ABSTRACT

**Purpose:** To assess the diagnostic usefulness of four linear discriminant functions for Heidelberg retina tomograph (HRT) II obtained in our population.

**Methods:** 450 eyes of 450 patients were studied. Eyes were classified as those of healthy subjects (101), those with ocular hypertension (126), those suspected to have glaucoma (121) or those with glaucoma (103). Intraocular pressure, optic nerve head appearance in stereophotographs, and standard automated perimetry results were assessed. ANOVAs among the groups were calculated for all global parameters and the 4 formulas. Receiver operating characteristic (ROC) curves were plotted for the 4 multivariate functions designed in our hospital and the areas under the ROC curves were compared. Sensitivity at 80% and 90% fixed specificities were also calculated.

**Results:** All functions discriminated well between healthy and glaucoma suspects, and between healthy and glaucomatous eyes. At 90% fixed specificity, sensitivities ranged from 54% to 60% in differentiating between healthy and glaucoma suspects, and from 65% to 68% in discriminating bet-

RESUMEN

**Objetivo:** Valorar la capacidad diagnóstica de cuatro fórmulas discriminantes diseñadas para el Heidelberg retina Tomograph (HRT) II a partir de nuestra población.

**Material y método:** Se incluyeron 450 ojos de 450 pacientes. Fueron clasificados según la presión intraocular, morfología papilar en estereofotografías y los resultados de la perimetria automatizada convencional en 101 ojos normales, 126 hipertensos oculares, 121 sospechosos de glaucoma y 102 ojos con glaucoma. Se calcularon las ANOVAs entre los grupos para todos los parámetros globales y las cuatro fórmulas. Se trazaron las curvas ROC para las cuatro funciones discriminantes diseñadas en nuestro hospital y se compararon las áreas bajo las curvas. Se calcularon los valores de sensibilidad para especificidades fijadas en el 80% y 90%.

**Resultados:** Las cuatro fórmulas discriminaron bien entre el grupo control y los sospechosos de glaucoma, y entre controles y glaucomas. Para una especificidad fijada en el 90%, las sensibilidades oscilaron entre el 54% y 60% para diferenciar entre sanos y sospechosos de glaucoma, y entre el 65% y...
ween control and glaucoma subjects. No differences were found between the areas under the ROC curves of these functions.

**Conclusions:** The evaluated linear discriminant functions increased the diagnostic ability of HRT II isolated parameters in detecting glaucoma. These functions performed better than the HRT-provided discriminant functions (Arch Soc Esp Oftalmol 2008; 83: 349-356).

**Key words:** Glaucoma, HRT, diagnosis, optic disc, discriminant function.

---

**INTRODUCTION**

Primary open-angle glaucoma is a multifactorial and progressive neuropathy characterized by the acquired loss of retinal ganglion cells and their axons. Together with the loss of nerve fibres typical changes commonly appear in the head of the optic nerve and defects in the visual field (1,2). The clinical assessment of the optic nerve head has been the absolutely necessary pillar for diagnosis and follow-ups of patients with glaucoma (3-8). However, stereophotographic or clinical assessment is subject to the experience and subjectivity of the observer. During the last few years these circumstances have conditioned the development of non-invasive techniques for studying the morphology of the optic nerve head in a quick, objective and reproducible manner, implying less dependency on the examiner.

The Heidelberg Retina Tomograph (HRT) is a confocal scanning laser that offers precise topographic maps of the optic disc and the peripapillary retina. It offers a series of global papillary morphometric parameters, and also for each of the six sectors in which it divides the papilla, based on a three-dimensional reconstruction. The HRT II has diagnostic classifications, such as Moorfields’ regression (9) and analysis tools for follow-ups.

Different studies (10,11) have sought to improve diagnostic ability of the device by combining the more heavily weighted stereometric parameters in discriminant equations. The linear discriminant analysis tools combine different scanning variables in function of their relative importance in producing a logarithmic regression equation (LRE). The HRT software includes two functions of this type: RB (10) and FSM (11). Bathija (12) et al. have proposed another LRE that the device software does not include but has shown to have a large ability to diagnose glaucoma (13). Other authors have also proposed different LRE for the HRT, but these are used less than the previous ones (1416).

