Comparison of visual field severity classification systems for glaucoma

M. Ng[1], **P.A. Sample**[1], **J.P. Pascual**[1], **L.M. Zangwill**[1], **C.A. Girkin**[2], **J.M. Liebmann**[3], **R.N. Weinreb**[1], and **L. Racette**[1]

[1]Hamilton Glaucoma Center, University of California, San Diego, La Jolla, CA

[2]Department of Ophthalmology, University of Alabama, Birmingham, AL

[3]New York Eye and Ear Infirmary, New York, NY

Abstract

Purpose—To compare categorical severity classification systems for glaucoma.

Methods—This cross-sectional study included 1,921 eyes (49.5% OD) from 1,137 participants from the Diagnostic Innovations in Glaucoma Study (DIGS) and African Descent and Glaucoma Evaluation Study (ADAGES). Standard automated perimetry fields were classified using the: (1) Advanced Glaucoma Intervention Study scoring system (AGIS), (2) Glaucoma Severity Staging system (GSS), and (3) Enhanced Glaucoma Severity Staging system (eGSS). Systems were characterized using the following continuous measures of severity: mean deviation (MD), pattern standard deviation (PSD) and visual field index (VFI). Classifications between systems and with optic disc stereophotograph assessment were compared (Kappa) and some stages were consolidated to evaluate severity classification across systems (Wilcoxon test).

Results—MD, PSD and VFI were significantly different between GSS and AGIS, and GSS and eGSS in normal and abnormal fields (P<0.005). Agreement between AGIS and eGSS was substantial (K=0.715±0.012); agreement between GSS and eGSS (K=0.559±0.014) and AGIS (K=0.519±0.016) was moderate. EGSS tended to stage abnormal fields most severely followed by GSS and then AGIS (P<0.001).

Conclusions—The presence of glaucomatous optic neuropathy increases with staging severity for all systems. However, different systems led to different severity staging. Of the systems examined in this study, eGSS may be the better choice for its ease of use for both clinicians and researchers.

More than 20 severity classification methods for visual fields have been proposed since the American Medical Association published its scoring system in 1958.[1] Visual field severity affects patient treatment decisions and monitoring progression is an integral part of patient care. In addition to the patient care perspective, understanding the severity of a study population helps clinicians and researchers make comparisons across research findings. However, only a few in-depth comparisons between systems have been made.[1, 2]

Classification systems measure severity using either categorical or continuous variables. The systems vary by their number of severity stages, the ease in their implementation, and whether subjective judgments or quantitative parameters determine a particular stage of severity. These differences should influence the choice of a severity classification system or make comparisons across systems difficult. We applied three categorical systems to a large number of visual fields to determine which system tends to stage more severely and how continuous global indices vary with progressing severity within these categorical systems. We further examined the incidence of glaucomatous optic neuropathy with severity across systems using optic disc stereophotograph assessments. The categorical severity classification systems we examined were: (1) the Advanced Glaucoma Intervention Study (AGIS) scoring method,[3] (2) the Glaucoma Severity Staging (GSS) system by Mills et al.,[4] and (3) the Enhanced Glaucoma Severity Staging (eGSS) system by Brusini and Filacorda.[5]

METHODS

Participants

All participants included in this cross-sectional study were selected from the ongoing longitudinal Diagnostic Innovations in Glaucoma Study (DIGS), conducted at the Hamilton Glaucoma Center at the University of California at San Diego (UCSD) and the African Descent and Glaucoma Evaluation Study (ADAGES), a multi-center study conducted at UCSD, the University of Alabama at Birmingham, and the New York Eye and Ear Infirmary. These ongoing studies are prospectively designed to assess structure and function in glaucoma. The methods, inclusion and exclusion criteria for participation in DIGS and ADAGES are the same (detailed below).[6] Healthy participants were recruited from the general population through advertisement, from referring practices and from the staff and employees at each of the study institutions. Informed consent was obtained from all participants and the Institutional Review Board of each pertinent institute approved the study, which adheres to the tenets of the Declaration of Helsinki and to the Health Insurance Portability and Accountability Act (HIPAA) for research involving human subjects.

Inclusion criteria for DIGS/ADAGES—Participants underwent complete ophthalmologic examinations including slit lamp biomicroscopy, intraocular pressure (IOP) measurement, and dilated stereoscopic fundus examination. Simultaneous stereoscopic photographs with adequate clarity were obtained for all participants. At study entry, all participants had open angles, a best-corrected acuity of 20/40 or better, a spherical refraction within 5.0 D, and cylinder correction within 3.0 D. A family history of glaucoma was allowed.

Exclusion criteria for DIGS/ADAGES—Participants were excluded if they had: (1) a history of intraocular surgery other than uncomplicated cataract or glaucoma surgery, (2) non-glaucomatous secondary causes of elevated IOP (e.g. iridocyclitis, trauma), (3) other

intraocular diseases affecting the visual field (e.g. pituitary lesions, demyelinating diseases, HIV+, AIDS, or diabetic retinopathy), (4) been taking medications known to affect visual field sensitivity, and/or (5) deficiencies other than glaucoma affecting color vision (screened with the Farnsworth-Munsell-D15 test).

All participants underwent standard automated perimetry (Humphrey Visual Field Analyzer II-i, Carl Zeiss Meditec, Inc., Dublin, CA) using program 24-2 and the Swedish Interactive Thresholding Algorithm. The two locations just above and below the blind spot were not included in the analysis, leaving a total of 52 test locations. Adequate refraction was provided according to the manufacturer's specifications and the pupils had a diameter of ≥ 3 mm. The pupils were dilated when this requirement was not met. The perimeter provides a STATPAC analysis for each visual field exam that includes the global indices: mean deviation (MD), pattern standard deviation (PSD), visual field index (VFI), and the Glaucoma Hemifield Test (GHT) output. It also indicates which location on the pattern deviation plot (PDP) and total deviation plot (TDP) are flagged as abnormal at <0.5%, <1%, <2% and <5% relative to the internal normative database[7].

