DIAGNOSTIC CAPABILITY OF PULSAR, FDT Y HRT-II IN GLAUCOMA SUSPECTS

CAPACIDAD DIAGNÓSTICA DEL HRT-II Y DE LAS PERIMETRÍAS TOP, PULSAR Y FDT EN PACIENTES SOSPECHOSOS DE SUFRIR GLAUCOMA

GONZÁLEZ-DE-LA-ROSA M1, GONZÁLEZ-HERNÁNDEZ M1, AGUILAR-ESTÉVEZ J2, DÍAZ-ALEMÁN T2, ARMAS-PLASENCIA R3

ABSTRACT

Purpose: To determine the diagnostic capability of PULSAR-T30W, FDT-Threshold-N30 and HRT-II in glaucoma suspects.

Methods: Forty-seven eyes from 47 referred glaucoma suspects (GS) were examined twice with each technique. Cases with TOP-WW-MD>6dB were excluded. Results were compared with those of 70 eyes from 70 normal controls (C).

Results: Mean MD value using TOP-WW in the GS group (0.96dB. SD=1.7) was not significantly different from C (0.8dB. SD=1.77) (p>0.05). Disc area in GS group (2.12 mm². SD=0.34) was significantly greater than in C (1.97 mm². SD=0.45) (p<0.01). For 95.7% specificity, PULSAR-sLV showed the highest sensitivity of 30.9% in individual examinations. The highest reproducible sensitivity in the two examinations was obtained using HRT-II maximum contour elevation (23.4%) and reference height (23.4%), and was 14.9% for various indices after correcting for the influence of disc area (cup area, cup/disc area ratio, maximum contour depression and mean RNFL thickness). Reproducible sen-

RESUMEN

Objetivos: Determinar la capacidad diagnóstica de TOP-32, PULSAR-T30W, FDT-Umbral-N30 y HRT-II en glaucoma de sospecha.

Métodos: 47 ojos de 47 sujetos remitidos por sospecha de glaucoma (SG) se examinaron dos veces. Se excluyeron los casos con defecto medio (MD) superior a 6dB en TOP-32. Los resultados se compararon con los obtenidos en 70 sujetos normales control (C).

Resultados: No se observaron diferencias significativas entre los valores de MD obtenidos en TOP-32, en los grupos SG (0.96dB. DE=1,7) y C (0,8dB. DE=1,77) (p>0.05). El área papilar de SG (2,12 mm². DE=0.34) fue significativamente superior que en C (1,97 mm². DE=0,45) (p<0.01). Para una especificidad del 95%, la raíz cuadrada de la varianza de pérdida (sLV) de PULSAR presentó la mayor sensibilidad (30,9%) en exámenes individuales. La mayor reproducibilidad diagnóstica se obtuvo con la máxima elevación del contorno de HRT-II (23%) y con el Plano de referencia (23,4%), siendo del 14,9% para varios índices, después de corregir la

Received: 31/5/06. Accepted: 12/6/07.

1 Ph.D. in Medicine.
2 Ph.D. in Medicine.
3 Graduate in Optics.
The first author has a commercial interest in the TOP program and in the PULSAR perimeter used in this paper. Partly funded by the Health Research Fund (HRF). Carlos III Institute. Health and Consumer Ministry.

Correspondence:
Manuel González de la Rosa
C/ 25 de Julio, 34
38004 Santa Cruz de Tenerife
Spain
E-mail: mgelarosa@jet.es
INTRODUCTION

Drance (1) warns that a diagnosis that no lesion exists is much more complex than our current notions of hypertension suggest. In fact, as our knowledge of the physiopathology of the condition improves and our diagnostic methods advance, the frontier between normality and glaucoma begins to move. If we do without the current criteria established in practice, we have to admit that this limit is not clearly defined. If it were, our attempts to discover ever more precocious and sensitive procedures would be unnecessary.