The discriminant ability of the analytical tools is usually less than that described for the original populations. The equations are optimal for the population used in their calculation, but they cannot always be generalized for other populations. The aim of this study is to evaluate the diagnostic ability of the four LREs developed in our hospital (17), through multivariate analysis of the topographic parameters obtained with the HRT II in a population sample different from the one used for reaching the equations.

Different factors may alter the internal and external validity of a diagnostic precision study. Due to this the STARD (Standards for Reporting of Diagnostic Accuracy) (18) initiative was developed. The purpose of the STARD initiative is to improve the quality of diagnostic ability studies. The design of this study will be done according to the 25 points defined in the STARD initiative.

---

**SUBJECTS, MATERIALS AND METHODS**

**Subjects**

From January, 2005, to June, 2006, a total of 450 eyes were chosen prospectively and consecutively from the outpatient doctors’ visits at the Ophthalmology Service of the «Miguel Servet» University Hospital in Zaragoza.

The study design was accepted by the hospital’s Ethical Review Board. All participants signed the
informed consent form and the study methodology was performed according to the Declaration of Helsinki guidelines.

Each subject, regardless of his/her classification group, met a series of inclusion criteria, namely age between 30 and 75 years, visual acuity equal to or over 20/30 (Snellen scale), refraction defects under 5 spheric dioptres and astigmatism under 3 spheric-equivalent dioptres and a transparent optical media (nuclear color or opalescence, cortical or posterior subcapsular crystalline <1 according to the Lens Opacities Classification System III) (19).

Those individuals who had a history of eye surgery or serious trauma, systemic diseases with ophthalmic repercussions, the inability to perform or be tested on any of the tests included in exploratory protocol (perimeter test, HRT, etc.) or those that did not meet any of the inclusion criteria were excluded.

One eye per subject was taken into account for the study. The choice was made randomly, except when only one eye met the inclusion criteria.

**Exploratory procedure**

All subjects underwent a complete eye exam, which included biomicroscopy with a slit lamp, gonioscopy, base measurements of the intraocular pressure levels (IOP) using applanation tonometry (average of three measurements on different days without hypotensor treatment), central cornea pachymetry (DGH Technology, DGH 500 model), assessment of the fundus using indirect ophthalmoscopy and with a slit lamp (with a 78 dioptre Volk lens), assessment of the fundus using indirect ophthalmoscopy and with a slit lamp (with a 78 dioptre Volk lens), stereophotographs of the papilla, conventional automated perimetry (AP) and topography of the optic nerve head with the Heidelberg Retina Tomograph (HRT II).

The clinical evaluation of the optic nerve was performed with the slit lamp, indirect ophthalmoscopy and papillary stereographs by two glaucoma specialists unaware of the patient’s medical history. The discrepancies between the observers were resolved by consensus. A glaucoma-compatible optic nerve appearance was defined as the existence of neuroretinal rim thinning, focal or diffuse with an increase in the cup, the presence of notches, or both (20).

The APs were taken with the «Humphrey 745» Field Analyzer (24-2 SITA Standard Strategy). Perimeter defects were defined as the presence of a group of at least three altered points with a probability level under 5% or a group with at least two altered points with a probability level under 1% (those points found at the blind-spot poles were excluded) (21) and/or standard deviation from the mean (SDM) with a probability level under 5% and/or proof of the glaucomatous hemifield outside normal limits. At least two perimeties were performed to decrease the “learning” effect and if any of them did not meet the validity criteria defined by the same perimeter (false positives, false negatives and loss of focus), the test was repeated.