Participants who were new to perimetry were given practice tests prior to the visual fields that were included in the analysis. Only reliable visual fields (<33% fixation losses and false negatives and <15% false positive responses) were included. Both eyes from each subject were included, except in cases in which only one eye was tested or a visual field for one eye was unreliable. All visual field tests from DIGS and ADAGES are reviewed for artifacts and suspected learning effects by the Visual Field Assessment Center (VisFACT), UCSD.[8]

Optic disc stereophotographs were obtained for all eyes (TRC-SS; Topcon, Paramus, NJ, USA or Nidek 3-DX, Fremont, CA, USA) taken within six months of the visual field tests. Stereophotographs were assessed for the presence of glaucomatous optic neuropathy by two trained graders, who were masked to the identity of the participant as well as to other grader determinations. A third trained grader, also masked to the identity of the participant and other graders, adjudicated any disagreements between the first two graders.

Severity classification systems

Each visual field was staged with the 3 different categorical visual field classification systems described below.

The Advanced Glaucoma Intervention Study scoring system (AGIS).[3]—This scoring system is based on the number and depth of neighboring depressed test locations on the total deviation plot in the nasal area, upper hemifield, and lower hemifield (Appendix Table A1). The visual field is divided into 4 concentric-like regions, where the outermost region is most sensitive to a depression. Each visual field is given a score between 0 and 20, where 0 indicates that no defective points were measured and 20 indicates that a total of at least 2 depressed locations in the nasal area and 9 depressed locations in each hemifield were measured. The severity of a visual field is broken into categories based on this score: (i) a score of 0 is a normal field categorized into stage 1; (ii) scores from 1–5 are fields with mild damage categorized into stage 2; (iii) scores from 6–11 are fields with moderate damage categorized into stage 3; (iv) scores from 12–17 are fields with severe damage categorized into stage 4; (v) scores from 18–20 are end-stage fields categorized into stage 5. All fields could be unambiguously classified.

The Glaucoma Severity Staging system (GSS).[4]—GSS, a modified version of the Hodapp-Anderson-Parrish system,[9] is based on mean deviation (MD), the location and number of points depressed on the pattern deviation plot at P<0.01 and P<0.05, the Glaucoma Hemifield Test, and visual acuity (Appendix Table A2). GSS has a total of 6

stages: (i) fields with no defect are categorized into stage 0; (ii) fields with early defect are categorized into stage 1; (iii) fields with moderate defect are categorized into stage 2; (iv) fields with advanced defect are categorized into stage 3; (v) fields with severe defect are categorized into stage 4; (vi) and fields with end-stage disease are categorized into stage 5. Some fields could be classified ambiguously.

The Enhanced Glaucoma Severity Staging system (eGSS).[5]—This staging system relies on two indices: (1) mean defect (MD), and (2) pattern standard deviation (PSD). See example visual fields staged with Brusini & Filacorda's chart in the Appendix (Figure A1). EGSS has a total of 7 stages: 0, the border between 0 & 1, 1, 2, 3, 4 and 5, where stage 0 are fields with no defect and stage 4 are fields with the greatest defect. Visual fields categorized into stages 1 and higher are also characterized by their defect type: (a) localized, (b) mixed, and (c) generalized. Defect characterization also relies on PSD and MD, but different criteria are used than for the severity staging. Note that the authors provided constants and conversion factors for formulas and estimated values, and that corrected loss variance (CLV) or LV can be used instead of CPSD or PSD, respectively.[5] It is also important to note that GSS and eGSS were developed independently despite the similarity in names.[1] All fields could be unambiguously classified using this system.

Analyses

Statistical analyses and staging implementation were performed using JMP version 5.1.2 (SAS Institute, Inc., Cary, NC).

Descriptive statistics—We compared the mean, median, and distributions of continuous STATPAC global indices MD and PSD and the Visual Field Index (VFI)[10] across all staging systems. The VFI is based on the pattern deviation values, which adjust for an overall drop in height of the visual field[11] and are less affected by cataract than the total deviation values. VFI is also a measure of severity and is a continuous variable given as a percentage, where 100% represents a normal field and 0% represents a perimetrically blind field. Rate of progression (assessed using linear regression) presented as a yearly change in VFI, can be calculated with the Guided Progression Analysis available for Standard Automated Perimetry on the HFA II-i software, version 4.2 or higher.

Agreement of normal versus abnormal classification of visual fields—We compared the agreement of each categorical classification system's categorization of visual fields into the non-defective (normal) stage versus the defective stages (all levels of severity). For example, the visual fields that had no defect using AGIS, GSS, and eGSS were fields in stage 1, 0, and 0, respectively; the visual fields with defects were categorized into stages 2 to 5, 1 to 5, and border to 4, respectively. Agreement was assessed using the Kappa statistic,[12, 13] which rates the strength of agreement as: poor (K=0.00), slight (K=0.01–0.20), fair (K=0.21–0.40), moderate (K=0.41–0.60), substantial (K=0.61–0.80), or almost perfect (K=0.81–1.00).

Comparison of global indices between severity staging systems—The continuous values of MD, PSD and VFI were also compared between the stages with no defect and between the defective stages across systems (Wilcoxon non-parametric test, 2-tailed).

Comparing severity staging with optic disc stereophotographs—We also compared agreement of each categorical visual field classification system with the optic disc stereophotograph outcome using the Kappa statistic.[12] Thus, visual fields that were classified as abnormal by the staging systems were grouped together to see whether there

was a significant agreement with presence of glaucomatous optic neuropathy as assessed using stereophotograph. We further report the incidence of abnormal optic disc stereophotographs for each system's severity stage.

Consolidated stages of severity—AGIS, GSS and eGSS each have a different number of stages, making it impossible to make a direct comparison of the stages. To circumvent this, we combined some stages so that all systems had 4 consolidated stages (CS) to allow for comparisons (Table 1). Correlation analyses were used to determine the consolidation scheme that resulted in maximal agreement for comparison. CS stage 1, for example, includes fields without defect or that are borderline; thus, eGSS stages 0 and "border" were combined into CS stage 1. CS stage 3 includes fields with "moderate" damage, representing AGIS stage 3 and GSS stage 2 fields because they were described as being moderate in the original papers.[3,4] To maximize the overlap, eGSS stages 2 and 3 were consolidated and labeled as CS stage 3.