Many researchers consider «ocular hypertension» and «glaucoma» to be sufficiently well-defined concepts and so no doubt need be raised. Such an attitude would render works such as the one we propose here pointless. Our hypothesis is that such a criterion is, on the contrary, a disadvantage when it comes to investigating new early diagnostic procedures. Thus for example the custom of establishing a benchmark standard is more of a hindrance than an aid in the structuring and interpretation of many of these works. It is evident that, if we assume that there is a diagnostic procedure that should be recognized as a standard, then, in the case where a better procedure exists, its results will be undervalued in comparison with those offered by the first. When a standard is applied, the normal statistical procedures for sensitivity and specificity analysis are only relatively certain if this procedure is more perfect than the one assessed. If not, such an analysis becomes a mathematical sophism.

Ideally, none of the procedures evaluated should interfere in the selection process. Our intention has been to analyze several of the examination procedures developed in the last few years in order to estimate their ability to provide an early diagnosis, attempting to avoid the logical labyrinth constituted by the use of a rigid standard and avoiding, as far as possible, preliminary selection biases.

In this work we have attempted to investigate the diagnostic ability and reproducibility of a series of tests applied to patients referred with suspected glaucoma.

SUBJECTS, MATERIAL AND METHODS

All the normal controls (Group C) and those with suspected glaucoma (SG) underwent a complete ophthalmologic exam, including:
— Family and personal history related to glaucoma, such as migraine, Raynaud’s syndrome, arterial hypo- or hypertension, diabetes, cardiovascular illnesses, smoking or medication.
— Refraction and visual acuity.
— Anterior segment examination and gonioscopy.
— Intraocular pressure (Goldmann tonometer).
— Topography of the optic nerve with confocal laser (HRT-II) *(Heidelberg Engineering, Gerhart-Hauptmann-Strasse 30. 69221 Dossenheim, Germany).*

— FDT Perimetry (Threshold N30) *(Carl Zeiss Pty Ltd, 114 Pyrmont Bridge Road, Camperdown, NSW 2050, Australia).*

— Conventional white-on-white luminous threshold perimetry (TOP «32» WW) *(2)* using the Octopus 311 perimeter *(HAAG-STREIT, Gartenstadtstrasse 10 CH-3098 Köniz. Berne. Switzerland).*

— PULSAR perimetry (T30W) *(Octopus experimental prototype. HAAG-STREIT, Gartenstadtstrasse 10 CH-3098 Köniz. Berne. Switzerland).*

— Exam of posterior pole in mydriasis.

**Inclusion Criteria**

The patients in the SG group were recruited consecutively between 07/03/2003 and 08/08/2003 at our Glaucoma Department from those referred by different out-patient clinicians and other professionals. They were especially recruited depending on their interocular pressure figures. Patients with and without hypertension treatment were included, as were those with limit pressures associated with asymmetric glaucoma, family history of glaucoma or suspected papilla. The suspicion of glaucoma was established by consideration of all relevant data. The centers referring patients did not have a perimeter or papillary topography equipment.

Group C was constituted by normal volunteers who came to the clinic during the same time period for refraction problems or were accompanying other patients. The field of vision was not taken into account for their inclusion.

**Exclusion Criteria**

Subjects with associated pathology were excluded, as well as those with treatment that might affect vision or field of vision, significant loss of crystalline transparency, corrected visual acuity of less than 0.8, pupil diameter of less than 3 mm or a lot of errors in the perimetric check-ups. An open anterior chamber was required in the gonioscopic exam. All subjects had refractive errors of less than 5 sphere dioptres and two cylinder dioptres.

For the selection of the SG group, we tried to ensure that our research was limited to early glaucoma by excluding those cases that presented mean defect (MD) perimetric values in excess of 6dB at the first exam. We excluded those suffering from a condition or taking any medication that might affect vision, and those that presented a family history of glaucoma or ocular hypertension, or signs associated with the illness (migraine, Raynaud’s syndrome, vasculopathy, etc).

Those subjects that did not complete all exams were also excluded.

**Techniques**

Examination procedures with HRT-II, FDT and TOP are well-known in the literature. In contrast, PULSAR perimetry is an experimental procedure that needs to be briefly described:

A digital photometer placed in the corner of the 19-inch Samsung SyncMaster 959NF monitor *(Samsung Electronics Co, 416 Maetan-3Dong, Paldal-Gu, Suwon City, Kyungki-Do, Korea)* reports the luminance to a computer via a USB connection. The computer periodically regulates the brightness and contrast according to a pre-established scale. In this way, it can show a scale with the 256 levels commonly used in computing.