The papillary topographic study was done using a version II Heidelberg Retina Tomograph confocal scanning laser that allows a three-dimensional image to be obtained of the papilla from three series of 16 to 64 tomography slices done at different depths and centered on the head of the optic nerve. The HRT II exploration was conducted in a standard manner and by only one experienced examiner. For this the patients’ pupils were dilated with 1% tropicamide. Once the digital image of the optic disc is obtained it is necessary to mark its outline. The same planimetry expert, who was unaware of any medical history of the patient, marked the papilla outline following the inner edge of Elschnig’s scleral ring. From here, the HRT II software (Heidelberg Eye Explorer version 1.4.1.0) automatically set a reference plane at 50 microns under the retinal surface of the temporal sector found between 350º and 356º and calculates a series of global parameters and also sectoral parameters of the head of the optic nerve.

**Classification into groups**

The subjects included were classified into four study groups: normal or control, ocular hypertension (OHT), those suspected of having glaucoma and those with simple chronic glaucoma.

- Control group (101 eyes): Normal eye exam, IOP under 20 mmHg, normal morphology of the optic nerve and AP with no defects.
- OHT group (126 eyes): IOP above 21 mmHg, normal optic nerve and AP with no defects.
- Group of those suspected of having glaucoma (121 eyes): optic nerve compatible with glaucoma and AP with no defects.
- Glaucoma group (102 eyes): IOP over 21 mmHg and AP with reproducible glaucomatous defects.
In order to avoid the possible influence of the prior papillary morphometric information on the analysis of the results, subjects with glaucoma were chosen based not on the papillary clinical evaluation but on the AP result instead.

**Discriminant equations**

This study aims to analyze the diagnostic ability of the four LREs (17) developed for our population (Zaragoza) in a sample independent from that used for its design. These equations are:

**LRE 1**

\[
P(Si) = \frac{1}{1 + e^{(2.348-1.826 \cdot \text{Rim Area} + 0.781 \cdot \text{Cup Disc Ratio} + 8.607 \cdot \text{Cup Shape Measure})}}
\]

Variables included: rim area, cup/disc area ratio, and the cup shape measure.

**LRE 2**

\[
P(Si) = \frac{1}{1 + e^{(4.176 - 2.511 \cdot \text{Rim Area} + 4.310 \cdot \text{Cup Volume} + 9.432 \cdot \text{Cup Shape Measure})}}
\]

Variables included: rim area, cup volume, and the cup shape measure.

**LRE 3**

\[
P(Si) = \frac{1}{1 + e^{(2.157 - 4.339 \cdot \text{Rim Volume} + 1.358 \cdot \text{Cup Area} + 10.863 \cdot \text{Cup Shape Measure})}}
\]

Variables included: rim volume, cup area, and the cup shape measure.

**LRE 4**

\[
P(Si) = \frac{1}{1 + e^{(3.779 - 2.534 \cdot \text{Rim Area} + 1.507 \cdot \text{Cup Area} + 8.404 \cdot \text{Cup Shape Measure})}}
\]

Variables included: rim volume, cup area, and the cup shape measure.

**Statistical analysis**

In order to perform the statistical calculations and analysis, the SPSS 15.0.1. and MedCalc 8.0.1. statistical programs were used.

An analysis of variance (ANOVA) was done for the four LREs in order to assess the ability to discriminate among the different diagnostic groups. Significant differences of p<0.05 were accepted.

Between the groups of normal subjects and those suspected of having glaucoma, and between the normal group and that with glaucoma, the ROC curves (receiver operating characteristic curves) were calculated for each of the LREs. The ROC curve is a representation of the rate of false positives (1-specificity) against the rate of true positives (sensitivity). It is a useful way of showing the continuous link between sensitivity and specificity of a test or measure. When the ROC curve is closer to the upper left corner of the graph, this means that the test will have a better diagnostic performance.

The area under the curve (AUC) is a number that gives us a measure of a test’s diagnostic ability. A perfect test will have an AUC of 1 (100% sensitivity and 100% specificity) while a test with no diagnostic value will have an AUC of 0.5. In order to establish whether or not there were differences in the AUCs the method described by Hanley McNeil (22) was used.