Agreement between classification systems—An important change in severity is when a patient's visual field is staged as abnormal from normal. To examine this agreement across systems, we compared the categorization of visual fields in the normal (CS1) stage versus all the other defective stages combined (CS2, CS3 and CS4). Agreement was assessed using the Kappa statistic.

For comparisons between systems across all severity stages, we used the non-parametric two-tailed Wilcoxon signed-rank matched-pair test. We also used the Spearman's rank correlation to assess how well each system's classification predicted classification in the other systems. Finally, we quantified the agreement between severity systems using linear weighted kappa statistics,[14] which is equivalent to the intraclass correlation coefficient.[15]

RESULTS

Descriptive statistics

Table 2 shows descriptive measures for MD, PSD, VFI and age. No significant difference was found between the number of OD (951) and OS (970) eyes (P=0.665). There were 1,921 eyes from 1,137 participants that met the study criteria ($7.9\pm8.4\%$ fixation losses; $2.6\pm8.4\%$ false negatives; $2.5\pm2.9\%$ false positives). Figure 1 presents four examples of visual fields staged by AGIS, GSS, and eGSS. The associated MD, PSD and VFI values are also presented. Figure 2 shows the staging for each categorical classification system according to the three global indices. The MD and PSD distributions appear broader in AGIS and GSS than in eGSS; this is likely due to the fact that eGSS relies on these two global indices in staging severity. The bubble graph shown in Figure 3 illustrates the overlap in the number of eyes classified in each combination of consolidated stages and MD, PSD and VFI categorical classification (binned data). This figure shows that there is significant overlap between the categorical staging systems (AGIS, GSS and eGSS) and the categories based on the binned MD, PSD and VFI values.

Agreement of normal versus abnormal classification of visual fields

The highest percentage of fields was classified as normal by each staging system with decreasing percentages as severity advanced (Appendix Table A3). The percentage of fields staged as normal with AGIS, GSS and eGSS were 68% (1311), 47% (905) and 69% (1314), respectively (Appendix Table A3). Agreement between AGIS and GSS in classifying a field as normal or abnormal was moderate (K=0.515±0.018). Between AGIS and eGSS, agreement was substantial (K=0.778±0.016). For example, 4.7% (91) of visual fields were categorized as having a defect by eGSS and no defect by AGIS, compared to 23.0% (441)

categorized as having no defect by AGIS and a defect by GSS (Table 3). The agreement between GSS and eGSS was moderate (K=0.516±0.018). The percentage of disagreements in categorizing a field as being normal or abnormal was 24.7% (475) (Table 3).

Comparison of global indices between severity staging systems

Table A3 (Appendix) and Figure 2 show the distribution of mean defect (MD), pattern standard deviation (PSD) and visual field index (VFI) for each severity stage for each categorical system. The percentage of fields staged as having the greatest defect with AGIS, GSS and eGSS were 0.8% (15) (stage 5), 0.8% (16) (stage 4), and 1.9% (36) (stage 5) (Appendix Table A3). No visual fields in this study were categorized in the most severe GSS stage (stage 5). This is likely due to the DIGS/ADAGES eligibility criteria, which requires successful completion of each test and a best-corrected visual acuity of 20/40 or better at baseline.

We compared the MD, PSD and VFI for visual fields staged as normal (stage 1 in the consolidated stages) across classification systems and found significant differences in all global indices between AGIS and GSS (all P<0.001), and between GSS and eGSS (P=0.007, P<0.001, P<0.001); no significant differences between global indices for AGIS and eGSS were found (PSD, P=0.796; MD, P=0.159; VFI, P=0.637)(Wilcoxon non-parametric test, 2-tailed). We also compared the MD, PSD and VFI of fields categorized in the abnormal stages across classification systems. Significant differences were found in all of these global indices between AGIS and GSS, and between GSS and eGSS (all P<0.001); between AGIS and eGSS, no significant differences were found (PSD, P=0.770; MD, P=0.051, VFI, P=0.305).

Comparing of severity staging with optic disc stereophotographs

The presence of glaucomatous optic neuropathy in the stereophotograph of an eye whose field was defective was considered a match, as was stereophotograph without glaucomatous optic neuropathy and categorization of a visual field in a normal stage. There was moderate agreement between AGIS (K= 0.457±0.022) and eGSS (K=0.459±0.22) stereophotograph assessments; agreement between stereophotograph assessment and GSS (K= 0.270±0.015) was fair. The percentage of disagreements found between stereophotograph and classification outcomes was greatest with GSS (37.9%), followed by AGIS (20.6%) and eGSS (20.4%). The incidence of glaucomatous optic neuropathy increases with increasing severity within each system (Figure 4).

Comparisons between systems across all severity stages

The percentage of eyes categorized in each consolidated stage for each classification system is presented in Figure 5. The agreement between AGIS and eGSS across consolidated stages was substantial (weighted K=0.715±0.012), and the agreement between GSS and eGSS (weighted K=0.559±0.014) and AGIS (weighted K=0.519±0.016) were moderate.

The eGSS staged abnormal visual fields more severely than GSS and AGIS, and GSS staged abnormal fields more severely than AGIS (all P<0.001). The correlations between each set of consolidated stages were as follows: AGIS and GSS, r=0.67; AGIS and eGSS, r=0.83; GSS and eGSS, r=0.68 (all P<0.001). Table 4 shows descriptive measures for the consolidated stages (1–4) by global index and staging system.

We further compared the similarity of categorical classification systems within each consolidated stage (CS, Table 1). No significant difference was found between AGIS and eGSS CS 1 (P=0.953), between GSS and eGSS CS 3 (P=0.070), and between AGIS and GSS CS 4 (P=0.397). All other combinations were significantly different (P<0.05).

DISCUSSION

In this study, we examined the severity staging of visual fields across three different categorical classification systems: the Advanced Glaucoma Intervention Study scoring system (AGIS),[3] the Glaucoma Severity Staging system (GSS),[4] and the Enhanced Glaucoma Severity Staging system (known as GSS-2, and in this paper as eGSS).[5] Each system has its own distinct advantages and disadvantages.

