before the first VF meeting with at least one of Anderson-Patella’s criteria (without a need to be confirmed in the consecutive VF test).22 As a result, 171 PPGVFs from 53 eyes of 51 patients were used in the analysis.

Data from 108 normal eyes of 87 subjects were acquired at the University of Tokyo Hospital (Tokyo, Japan), and Tajimi Iwase Eye Clinic (Gifu, Japan) between April 1997 and October 2006. Inclusion criteria included no abnormal eye-related findings except for clinically insignificant senile cataract on biomicroscopy, gonioscopy, and funduscopy, and no history of ocular diseases, such as diabetic retinopathy or age-related macular degeneration; other inclusion criteria were age \( \geq 20 \) years, spherical equivalent refractive error \( \geq -6.0 \) D and \( <6.0 \) D, intraocular pressure \( \leq 21 \) mm Hg by Goldmann applanation tonometry, having undergone HFA measurement at least one time, and normal VF test results that did not meet any of Anderson-Patella’s criteria and no family history of glaucoma.22 The study protocol was approved by the institutional review board of each institution and adhered to the tenets of the Declaration of Helsinki. Written informed consent was obtained from each subject.

Visual Field Measurements
All VFs were obtained by using the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA, USA) with the 30-2 SITA Standard strategy (HFA 30-2) and the Goldmann size III target. Near refractive correction was used as necessary. Reliable VFs were defined as results with less than 20% fixation losses and less than 15% false-positive error, following the criteria used by the HFA software.

Statistical Analyses
The Random Forests method was used to classify VFs as PPGVFs or healthy VFs, by using the 52 TD values that correspond to the HFA 24-2 program test points, MD, and PSD as predictors; thus, TD values of the outermost 22 test points in the 30-2 test pattern were excluded. In the Random Forests method, 10,000 decision trees were grown to build an “ensemble classifier.” The leave-one-out cross-validation method was then performed and the area under the receiver operating characteristic curve (AROC) was calculated by changing the cutoff value for the probability of a VF being classified as a PPGVF (as indicated by the proportion of tree votes in the Random Forests classifier).39 In leave-one-out cross-validation, one or both eyes of a single subject were used as validation data, and the remaining subjects were used as training data. This procedure was then repeated until each OAG patient and healthy subject in the original sample was used once as validation data. In other words, for each individual, only the data from all other subjects was used in the prediction.

All statistical analyses were carried out by using the statistical programming language R (version 2.15.1; The R Foundation for Statistical Computing, Vienna, Austria). The R package “randomForest” was used to generate the Random Forests classifier. The importance VF parameters used in the Random Forests method was determined by using the Random Forests “Variable Importance” (VIMP) measure; this was calculated by randomly permuting a variable at each decision tree and measuring whether the squared errors decreased.26 The optimum cutoff point in the receiver operating characteristic curve was calculated by using Youden’s method.40 The difference between the two groups was investigated by using a linear mixed-effects model when two or more measurement results in the same eye and same subject were included, with the R package “lme4”; this model inherently accounts for imbalances in the numbers of VFs per eye and of a subject.

RESULTS

Demographic data of the subjects are given in the Table. The mean interval between patients’ last PPGVF to the time of conversion to the first glaucomatous VF (i.e., Anderson-Patella’s criteria22 met on two consecutive occasions) averaged 315.8 ± 184.3 (standard deviation, SD) days. The average interval between consecutive PPGVFs (i.e., until Anderson-Patella’s criteria were met at any occasion without the need of being repeated in the consecutive VF) was 271.1 ± 188.1 days.

There were no intergroup differences in right eye/left eye, male/female ratios (\( P = 0.78 \) and 0.55, respectively; \( \chi^2 \) test), and in refractive error (\( P = 0.21 \), unpaired t-test), except for age (\( P = 0.008 \)). The average number of PPGVFs per eye was 3.2 ± 2.0 (mean ± SD).