The sensitivity values of the LREs were calculated for the specificities fixed at 80% and 90%, to discriminate between normal and glaucomatous eyes.

**RESULTS**

The clinical characteristics of the groups under study are found in table I. No statistically significant differences were found in age, visual acuity and the central cornea thickness value among the four groups. The cup/vertical disc ratio measured by stereophotographs was different in the four groups and the base IOP showed differences among the normal group and the rest of the groups, and among the pre-perimetric glaucoma groups and the glaucoma groups. The mean deviation (MD) and the standard deviation from the mean showed differences between the glaucoma group and the other three groups.

The ANOVAs of the LREs showed statistically significant differences between the OHT and control groups, on the one hand, with the groups of subjects suspected of having glaucoma and the glaucomatous group, on the other, in all of the equations (table II and fig. 1). Also, the LRE 1 and LRE 2 equations showed statistically significant differences between the control group and the OHT group. No differences were found between the group of subjects suspected of having glaucoma and those with glaucoma for the values in the four equations, nor for the control groups and the OHT for equations LRE 3 and LRE 4.
When it came to assessing the diagnostic ability of our equations for normal eyes and those with glaucoma, the ROC analysis (Table III, Figure 2) of the LREs showed, for a specificity set at 90%, a sensitivity of 65% for the LRE 1, 68% for the LRE 2 and 67% for both LRE 3 and LRE 4. In the same way, for 80% specificity, the sensitivities found for the four LREs increased to 76% and 77%. The LRE 4 presented the highest AUC (0.865). No significant differences were found between the AUCs of the four LREs.

The AUCs of the LREs for discriminating between normal subjects and those suspected of having glaucoma (Table IV) were slightly lower than in the previous case. Significant differences were not found between them. The best AUC was seen for LRE 1 (0.864). With specificity set at 90%, the sensitivities of the four LREs were between 54% and 60%. Whereas for a specificity of 80%, the sensitivities fluctuated between 70% and 75%.

Between the control group and the glaucoma patients, the AUCs of the discriminant functions included in the software marketed by HRT II were 0.850 and 0.798 for the FSM and RB, respectively. For specificity set at 90%, the sensitivity was 68.6% for the FSM and 50% for the RB. Between the control group and subjects suspected of having glaucoma, the AUCs were 0.820 for the FSM and 0.689 for the RB.

Table I. Demographic and clinical characteristics of the groups under study

<table>
<thead>
<tr>
<th></th>
<th>Control group (n=101)</th>
<th>OHT group (n=126)</th>
<th>Suspected glaucoma (n=121)</th>
<th>Glaucoma group (n=102)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age</td>
<td>55.5 11.5</td>
<td>53.8 12.7</td>
<td>55.7 11.2</td>
<td>58.9 9.50</td>
</tr>
<tr>
<td>V.A.</td>
<td>0.89 0.188</td>
<td>0.93 0.08</td>
<td>0.92 0.1</td>
<td>0.85 0.12</td>
</tr>
<tr>
<td>Baseline IOP</td>
<td>14.6 2.79 O, S, G</td>
<td>22.9 1.55 N</td>
<td>21.3 4.25 N, G</td>
<td>24.2 4.53 N, S</td>
</tr>
<tr>
<td>Vertical C/D</td>
<td>0.27 0.22 O, S, G</td>
<td>0.43 0.27 N, S, G</td>
<td>0.57 0.20 N, O, G</td>
<td>0.71 0.18 N, O, S</td>
</tr>
<tr>
<td>Pachymetry</td>
<td>560.4 21.8</td>
<td>563.35 37.8</td>
<td>558.5 36.1</td>
<td>557.90 34.7</td>
</tr>
<tr>
<td>MD of the AP</td>
<td>-0.56 1.54 G</td>
<td>-0.42 1.25 G</td>
<td>-0.33 1.29 G</td>
<td>-6.81 6.79 N, O, S</td>
</tr>
<tr>
<td>SDM of the AP</td>
<td>1.03 1.06 G</td>
<td>1.11 0.75 G</td>
<td>1.09 1.03 G</td>
<td>4.95 3.75 N, O, S</td>
</tr>
</tbody>
</table>