AGIS was developed using a relatively large sample size (n=562) for the Advanced Glaucoma Intervention Study.[3] The AGIS scoring system has been implemented in several studies.[16–17, 18] AGIS visual fields are placed within five stages of severity based on the 20-point scale. This provides a way to classify severity by categories as well as by relatively fine ordinal steps. Once the rules are automated in a software program, large volumes of visual fields can be classified. However, due to the number of parameters required for AGIS scoring (Appendix Table A1) it is difficult to use manually in the clinic.

In contrast, Brusini and Filacorda's eGSS easily implemented in the clinic on a case-by-case basis as it relies on only two indices that translate to coordinates on a plot (Appendix Figure A1). In addition to its seven stages, eGSS also characterizes the defect into three categories (localized, mixed, and generalized), a simple descriptor for relaying diagnostic information to patients and between clinicians.

GSS incorporates several parameters and characterizes the damage level in simple terms. An extended version of the Hodapp-Anderson-Parrish (HAP) scoring system,[9] GSS has one normal and five abnormal stages (rather than just four) allowing for finer distinctions than HAP during the course of disease progression. However, there are non-mutually exclusive criteria that result in ambiguity when categorizing a visual field in stage 2 versus 3, and stage 3 versus 4. For example, to meet criteria A in the pattern deviation plot for stage 2, both of the following conditions must be met: 25% to 50% points depressed below the 5% level, and 15% to 25% points depressed below the 1% level. To meet criteria A for stage 3, the following two conditions must be met: 50% to 75% points depressed below the 5% level, and 25% to 50% point depressed below the 1% level (see Appendix, Table A2). If a visual field had 52% points depressed below the 5% level and 19% points depressed below the 1% level, it would fulfill only one condition for stage 2 and only one condition for stage 3. As a result, this visual field would be placed in stage 1. This situation is shown in the field represented by the diamond symbol in Figure A1. An example of one slight adjustment to these conditions, among others, would be to require 15% (instead of 25%) to 50% points depressed below the 1% level in the second condition of criteria A for stage 3. Such a change would categorize this visual field in stage 3 and make the staging mutually exclusive.

Since there is no gold standard for staging severity, we used three global indices to characterize and compare the categorical classification systems: mean defect (MD), pattern standard deviation (PSD), and visual field index (VFI). Both MD and PSD are STATPAC measures routinely used to describe visual fields and the VFI was recently developed by Bengtsson and Heijl.[10] eGSS and AGIS were most similar, as they showed no significant difference in each of the three global indices across normal and abnormal fields. It is of interest to note that the continuous global indices are also useful indicators of severity. While these indices are readily available from STATPAC, they are based on complex statistical analyses. Choosing between continuous and categorical measures of severity depends on several factors. For example, if a clinician wants to determine whether progression has occurred since the last visit of a patient, a continuous measure may be more useful than a categorical staging system. On the other hand, if a study wants to assess the

effectiveness of a treatment at different stages of glaucoma, a categorical staging system may be more appropriate.

It is important to note that eGSS and GSS incorporate both MD and PSD, although with different degrees of dependence; and that AGIS is based on total deviation plot values. Therefore, use of MD and PSD to compare the severity classification systems may be influenced by the fact that these indices were used to compute a score on one system but not the other. However, the presence or absence of significant differences between scoring systems was consistent across these global indices.

Another difference between the systems examined here includes the number of severity stages. We devised a consolidated staging scheme so that each system had four levels of severity. Our study sample included more early stage than late-stage visual fields. Nonetheless, the scheme was based on maximizing the overlap between systems. Yet, the systems staged severity significantly differently. Staging of the abnormal fields was significantly more severe with eGSS compared to AGIS or GSS.

The results of this study should be interpreted in light of its limitations. First, there is no gold standard available for glaucoma severity. Therefore, the severity staging systems were compared to each other and to continuous measures of glaucoma severity (MD, PSD, and VFI). Second, it is not trivial to determine what constitutes a "better" severity staging system. In this cross-sectional study, all systems staged severity in an approximately similar manner; the eGSS may be better in that it is easier to use in busy clinical settings. However, other objective and quantifiable definitions of the accuracy of staging systems could be used. For example, which system correlates best with glaucomatous axonal loss and/or dysfunction, or which system correlates best with impairment in performing activities of daily living? Another definition of what makes a staging system better could be based on the reproducibility of its results. The present study wasn't designed to assess the severity staging systems on these parameters. Instead, we wanted to compare the staging systems cross-sectionally.

Using one common severity staging system would be desirable, as it would enable researchers and clinicians to directly compare diagnostic devices, treatments and other important factors mentioned earlier that are associated with this progressive disease. The systems we investigated showed increased incidence of glaucomatous optic neuropathy with increasing severity, suggesting the value in using classification systems for assessing patients. However, our study showed that each classification system could lead to different severity staging. Given our overall results and its ease of use, Brusini and Filacorda's Enhanced Glaucoma Severity Staging system may be the best choice for its ease of use for clinicians and researchers alike.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

None.

REFERENCES

1. Brusini P, Johnson CA. Staging functional damage in glaucoma: review of different classification methods. Surv Ophthalmol. 2007; 52:156–179. [PubMed: 17355855]

2. Katz J, Sommer A, Gaasterland DE, Anderson DR. Comparison of analytic algorithms for detecting glaucomatous visual field loss. Arch Ophthalmol. 1991; 109:1684–1689. [PubMed: 1841576]

3. Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. Ophthalmology. 1994; 101:1445–1455. [PubMed: 7741836]

4. Mills RP, Budenz DL, Lee PP, Noecker RJ, Walt JG, Siegartel LR, Evans SJ, Doyle JJ. Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. Am J Ophthalmol. 2006; 141:24–30. [PubMed: 16386972]

5. Brusini P, Filacorda S. Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. J Glaucoma. 2006; 15:40–46. [PubMed: 16378017]

6. Sample PA, Girkin CA, Zangwill LM, Jain S, Racette L, Becerra LM, Weinreb RN, Medeiros FA, Wilson MR, De Leon-Ortega J, Tello C, Bowd C, Liebmann JM. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. Arch Ophthalmol. 2009; 127:1136–1145. [PubMed: 19752422]

7. Anderson, DR.; Patella, VM. Automated Static Perimetry. 2 ed.. Mosby: 1998.

8. Racette L, Liebmann JM, Girkin CA, Zangwill LM, Jain S, Becerra LM, Medeiros FA, Bowd C, Weinreb RN, Boden C, Sample PA. African Descent and Glaucoma Evaluation Study (ADAGES): III. Ancestry differences in visual function in healthy eyes. Arch Ophthalmol. 2010; 128:551–559. [PubMed: 20457975]

9. Hodapp, E.; Parrish, RK., 2nd; Anderson, DR. Clinical decisions in glaucoma. 1st ed.. St. Louis: Mosby-Year Book; 1993.

10. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. Am J Ophthalmol. 2008; 145:343–353. [PubMed: 18078852]

11. Patella, VM.; Heijl, A. The Field Analyzer Primer - Essential Perimetry. 3 ed.. Dublin: Carl Zeiss Meditec; 2002.

12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

13. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960; 20:37–46.

14. Cohen J. Nominal scale agreement with provision for scaled disagreement or partial credit. Psych Bulletin. 1968; 70:213–220.

15. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as a measures of reliability. Educ Psychol Meas. 1973; 33:613–619.

16. Girkin CA, Emdadi A, Sample PA, Blumenthal EZ, Lee AC, Zangwill LM, Weinreb RN. Short-wavelength automated perimetry and standard perimetry in the detection of progressive optic disc cupping. Arch Ophthalmol. 2000; 118:1121–1236. [PubMed: 10922210]

17. Patel A, Wollstein G, Ishikawa H, Schuman JS. Comparison of visual field defects using matrix perimetry and standard achromatic perimetry. Ophthalmology. 2007; 114:480–487. [PubMed: 17123623]

18. Nouri-Mahdavi K, Hoffman D, Coleman AL, Liu G, Li G, Gaasterland D, Caprioli J. Predictive factors for glaucomatous visual field progression in the Advanced Glaucoma Intervention Study. Ophthalmology. 2004; 111:1627–1635. [PubMed: 15350314]