### Table. Characteristics of the Study Participants

<table>
<thead>
<tr>
<th></th>
<th>Normal Group</th>
<th>Glaucoma Group</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, male/female</td>
<td>35/52</td>
<td>26/25</td>
<td>0.09</td>
</tr>
<tr>
<td>Eye, right/left</td>
<td>51/57</td>
<td>27/26</td>
<td>0.78</td>
</tr>
<tr>
<td>Age, mean ± SD, y</td>
<td>50 ± 16</td>
<td>55 ± 15</td>
<td>0.008</td>
</tr>
<tr>
<td>Refractive error, diopters, mean ± SD</td>
<td>-0.15 ± 1.1</td>
<td>-0.55 ± 0.33</td>
<td>0.21</td>
</tr>
</tbody>
</table>

\( P \) value: comparison between glaucoma and normal groups (unpaired t-test for numerical data and \( \chi^2 \) test for categorical data).
Figures 1a and 1b show the mean TD value at each location across all healthy VFs and all PPGVFs, respectively. A significant difference was observed in TD values between healthy VFs (Fig. 1a) and PPGVFs (Fig. 1b) at 25 of 52 test locations with a significance level of $P < 0.001$ (linear mixed model); none of the mean TD values in the healthy VFs were lower than $1.0$ dB, whereas in the PPGVFs 22 of 52 test points were lower than this threshold. There was a significant intergroup difference in the MDs ($0.03$ $1.11$ and $0.91$ $1.56$ dB in healthy VFs and in PPGVFs, respectively; $P < 0.001$, linear mixed model) and PSDs ($1.56$ $0.33$ and $1.97$ $0.43$ dB, respectively; $P < 0.001$; Figs. 2a, 2b). Only $2.8\%$ (3 VFs) of the healthy VFs, as compared with $12.3\%$ (21 VFs) of PPGVFs, had a PSD above $2.5$ dB. Pattern standard deviation was above $2.0$ dB in $10.2\%$ (11 VFs) of the healthy VFs and $39.2\%$ (67 VFs) of PPGVFs. None of the healthy VFs with a PSD above $2.0$ dB had GHT results ‘‘outside normal limits’’; moreover, the deterioration of VF thresholds appeared simply to be measurement noise. Glaucoma hemifield test results were ‘‘borderline’’ in $16.7\%$ (18) of healthy VFs and in $8.2\%$ (14) of PPGVFs. There was a ‘‘general reduction of sensitivity’’ in GHT result in two PPGVFs (1.1%). Two (1.9%) healthy VFs and 10 (5.8%) PPGVFs had a PSD with a $P$ value between 0.05 and 0.10.

As shown in Figure 3, the AROC obtained by using the Random Forests method was 79.0% (confidence interval [CI], 73.5%–84.5%). Sensitivities at specificities of 80%, 90%, and 95% were 59.6%, 40.4%, and 25.1%, respectively. The optimum discrimination, using Youden’s method, was obtained when sensitivity was equal to 86.5% and specificity was equal to 62.0%; this corresponded to a voting rate of $>53.8\%$ in the Random Forests classifier. The 15 most important parameters (according to VIMP score) in the Random Forest system are depicted in Figure 4.

**DISCUSSION**

In the current study, series of VFs of OAG patients were examined and the properties of VFs before an eye’s conversion to glaucomatous VF, following Anderson-Patella’s criteria, were investigated by using the Random Forests classifier. Significant deterioration of TD values was observed at 25 ($P < 0.001$) of 52 test points in the PPGVFs; in addition, a significant difference in MD and PSD was also observed between PPGVFs and healthy VFs. Furthermore, a sizeable number of PPGVFs could be discriminated from healthy VFs by using the Random Forests machine-learning classifier.

Previous attempts have been made to discriminate PPG from healthy eyes by using various devices and special forms of perimetry, in a cross-sectional manner. Choi et al. have used MD and PSD of the Humphrey Matrix 24-2 test and obtained an

**FIGURE 1.** Mean and standard deviation of the total deviation values of healthy and preperimetric glaucoma visual fields (PPGVFs). The number in brackets below the mean value represents the standard deviation. (a) Normal eyes. (b) Preperimetric glaucoma eyes. *Significant difference between the healthy and PPGVFs with the significance level of $P < 0.001$ (written in red; linear mixed-effects model). TD values in the 22 outermost test points (gray) were excluded from analysis.