N: differences with the normal group (p<0.05); O: differences with the OHT group (p<0.05); P: differences with the pre-perimetric group (p<0.05); G: differences with the glaucoma group (p<0.05). OHT: ocular hypertension; MD: mean deviation; VA: best corrected visual acuity; IOP: intraocular pressure; C/D: vertical cup/disc ratio in stereophotographs; AP: conventional automated perimetry; SDM: standard deviation from the mean.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRE 1 Control</td>
<td>101</td>
<td>0.19</td>
<td>O, S, G</td>
</tr>
<tr>
<td>OHT</td>
<td>126</td>
<td>0.36</td>
<td>N, S, G</td>
</tr>
<tr>
<td>Suspected</td>
<td>121</td>
<td>0.74</td>
<td>N, O</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>102</td>
<td>0.76</td>
<td>N, O</td>
</tr>
<tr>
<td>LRE 2 Control</td>
<td>101</td>
<td>0.18</td>
<td>O, S, G</td>
</tr>
<tr>
<td>OHT</td>
<td>126</td>
<td>0.37</td>
<td>N, S, G</td>
</tr>
<tr>
<td>Suspected</td>
<td>121</td>
<td>0.72</td>
<td>N, O</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>102</td>
<td>0.75</td>
<td>N, O</td>
</tr>
<tr>
<td>LRE 3 Control</td>
<td>101</td>
<td>0.19</td>
<td>S, G</td>
</tr>
<tr>
<td>OHT</td>
<td>126</td>
<td>0.24</td>
<td>S, G</td>
</tr>
<tr>
<td>Suspected</td>
<td>121</td>
<td>0.70</td>
<td>N, O</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>102</td>
<td>0.75</td>
<td>N, O</td>
</tr>
<tr>
<td>LRE 4 Control</td>
<td>101</td>
<td>0.21</td>
<td>S, G</td>
</tr>
<tr>
<td>OHT</td>
<td>126</td>
<td>0.36</td>
<td>S, G</td>
</tr>
<tr>
<td>Suspected</td>
<td>121</td>
<td>0.73</td>
<td>N, O</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>102</td>
<td>0.76</td>
<td>N, O</td>
</tr>
</tbody>
</table>

N: differences with the normal group (p<0.05); O: differences with the OHT group (p<0.05); P: differences with the pre-perimetric group (p<0.05); G: differences with the glaucoma group (p<0.05). Column P represents the ANOVA result between the groups (statistical significance p<0.05).

Fig. 1: Average values for the four LREs in the different diagnostic groups.
DISCUSSION

In order to increase the diagnostic usefulness of the papillary morphometric parameters, the software integrated in the HRT II includes two discriminant equations (RB and FSM). It has been observed that the sensitivity and specificity values of those equations are lower when they are used in populations other than those used in their design (23).

The AUCs obtained between the normal and glaucoma groups for the different equations ranged from 0.845 for LRE 3 and 0.865 for LRE 4. These numbers were also somewhat lower than those described in the original population, which showed AUCs between 0.881 and 0.900, with the best equation being LRE 2 (17).

Bowd et al. (14) compared the different discriminant equations published by the HRT II. The best equation was one developed in that same population (AUC of 0.906 ± 0.02) but the AUCs of the rest of the equations studied ranged between 0.848 and 0.890. They observed that the worst diagnostic performance was that described by Mikelber et al. (11) with 64% sensitivity for a specificity set at 90%.

Mardin (15) proposed a discriminant equation with 83.6% sensitivity for specificity set at 95% to differentiate between normal and glaucomatous eyes. But this ability decreased slightly when trying to differentiate between normal and pre-perimetric glaucoma eyes (42.2% sensitivity for the same 95% specificity). In our study, the four LREs showed a better diagnostic performance in differentiating between normal eyes and those suspected of having glaucoma.