19. The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. Control Clin Trials. 1994; 15:299–325. [PubMed: 7956270]
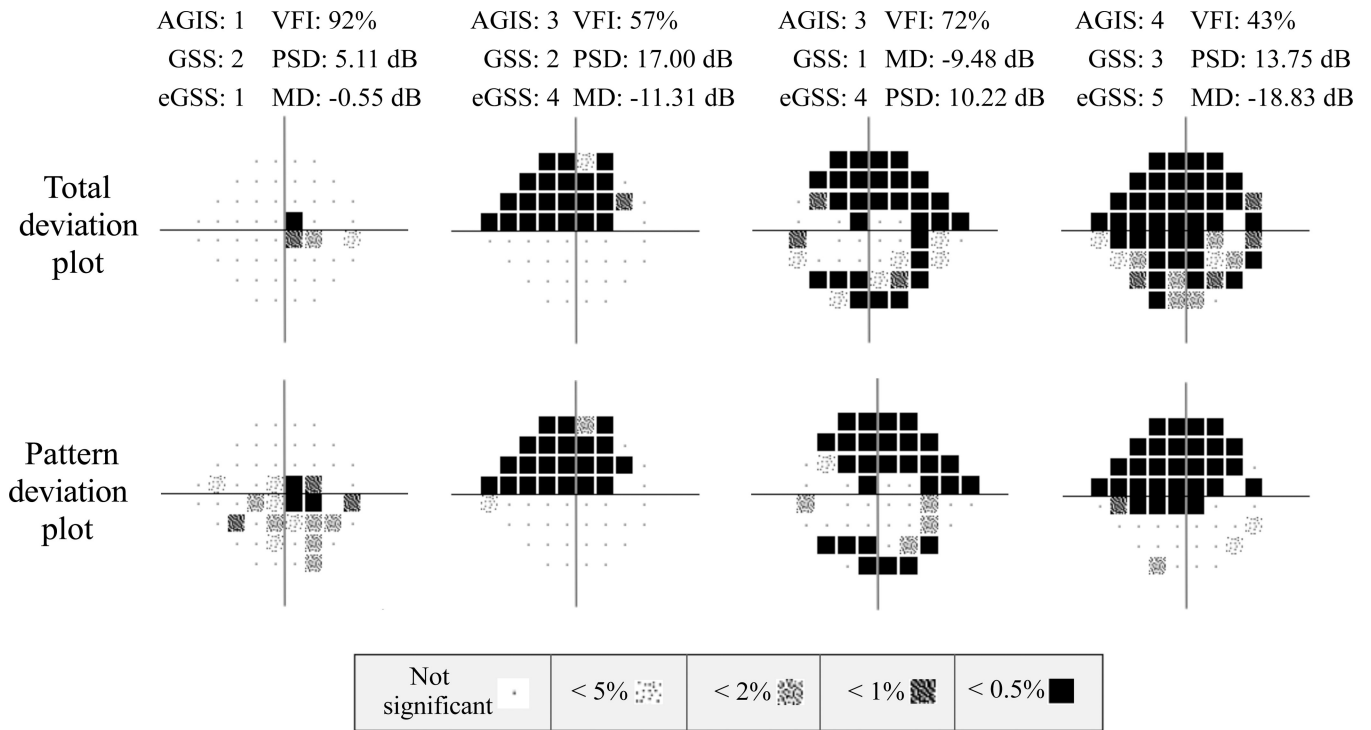
Figure 1.
Four example visual fields staged by AGIS, GSS and eGSS. The total deviation and pattern deviation plots are shown. The MD, PSD, and VFI values are also presented.
(AGIS=Advanced Glaucoma Intervention Study scoring system; GSS=Glaucoma Severity Staging system; eGSS=Enhanced Glaucoma Severity Staging system)
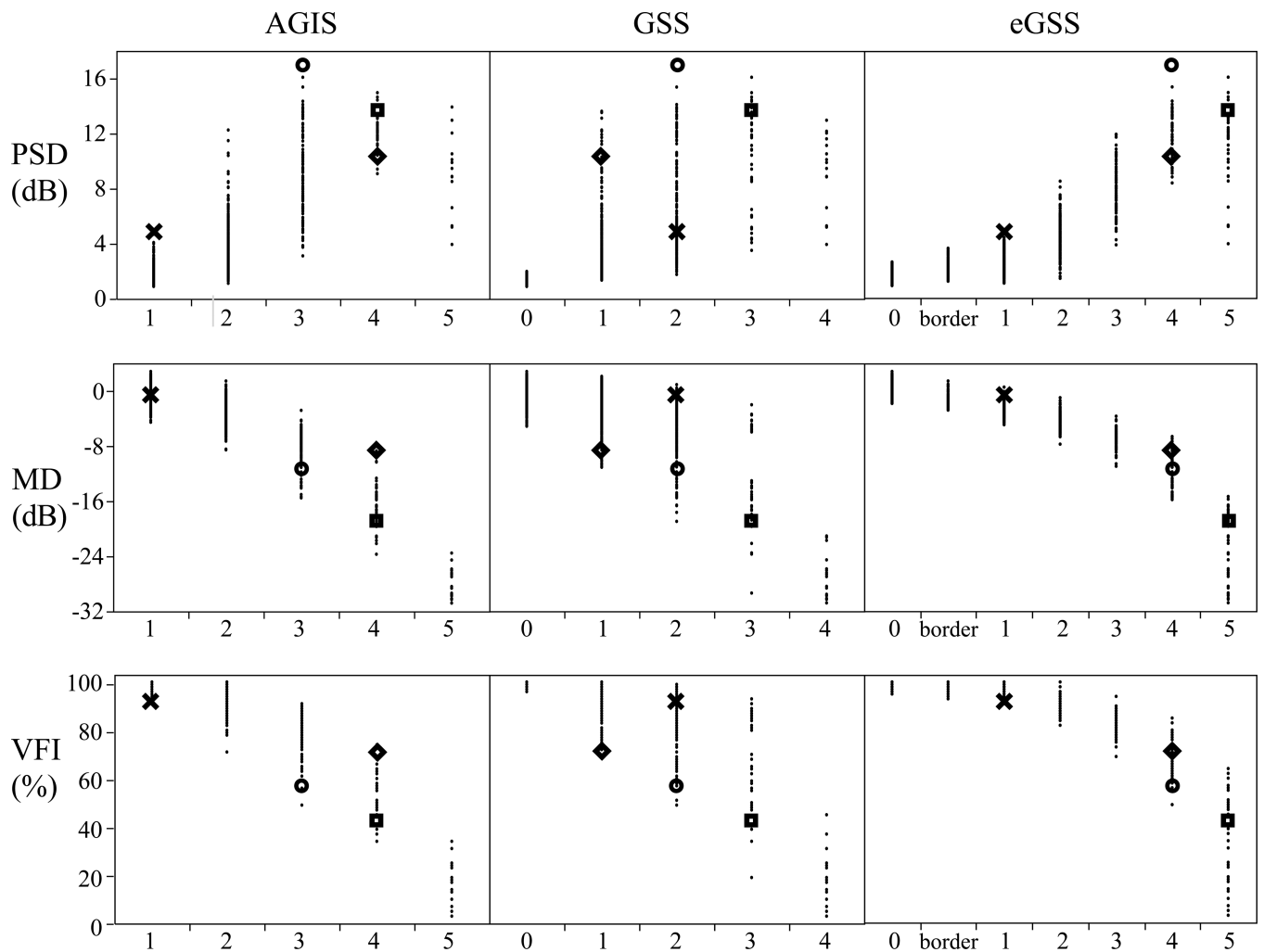
Figure 2.
Pattern standard deviation (PSD), mean defect (MD) and visual field index (VFI) distributions by severity stage according to three classification systems: Advanced Glaucoma Intervention Study scoring system (AGIS), Glaucoma Severity Staging system (GSS) and the Enhanced Glaucoma Severity Staging system (eGSS). (N=1,921 visual fields; four examples are highlighted, which correspond to examples shown in Figure 1.)
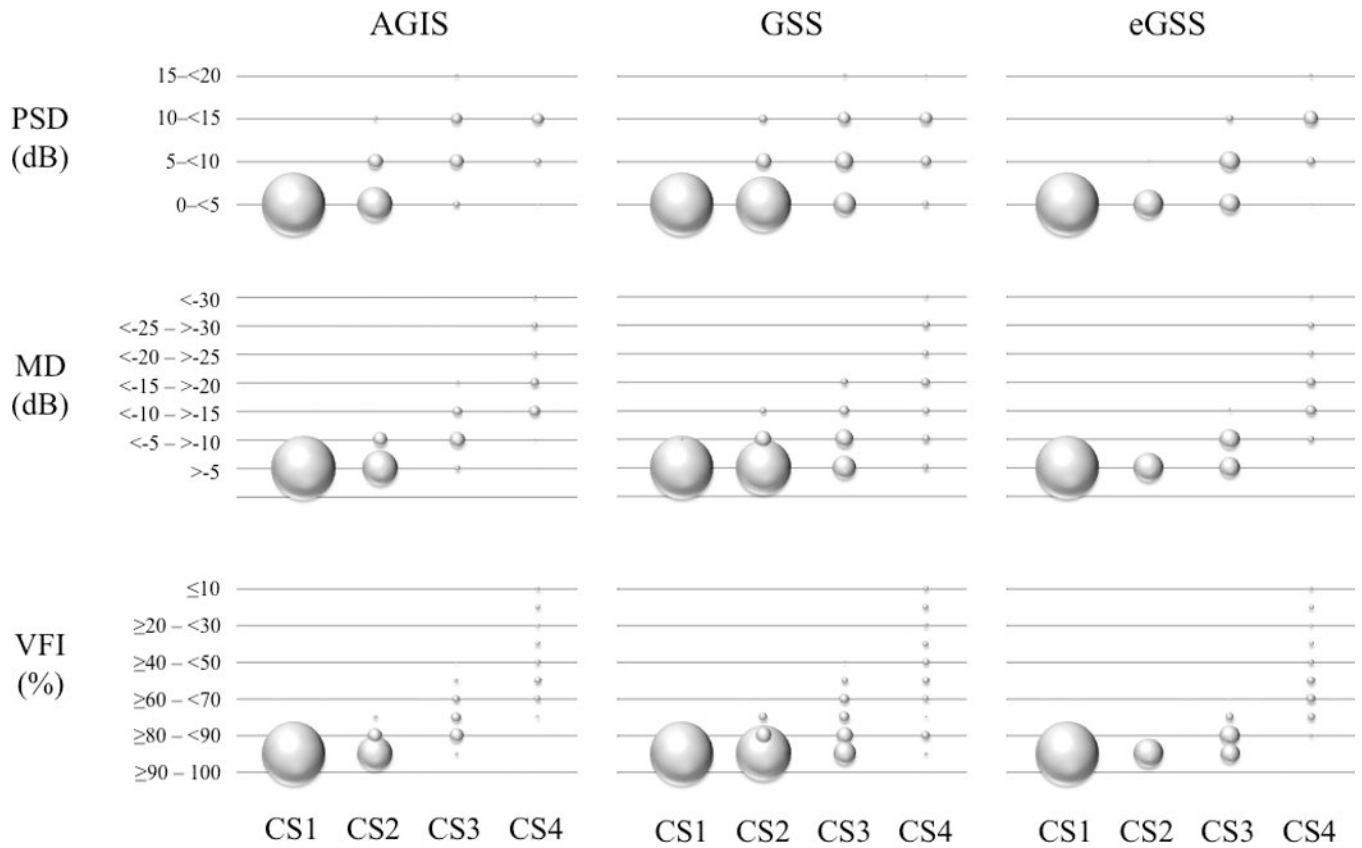
Figure 3.
A bubble graph depicting the overlap in the number of eyes in each combination of consolidated stage and MD, PSD and VFI categorical classification (based on the binned data).
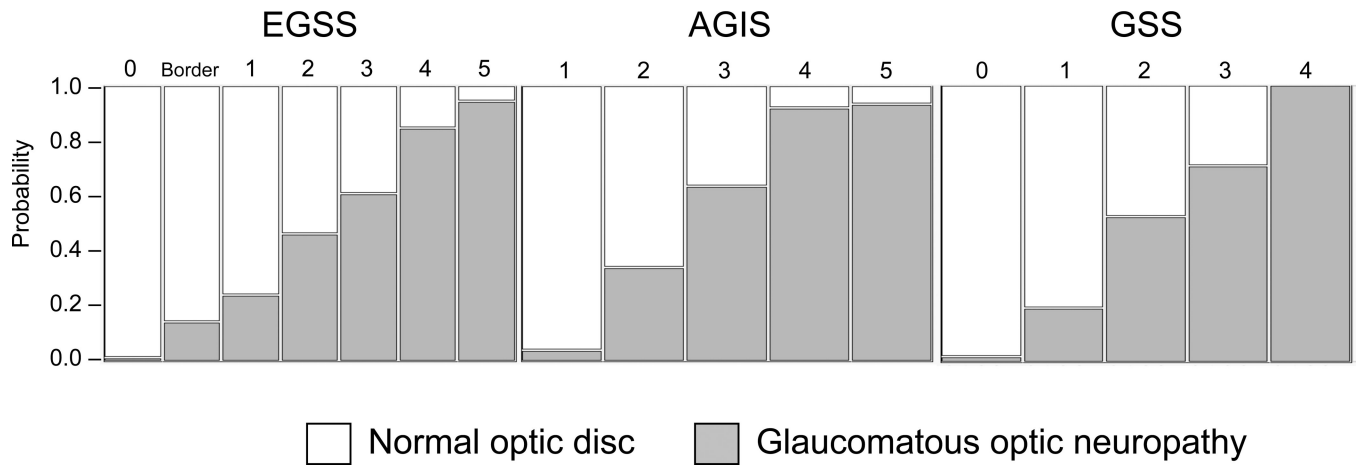
Figure 4.
The incidence of glaucomatous optic neuropathy increases with severity across all systems.
(AGIS=Advanced Glaucoma Intervention Study scoring system; GSS=Glaucoma Severity
Staging system; eGSS=Enhanced Glaucoma Severity Staging system)

