Several studies have investigated the ability of machine learning classifiers to interpret visual field test results. In an earlier study we assessed the diagnostic accuracy of a previously trained Artificial Neural Network (ANN), with its input based on pattern deviation probability scores, in an independent data set of fields from both healthy and glaucomatous eyes (Bzios et al 2007). The performance of our ANN was very good (AROC 0.98) and similar to that obtained when training the network. The output of our ANN is a number ranging between 0 and 1, where a cut off value is set to best separate glaucomatous from healthy fields. Outputs close to the endpoints, 0 and 1, indicate high classification certainty, while an output close to the middle of the range indicates uncertainty. Therefore, it is possible to let the ANN provide not only a classification of the visual field test, but also an estimation of the certainty of classification.

To compare diagnostic accuracy and certainty of an ANN trained for classification of visual fields for the diagnosis of glaucoma with that of clinicians having full access to HFA Statpac Single-Field Analysis.

Glaucoma patients

Visual fields were selected in a pseudo-random order from our perimetric database. Only fields from patients with a diagnosis of primary open angle glaucoma including normal tension glaucoma, exfoliation or pigmentary glaucoma were eligible. Visual fields of glaucoma patients with comitant cataract or diabetes without retinopathy could be included. Glaucoma was defined as optic disc damage compatible with glaucoma as assessed by either expert evaluation of optic disc photographs or a comprehensive description of the optic disc in the patient record. Fields with Mean Deviation (MD) values worse than -10 dB were excluded.

Healthy subjects

Visual fields from healthy subjects were randomly selected from a database of visual fields originated and established to establish normal limits for SFTA strategies (<20% of the full normal database).

Subjective assessment

Clinicians with different experience of glaucoma management independently assessed Single-Field Analysis printouts with full Statpac information.

One visual field from each of 99 glaucoma patients with a mean age of 75.8 years (ranging from 55 to 95 years), and 66 healthy subjects with a mean age of 64.5 years (ranging from 51 to 83 years) were included. Mean MD was -5.8 dB (ranging from -10 to +0.3 dB) for glaucoma patients and +0.2 dB (ranging from -3.5 to +2.4 dB) for healthy subjects.

Thirty clinicians participated: 4 glaucoma specialists, 8 general ophthalmologists, 10 ophthalmologists with other subspecialties and 8 glaucoma were eligible. Visual fields of glaucoma patients with concomitant cataract or diabetes without retinopathy could be included. Glaucoma was defined as optic disc damage compatible with glaucoma as assessed by either expert evaluation of optic disc photographs or a comprehensive description of the optic disc in the patient record. Fields with Mean Deviation (MD) values worse than -10 dB were excluded.

Healthy subjects

Visual fields from healthy subjects were randomly selected from a database of visual fields originated and established to establish normal limits for SFTA strategies (<20% of the full normal database).

Subjective assessment

Clinicians with different experience of glaucoma management independently assessed Single-Field Analysis printouts with full Statpac information.

One visual field from each of 99 glaucoma patients with a mean age of 75.8 years (ranging from 55 to 95 years), and 66 healthy subjects with a mean age of 64.5 years (ranging from 51 to 83 years) were included. Mean MD was -5.8 dB (ranging from -10 to +0.3 dB) for glaucoma patients and +0.2 dB (ranging from -3.5 to +2.4 dB) for healthy subjects.

Thirty clinicians participated: 4 glaucoma specialists, 8 general ophthalmologists, 10 ophthalmologists with other subspecialties and 8 glaucoma were eligible. Visual fields of glaucoma patients with concomitant cataract or diabetes without retinopathy could be included. Glaucoma was defined as optic disc damage compatible with glaucoma as assessed by either expert evaluation of optic disc photographs or a comprehensive description of the optic disc in the patient record. Fields with Mean Deviation (MD) values worse than -10 dB were excluded.

Healthy subjects

Visual fields from healthy subjects were randomly selected from a database of visual fields originated and established to establish normal limits for SFTA strategies (<20% of the full normal database).

Subjective assessment

Clinicians with different experience of glaucoma management independently assessed Single-Field Analysis printouts with full Statpac information.

Table 1 Sensitivity and specificity of the artificial neural network (ANN) and subjective clinicians with varying experience of glaucoma.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (min, max)</td>
<td>Mean (min, max)</td>
</tr>
<tr>
<td>ANN</td>
<td>95% (91%, 100%)</td>
</tr>
<tr>
<td>Glaucoma experts</td>
<td>91% (79%, 97%)</td>
</tr>
<tr>
<td>General ophthalmologists</td>
<td>83% (73%, 93%)</td>
</tr>
<tr>
<td>Other subspecialists</td>
<td>84% (76%, 96%)</td>
</tr>
<tr>
<td>Residents</td>
<td>83% (73%, 93%)</td>
</tr>
</tbody>
</table>

Significantly better compared to the average clinicians, p < 0.001.

CONCLUSIONS

Our results suggest that a well-designed trained and validated ANN performs at least as well as clinicians in assessments of visual fields for the diagnosis of glaucoma. Incorporating such a network as a standard analysis tool for interpretation of perimetric test results could facilitate maintaining high standards of such interpretation in many glaucoma care settings.

ACKNOWLEDGEMENT

This study was supported by the Swedish Research Council, grant K2011-65X-10426-19-3, the Herman Jarnsårdt Foundation and Crown Princess Margareta’s Foundation.

REFERENCES