The stimulus consists of a circular wave, waning in contrast towards the periphery, similar to that of a drop falling on water. The background lighting is 100 asb and the stimulus oscillates above and below the background levels, in such a way that it is iso-illuminant with this on the whole (Fig. 1). The duration of the stimulus is 500 msec and the diameter 5°. The stimulus wave can be modulated in terms of spatial resolution (from 0.5 to 6.3 cycles per degree on a scale of 12 logarithmic levels), contrast (32 logarithmic levels between 3 and 100%), color (white, red, green and blue), centrifugal movement (from 2 to 20 cycles per second on a logarithmic scale of 11 levels) or time frequency, oscillating in-phase and counter-phase (10, 15 or 30 Hz).

For spatial resolution, the scale is established from a minimum spatial frequency value of 0.5 cycles/degree, using the following formula: Spatial resolution (dLog) = 10 x Log (Spatial Frequency/0.5).

As the contrast is below and above the background level, the following formula was used: Con-
Contrast = –20 x Log (Central amplitude from the background/intensity of the background).

The instrument can examine several visual functions related to the Contrast Sensitivity Function (CSF) in the center of the field of vision, with the rest remaining constant, in order to measure the sensitivity of some with respect to others, using the Tendency Oriented Perimetry (TOP) strategy. On the basis of previous studies (3-5), the 36-level stimulus scale used had spatial resolution (sr) and contrast (c) that varied simultaneously (src units). This scale has the advantage of being approximately perpendicular to the contrast sensitivity curves, which are parallel to each other in the different regions of the field of vision (Fig. 1). In this way the threshold halfway down the descending portion of the curve towards the high spatial frequencies is examined, i.e. the area in which glaucomatous defects tend to occur, at least for central vision (6-8).

In the same way as for FDT (9), the thresholds in PULSAR are altered when refraction is inadequate [Fernandez-Baca G, Gonzalez de la Rosa M, Perez Fernandez JR & Gonzalez Hernandez M (2003): Influence of visual acuity on PULSAR T30W Perimetry. 79th Congress of the Spanish Ophthalmology Society. Valencia]. Because of that, the instrument shows a visual acuity test for myopia just before the start of each exam. In the PULSAR T30W programme, the stimulus is shown in-phase and counter-phase at 30 Hz, which has been shown to be useful in the early diagnosis of glaucoma (4,10-12).

In all cases, the first two perimetries were excluded to reduce the «learning effect». The three perimetric exams and the HRT-II were applied twice in each participant. All the exams were performed over a period of less than one month, with a minimum rest of 10 minutes. The field of vision test was performed with the refraction for hypermetropy for FDT and TOP, and for myopia in PULSAR. The working hypothesis was that the most efficient diagnostic technique would be the one that best discriminated between both SG and C groups, while maintaining high specificity.

Linear regression was used to analyze the correlation between the numerical indices for each technique. The association between perimetric and anatomical indices was evaluated using linear and logarithmic regression. The probability that the indices would be significantly different between the SG and C groups was evaluated using Student's t. An association was also evaluated for indices obtained previously through discriminating analysis (11).

An ROC (Receiving Operating Characteristic) analysis was performed for each index, calculating the ROC areas, confidence intervals, cut-off points, sensitivity, specificity and probability (Student's t) using an application designed with Excel (Microsoft Corporation). Our study was centred on the cut-off point providing a specificity closest to 95%, since early diagnosis requires high specificity. The result of this application was verified with the NCSS 2000 programme (NCSS, Kaysville, UT). The diagnosis was considered to be reproducible when it coincided in both examinations. Thus, the reproducibility in the diagnosis of abnormalities was calculated over the total of the sample, and not as a percentage of the...
sensitivity. In order to evaluate reproducibility, the intra-class correlation coefficient was also analyzed (13) using the SPSS v. 12.0.1. programme (SPSS Inc., Chicago, ILL). The same software was used to calculate the statistical power of the results in relation to the number of cases.

In this study the principals of the Helsinki Declaration were respected.