It has been noted that the stage of the disease (measured by the AGIS study scale) significantly affects the performance of different imaging equipment used for diagnosing glaucoma, including HRT II (24). The sensitivity of these devices increased when there was a higher level of damage from glaucoma according to the AGIS scale. Because of this it is difficult to compare the results obtained in different studies (different sample selection criteria and varying extent and degree of visual field defects in each glaucoma group).

The diagnostic ability of the LREs was slightly higher than those included in the HRT II software. More specifically, Burk’s discriminant function

Table III. Areas under the ROC curve for each of the LREs between normal subjects and those with glaucoma

<table>
<thead>
<tr>
<th>Area under the curve</th>
<th>95% CI</th>
<th>p</th>
<th>Sensitivity/Specificity Specificity ≥ 90%</th>
<th>Specificity ≥ 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRE 1</td>
<td>0.863</td>
<td>&lt;0.001</td>
<td>65%</td>
<td>77%</td>
</tr>
<tr>
<td>LRE 2</td>
<td>0.863</td>
<td>&lt;0.001</td>
<td>68%</td>
<td>76%</td>
</tr>
<tr>
<td>LRE 3</td>
<td>0.845</td>
<td>&lt;0.001</td>
<td>67%</td>
<td>76%</td>
</tr>
<tr>
<td>LRE 4</td>
<td>0.865</td>
<td>&lt;0.001</td>
<td>67%</td>
<td>76%</td>
</tr>
</tbody>
</table>

Table IV. Areas under the ROC curve for each of the LREs between normal subjects and those suspected of having glaucoma

<table>
<thead>
<tr>
<th>Area under the curve</th>
<th>p</th>
<th>Sensitivity/Specificity Specificity ≥ 90%</th>
<th>Specificity ≥ 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRE 1</td>
<td>0.864</td>
<td>&lt;0.001</td>
<td>60%</td>
</tr>
<tr>
<td>LRE 2</td>
<td>0.854</td>
<td>&lt;0.001</td>
<td>54%</td>
</tr>
<tr>
<td>LRE 3</td>
<td>0.847</td>
<td>&lt;0.001</td>
<td>55%</td>
</tr>
<tr>
<td>LRE 4</td>
<td>0.863</td>
<td>&lt;0.001</td>
<td>59%</td>
</tr>
</tbody>
</table>
(RB) showed a very low sensitivity when trying to discriminate between healthy eyes and those suspected of having glaucoma.

None of the four equations analyzed showed a diagnostic ability sufficient to allow the HRT to be used as the sole tool for diagnosing glaucoma. A specificity of around 95% (25) would be needed to avoid an excessively high number of false positives, especially with a disease like glaucoma, which presents a relatively low prevalence (26). When we adjusted the cut-off points to satisfy such a strict criterion, a little over half the patients were diagnosed. In clinical practice, glaucoma diagnosis should be based on a combined assessment of all the diagnostic tools available (perimetry, tonometry, papillary clinical evaluation, etc.).

The use of discriminant equations improves the diagnostic ability of the HRT II papillary topograph, greatest when the population under study is most similar to the model it is being compared with. The samples used for the regulatory basis included in the different imaging systems limit the ability to obtain higher diagnostic usefulness in populations with different demographic characteristics. The development of larger and more representative samples for the regulatory basis is recommended, as well as that of multivariate discriminant equations adapted to the target population. In this sense, the latest version of the HRT (HRT3) has increased its reference database by including more eyes, with a greater range in the disc size and different ethnic groups. Systems with the ability to adapt or learn, such as neural networks, may also help us improve the diagnostic ability of the HRT and other diagnostic imaging equipment. Further studies will be necessary to evaluate the clinical applicability of these advancements.

REFERENCES


ARCH SOC ESP OFTALMOL 2008; 83: 349-356