**FIGURE 2.** Boxplots comparing mean deviation and pattern standard deviation values in healthy visual fields and preperimetric glaucoma visual fields (PPGVFs). (a) Mean deviation. (b) Pattern standard deviation. $P$ values were calculated by using the linear mixed model.
AROC of 61.9% and 76.7% for each index, respectively; in their study PPG was diagnosed by using a combination of red-free fundus photography, Stratus optical coherence tomography (OCT), and stereophotography.\(^{17}\) Leeprechanon et al.\(^{16}\) have used FDT and SWAP and reported AROCs of 62% (FDT MD), 67% (FDT PSD), 74% (SWAP MD), and 53% (SWAP MD), in which PPG was defined by abnormal findings with either OCT or Heidelberg retina tomograph (HRT). In a study by Hirashima et al.,\(^{11}\) FDT is used to discriminate PPG defined as abnormal on stereophotography, and AROCs of 76.2% (MD) and 78.0% (PSD) are reported.

In addition, similar efforts have been made with imaging devices; the reported AROCs are 79% to 94% with OCT\(^{4–7,11}\) in which the diagnosis of PPG is made by using stereophotography\(^{5,7,11}\) and/or red-free photography,\(^{4,6}\) and 91% with SLP in which red-free photography, SLP, or OCT is used to diagnose PPG.\(^{8}\) The 95% CI of the AROC obtained in the current study (73.5%–84.5%) overlaps with those reported by FDT or SWAP and some of those reported with imaging devices; however, it is not appropriate to directly compare AROCs because of considerable differences in the study designs. In the current study, PPGVFs were categorized as the VFs undertaken before conversion to a glaucomatous VF according to Anderson-Patella’s criteria; PPGVFs were then discriminated from healthy normal VFs. On the other hand, PPG eyes, not PPGVFs, are distinguished from healthy normal eyes in the previous studies. In addition, the definition of PPG is not always the same across these studies, as detailed above. Moreover, the criteria for classifying glaucomatous VFs also vary. For instance, all conditions included in Anderson-Patella’s criteria are used in the studies of Hirashima et al.\(^{11}\) and Choi et al.\(^{17}\) while an abnormal result with GHT or PSD is used in the study by Leeprechanon et al.\(^{16}\) In the current study, an abnormal VF result was defined as meeting at least one of the three conditions included in Anderson-Patella’s criteria.\(^{22}\) There are only two studies that have analyzed the usefulness of OCT and HRT in diagnosing PPG with longitudinal series of VFs.\(^{5,7}\) In these studies, patients suspected of having glaucoma, based on the appearance of the optic disc on stereophotography, have been followed up longer than 10 years (VF remained normal throughout the follow-up); it was investigated whether HRT and OCT can discriminate between PPG eyes with stereophotographic evidence of progressive glaucomatous change and those without stereophotographic evidence of progressive glaucomatous change. As a result, the largest AROC is obtained with OCT temporal superior RNFL thickness (88%), followed by OCT global RNFL thickness (86%) and HRT rim area (72%)\(^{7}\) and average RNFL thickness (89%), followed by inferior hemisphere average RNFL thickness (87%) and inferior quadrant average RNFL thickness (85%).\(^{5}\)

The focus of the current study was to discriminate PPGVFs, defined as normal according to Anderson-Patella’s criteria, from PPG eyes with abnormal findings.
Visual Fields in Preperimetric Glaucoma

IOVS  December 2014  Vol. 55  No. 12  7818

VF parameters taken in healthy eyes. The PPGVFs were all taken from preperimetric OAG eyes that later went on to develop manifest glaucoma. Importantly, this study is unique, since any extra functional and/or structural measurements were not included; instead only SAP VFs were included. The Random Forests method was able to discriminate between PPGVFs and VFs from healthy eyes at least as well as FDT or SWAP can discriminate PPG eyes from normal eyes. This suggests that many PPGVFs have a retinal sensitivity distribution different from that observed in healthy eyes.

Pattern deviation (PD), TD values, MD, and PSD are displayed in VF test results alongside their associated significance levels; these are calculated by comparing the observed results against a normative database. Consequently, the diagnosis of glaucoma, based only on SAP measurements’ significance levels, is difficult in early glaucoma where there may be considerable overlap between the range of sensitivity values of normal eyes and early glaucomatous eyes. This phenomenon may be especially apparent with SAP in which all types of ganglion cells are stimulated.

Anderson-Patella’s criteria22 adopt PD values, instead of TD values, to diagnose glaucoma, since it is believed that these values help to eliminate the effects of cataract on the measured retinal sensitivity.44 Since eyes with cataract were excluded from both groups in the current study, using PD values in the Random Forests classifier instead of TD values is unlikely to offer any improvement; indeed, the use of PDs instead of TDs did not improve the AROC associated with the Random Forests method to discriminate PPGVFs and healthy VFs (AROC: 77.0% [CI: 70.9%-83.1%]).