**RESULTS**

Two participants from Group C were excluded for poor collaboration and one for not completing the second group of exams. Two in SG were excluded for cataracts, one for poor collaboration, one for not disclosing ocular hypertension and two for not completing the second group of exams.

Finally 70 eyes of 70 normal subjects in Group C (average age 40.1 years, SD=16.5, 39 women) were studied and 47 eyes of 47 patients in Group SG (average age 60.4 years, SD=10.8, 31 women). All the patients were of Caucasian race except for one negro belonging to Group SG.

In Group C, no correlation was observed between age and any of the indices analyzed.

The duration of the perimetric exams was as follows: TOP 32 WW = 2:52 minutes (SD=0:17), PULSAR T30W= 3:50 minutes (SD=0:12) and FDT N30= 4:31 minutes (SD=0:19).

The mean defect (MD) measured using TOP 32 WW was: Group C = 0.80dB (SD=1.8) and Group SG = 0.96dB (SD=1.7) (p>0.05) (Fig. 2).

The mean square root of the loss variance (sLV) using TOP 32WW was: Group C = 1.77dB (SD=0.68) and Group SG = 2.30dB (SD=1.19) (p<0.01).

The mean MD using PULSAR was: Group C = 0.13dB (SD=2.11) and Group SG = 1.25dB (SD=2.72) (p<0.01).

The mean square root of the loss variance (sLV) using PULSAR was: Group C = 2.25dB (SD=0.58) and Group SG = 2.94dB (SD=1.00) (p<0.01).

The mean MD using FDT was: Group C = -1.59dB (SD=2.42) and Group SG = 1.81dB (SD=3.03) (p>0.05).

The mean pattern standard deviation (PSD) using FDT was: Group C = 4.33dB (SD=1.97) and Group SG = 4.50dB (SD=2.28) (p>0.05). 39% of the subjects from Group C presented at least one defect in the five central points.

The average papillary area using HRT-II was: Group C = 1.97 mm$^2$ (SD=0.45) and Group SG = 2.12 mm$^2$ (SD=0.34) (p<0.01).

Table I shows the correlation coefficient (r) between the principal numeric indices of the three types of perimetry test. The greatest correlation was observed between the values of MD and sLV-PSD in each procedure, followed by the values of the respective MD values of PULSAR and FDT.

Using linear regression, MD PULSAR presented a significant correlation with 9 of the 23 main indices of HRT-II, FDT PSD with three indices, FDT MD with one and PULSAR sLV with one. The application of logarithmic regression did not change these results.

Many of the HRT-II indices presented a high degree of dependency on the size of the papilla.

![Fig. 2: Distribution of the MD TOP WW in the control and suspected glaucoma patients.](image)

**Table I Correction coefficients between the three perimetries**

<table>
<thead>
<tr>
<th></th>
<th>MD TOP WW</th>
<th>sLV TOP WW</th>
<th>MD PULSAR</th>
<th>sLV PULSAR</th>
<th>MD FDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>sLV TOP WW</td>
<td>0.72 (p&lt;0.01)</td>
<td>0.23 (p&lt;0.01)</td>
<td>0.67 (p&lt;0.01)</td>
<td>-0.20 (p&lt;0.01)</td>
<td>0.25 (p&lt;0.01)</td>
</tr>
<tr>
<td>DM PULSAR</td>
<td>0.11 (p&lt;0.05)</td>
<td>0.06 (p&lt;0.05)</td>
<td>0.25 (p&lt;0.01)</td>
<td>0.17 (p&gt;0.05)</td>
<td>-0.59 (p&lt;0.01)</td>
</tr>
<tr>
<td>sLV PULSAR</td>
<td>-0.021 (p&gt;0.05)</td>
<td>-0.25 (p&lt;0.01)</td>
<td>-0.36 (p&lt;0.01)</td>
<td>-0.27 (p&lt;0.01)</td>
<td></td>
</tr>
<tr>
<td>MD FDT</td>
<td>-0.20 (p&lt;0.01)</td>
<td>0.22 (p&lt;0.01)</td>
<td>0.67 (p&lt;0.01)</td>
<td>0.25 (p&lt;0.01)</td>
<td>-0.59 (p&lt;0.01)</td>
</tr>
<tr>
<td>PSD FDT</td>
<td>0.14 (p&gt;0.05)</td>
<td>0.22 (p&lt;0.01)</td>
<td>0.67 (p&lt;0.01)</td>
<td>0.25 (p&lt;0.01)</td>
<td>-0.59 (p&lt;0.01)</td>
</tr>
</tbody>
</table>
Figure 3 shows the correlation coefficients and the probability value in Group C. A significant association was observed between the maximum contour elevation and the reference height in both Group C ($r=0.41$, $p=0.0004$) and Group SG ($r=0.38$, $p=0.0002$).