AGIS-CS

3%
6%
24%
68%

GSS-CS

3%
11%
37%
47%

eGSS-CS

5%
13%
14%
69%

CS 1 ☐ CS 2 ☐ CS 3 ☐ CS 4 ■

Figure 5.
Percentage of eyes categorized by three classification systems into consolidated stages (CS). (Total study sample is N=1,921; AGIS=Advanced Glaucoma Intervention Study scoring system; GSS=Glaucoma Severity Staging system; eGSS=Enhanced Glaucoma Severity Staging system)
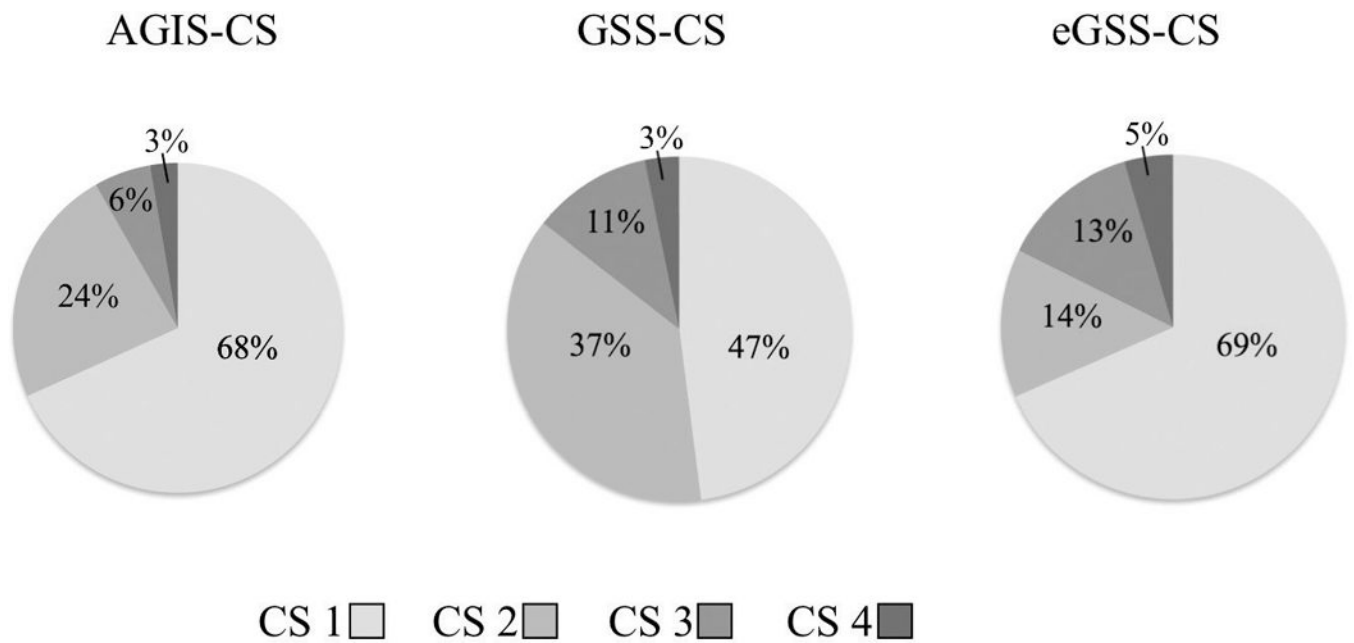
Table 1