We have shown that the Random Forests classifier can distinguish a considerable proportion of PPGVFs from healthy VFs (sensitivity was equal to 59.6%, 40.4%, and 25.1% at significance levels of 80%, 90%, and 95%, respectively). This level of discrimination could possibly be increased if SAP testing procedures were modified. For example, the 30-2 test grid of the HFA has test points located in 6-degree intervals, which is not ideal for detecting glaucomatous VF change because small, localized VF changes may be missed; instead, only less apparent deteriorated sensitivity is shown on VF if these occur in between test points. Indeed, recent research has shown that adoption of a HFA 10-2 test program reveals glaucomatous VF defects in 16% of eyes whose HFA 24-2 test results are normal.49

Interestingly, 12.5% of PPGVFs had a PSD value above +2.5 dB, whereas only 2.8% of the healthy VFs had a PSD above that value. There was a significant difference in the PSD values in 25 (P < 0.001) of 52 VF test points; however, the contribution of these test points to the discrimination was much less than that of PSD (Fig. 4). This is probably because the VF area primarily affected in glaucoma varies among patients and no one test point can serve as a universal landmark of glaucomatous VF deterioration. However, TD values in the superior nasal step and arcuate area close to the blind spot showed relatively strong contributions to the discrimination (Fig. 4), being compatible with the importance of Bjerrum and superior nasal step area as initial sites of glaucomatous damage. On the other hand, deterioration of PSD, representing un-uniformity of the damage at each test point, should be independent of the location of the initial VF deterioration. Thus, clinicians should be advised to treat patients carefully when PSD is above 2.5 dB, even if it is not deemed significant deterioration (i.e., P < 0.05). Agreeing with this result, the Ocular Hypertension Treatment Study has reported that PSD at baseline is the only VF parameter that significantly contributes to the future development of primary OAG.50

All of the analyses in the current study were carried out by using existing statistical software, specifically the language ”R”, which is an open source statistical program. In the current study, the leave-one-out cross-validation with the Random Forests method was performed, in which the original data set was split into validating (one subject) and training (all of other subjects) data sets, being repeated for the number of subjects in the original data set so that all of the subjects were used as validating data once. This is equivalent to predicting the diagnosis of a new patient (validating data set) by using the Random Forests diagnosis system trained by using the previous training data set in clinics. Thus, a Random Forests clinical support tool could be built simply on a personal computer; clinicians would then be able to diagnose a new patient’s VF with the sensitivity and specificity shown in Figure 3.

In the current study, only normal VFs from healthy eyes were included (abnormal VFs from healthy eyes were excluded from analyses), as the purpose of this study was to compare normal VFs against PPGVFs; however, in clinical settings, a substantial number of abnormal VFs are observed owing to high variability, but also as a result of various disorders, such as systemic disorders. Thus, a limitation of the current study is that the results cannot be directly applied to discriminate between normal and nonglaucomatous abnormal VFs.

A limitation of the current study is that the average age of the PPG group was significantly greater than that of the healthy group. This may have an influence on the PSD value; however, any impact is likely to be minimal because it is already corrected for age; indeed, we did not observe a significant relationship between age and PSD in the current population (P = 0.90, data not shown). A further limitation of the present study concerns the lack of external validation. Unfortunately, as far as we know, no previous studies have investigated PPGVFs longitudinally, but instead have collected PPGVFs in a cross-sectional manner with no confirmation that they convert to glaucomatous VFs. Thus, we have used leave-one-out cross-validation whereby, for each individual, only the data from other subjects were used in the model’s prediction. Thus, there is no overlap between the “training” and “validation” data sets. Nonetheless, it would be desirable to carry out a further study to validate the current results by using an external, independent data set.

In conclusion, we showed that considerable differences exist between SAP VFs of “preperimetric” glaucoma patients and the VFs of healthy subjects. Furthermore, eyes with such preperimetric glaucoma VFs can be discriminated from normal eyes, by using the Random Forests method with an AROC, at least as well as those reported for FDT or SWAP.11,16,17

Acknowledgments

Supported in part by Technology Agency (JST) CREST (RA and HM) and Grants 25861618 (HM) and 26462679 (RA) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Disclosure: R. Asaoka, None; A. Iwase, None; K. Hirasawa, None; H. Murata, None; M. Araie, None.

References