Figure 4 shows that the best results for the perimetric exams were obtained with the PULSAR sLV. The 95% percentile of the ROC areas in the MD FDT and PSD FDT was lower than the 5% percentile of sLV PULSAR.

A high sensitivity was observed (43.6%) for the formula obtained previously through discriminating analysis combining anatomical and functional indices (11) \[ \text{COMBINATION} = -4.95 + (\text{sLV TOP WW} \times 1.09) + (\text{sLV PULSAR} \times 0.61) + (\text{MD TOP WW} \times -0.33) + (\text{Max. depr. cont.} \times 3.56) \]. However, reproducibility was very low (12.8%).

Some HRT II indices were offset depending on papillary size. Garway-Heath (14) and Jonas (15) have proposed two methods for this. The regression equation between each index and papillary size was calculated for the control population. The formula for the resulting regression line was used to calculate the value that each index should reach in each specific patient as a function of papillary size. For the subjects in the SG Group, the deviation of their real values was subsequently calculated with respect to that expected in the normal population.

Table II shows the results for a specificity of 95.7% (5% and 95% confidence intervals = 95.7±4.8%). The English names have been kept for the indices, as no commonly accepted Spanish translation exists for all of these. The cup area/papilla ratio and the vertical cup/papilla ratio are shown in the table, with and without correction. The reproducibility indicates that the percentage of SG cases with significant differences with respect to Group C in the two exams. For example, with the Maximum Contour Elevation index a little over a quarter of the cases turned out abnormal (sensitivity = 26.6% on average in both exams) and in 23.4% of all the cases both exams came out abnormal. The probability (p value) of the intra-class correlation coefficient was less than 0.001 in all cases, except for the «Mean Variability» (p=0.015) and «Combination» (p=0.629).

When both perimetric examinations were abnormal, we verified if the same thing happened in both HRT-II exams. This occurred in three out of four cases in PULSAR MD, six out of eight in PULSAR sLV, one out of three in FDT MD and in one in two cases for FDT PSD.

The greatest ROC area corresponded to PULSAR sLV, but an important overlap was observed between its confidence intervals and those of the other best indices obtained (Fig. 5).

The power of the three psychophysical examinations was calculated to distinguish between their diagnostic abilities and those of the main HRT-II indices (Figs. 4 and 5). For the two exams, compar-
ing groups C and SG, the power values were: PUL-
SAR MD (0.727, 0.652), PULSAR sLV (0.995,
0.992), FDT MD (0.081, 0.066), FDT PSD (0.060,
0.189), Maximum contour depression (0.929,
0.918), Maximum contour elevation (0.836, 0.776),
Reference height (0.749, 0.905) and Mean variabil-
ity (0.880, 0.671).

### DISCUSSION

#### Discussion of the Functional Results

The mean MD values for TOP WW in our SG
sample are lower than those in most of the studies
carried out on early glaucoma (16-23), which
explains why the results are less than those of other
studies using cases with deeper defects.

The selection of cases for Group SG by a T OP
WW MD score of less than 6dB in the f irst exam
has introduced a small bias in the results of this
technique with respect to the others, where the more
advanced defects were not eliminated. The diagnos-
tic potential of TOP WW could thus have been
underestimated and it has not been included in the
study.

![Fig. 5: Confidence intervals of the ROC area for the best indices obtained.](image-url)
There is no reason to suppose that the age difference between the C and SG Groups have an influence on the results. MD is an age-adjusted index and no-one has reported any impact of age on sLV, PSD or the overall HRT-II indices.

On the contrary the present work confirms our earlier studies that indicate that PULSAR is sensitive and specific for the early diagnosis of glaucoma, just like our previous observations on sLV. TOP calculates thresholds taking physiopathological interrelations into consideration (24). The studies that have assumed that the softening of the visual field it produces and the reduction it induces in the LV values may lead to a diagnostic limitation are based on suppositions (25) or on simulations using isolated scotomas, without relation to the physiopathology of glaucoma (26).

The reality is just the opposite: with TOP few cases of elevated LV are observed in normal subjects. The principal components of LV are the threshold fluctuation, topographical irregularity, and errors. The errors in TOP are limited as the neurological «fatigue effect», attention lapses and the methodological defects caused by examining the points independently are reduced (27-29). Flammer (30) indicated that the first signs of glaucoma are not the focal or diffuse defects but the increase in threshold fluctuation. By reducing methodological errors, TOP seems to detect the increased fluctuation as a neuronal defect, probably prior to or at least independent of anatomical damage. PULSAR simultaneously examines functions associated with contrast and high spatial and timing frequencies that could facilitate the detection of these functional defects.

Discussion of the Morphological Results

If the papillary size is taken into account, the diagnostic sensitivity of some HRT-II indices is substantially reduced. The proportion of this reduction is variable and does not affect diagnostic reproducibility, which continues to remain high. Some of them, like the Maximum contour elevation and the Reference height, attained maximum diagnostic capacity. HRT-II calculates a specific value for the Reference height in each exam, 50 microns below the contour line in the temporal sector between 350° and 356°. This may explain the comparable results obtained with both exams.

Discussion of the Relation Between Function and Morphology

With respect to the results obtained with HRT-II, we must acknowledge that previous studies (11) had led us to think that its diagnostic capacity was below the value of sLV in TOP-WW and the PULSAR indices, but this opinion should be revised taking into account the new data obtained with respect to diagnostic reproducibility and the intra-case correlation coefficient for each index.

It has not been possible to avoid a certain degree of bias in the selection of the SG Group, since the doctors admitting patients knew the ophthalmoscopic appearance of the papilla. Thus, the differences between the best anatomical and perimetric indices may be circumstantial, caused by the characteristics and size of the sample. However, we have not wished to suppress the information provided by the HRT-II, so as not to exclude possible pre-perimetric glaucomas. The intra-class correlation was, in general, greater in the HRT-II indices than in the perimetric ones, in contradiction with the results obtained by other authors (31) who found no differences in the long-term fluctuation of the perimetric and HRT-II indices. However the bias produced by the papilla in the selection of the SG Group and the sensitivity differences observed between PULSAR and the best HRT-II indices introduce doubts with respect to the precocity of the morphological defects, since the ROC of HRT-II areas are lower and their confidence intervals overlap with those of sLV PULSAR.

We should add that the formula we obtained in a previous work through discriminant analysis (11) and which associated various perimetric and HRT-II indices has confirmed its high sensitivity in this work. However, this fact contrasts with a fairly poor diagnostic reproducibility, in such a way that its overall diagnostic ability is inferior to that of the best perimetric and HRT-II indices. We should conclude upon seeing this new information that the association of perimetric and papillary indices does not provide as many diagnostic advantages as we had supposed.

The absence of differences between linear and logarithmic regression could be related to the characteristics of our sample: incipient glaucoma, with a scant range of defects. The MD with all three types of perimetry presented a high correlation with papillary damage, but with less diagnostic ability.
than sLV PULSAR, which also reached a lower intra-class correlation with MD. This might indicate that sLV PULSAR provides pre-anatomical information, while MD would be associated with more advanced phases of the pathological process.

The SG that are differentiated from C using these techniques will not necessarily develop glaucoma in the future. This is particularly likely in many suspect cases with large papillae. However we should remember that, in the Ocular Hypertension Treatment Study (32,33), the PSD value in the Humphrey perimetry, equivalent to the sLV in Octopus perimetry, seems to exceed MD as a risk factor for developing glaucoma.

The methodology of this work has been similar to another recent publication of our Group, where some additional considerations can be consulted with respect to the possibility of reducing the usual biases in this kind of study (34).

REFERENCES