Consolidated stages (CS) 1, 2, 3, and 4 for severity of standard automated perimetry visual fields.

Consolidate Stages		AGIS		GSS		eGSS	
CS 1	Normal	1	Normal	0	No defect	0	
						Border	
CS 2	Early/Mild Defect	2	Mild damage	1	Early defect	1	
CS 3	Moderate	3	Moderate damage	2	Moderate	2	
						3	
CS 4	Advanced/Severe/ End-Stage	4	Severe damage	3	Advanced	4	
				4	Severe		
		5	End-stage	5	End-stage	5	

Table 2

Descriptive measures of our study sample (N=1,921 fields) using three global indices: mean defect (MD), pattern standard deviation (PSD) and visual field index (VFI).

N=1,921	Average	Std Dev	Median	Range
Mean defect (MD in dB)	−2.1	3.9	−1.1	−30.8 – 2.8
Pattern standard deviation (PSD in dB)	2.8	2.5	1.8	0.9 – 17.0
Visual field index (VFI in %)	95.2	10.7	99.0	3.0 – 100.0
Age (years)	51.4	15.3	50.7	18.3 – 90.5

Table 3

Contingency table of visual fields classified as normal and abnormal using the Advanced Glaucoma Intervention Study (AGIS) scoring system, Glaucoma Severity Staging (GSS) system, and Enhanced Glaucoma Severity Staging (eGSS) system.

Count		AGIS		eGSS	
		Abnormal	**Normal**	**Abnormal**	**Normal**
Total %	Abnormal	575 29.9	441 23.0	574 29.9	442 23.0
GSS	Normal	35 1.8	870 45.3	33 1.7	872 45.4

Count		AGIS	
		Abnormal	Normal
Total %	Abnormal	516 26.9	91 4.7
eGSS	Normal	94 4.9	1220 63.5

Table 4

Descriptive measures of the global indices mean deviation (MD), pattern standard deviation (PSD) and visual field index (VFI) by severity stage for each consolidated staging (CS) system. (AGIS=Advanced Glaucoma Intervention Study scoring system; GSS=Glaucoma Severity Staging system; eGSS=Enhanced Glaucoma Severity Staging system; sd = standard deviation)

			CS1	CS2	CS3	CS4
MD	AGIS CS	mean±sd	−0.5±1.2	−3.2±1.6	−8.4±2.8	−19.7±6.0
		median	−0.5	−3.2	−7.7	−17.8
		range	−4.6 – 2.8	−8.6 – 1.4	−15.5 – −2.8	−30.8 – 9.7
	GSS CS	mean±sd	−0.4±1.3	−2.1±2.1	−5.6±3.9	−16.1±8.8
		median	−0.2	−1.7	−4.6	−15.8
		range	−5.2 – 2.8	−11.1 – 2.1	−19 – 0.9	−30.8 – −2.0
	eGSS CS	mean±sd	−0.4±1.1	−2.9±0.8	−5.2±1.7	−16.2±6.5
		median	−0.4	−3.0	−5.0	−14.9
		range	−2.8 – 2.8	−4.9 – 0.5	−11.0 – −1.0	−30.8 – −6.7
PSD	AGIS CS	mean±sd	1.7±0.5	3.5±1.6	8.9±3.2	11.1±2.5
		median	1.6	3.0	8.6	11.7
		range	0.9 – 5.1	1.1 – 12.3	3.1 – 17.0	3.9 – 15.0
	GSS CS	mean±sd	1.5±0.2	2.8±1.8	6.1±3.4	9.8±3.5
		median	1.4	2.2	4.8	10.5
		range	0.9 – 2.0	1.4 – 13.6	1.8 – 17.0	3.5 – 16.1
	eGSS CS	mean±sd	1.7±0.4	2.7±0.8	5.5±2.1	11.6±2.4
		median	1.6	2.6	5.1	12.1
		range	0.9 – 3.6	1.1 – 5.1	1.4 – 11.9	3.9 – 17.0
VFI	AGIS CS	mean±sd	99.0±1.3	94.1±3.9	79.0±9.2	44.6±20.0
		median	99.0	95.0	81.0	49.5
		range	92 – 100	71 – 100	49 – 91	3 – 74
	GSS CS	mean±sd	99.5±0.7	96.0±4.2	86.1±10.9	54.1±26.6
		median	100.0	97.0	91.0	56.0
		range	96 – 100	72 – 100	49 – 99	3 – 93
	eGSS CS	mean±sd	99.0±1.1	95.9±1.8	88.8±5.0	54.7±20.4

	CS1	CS2	CS3	CS4
median	99.0	96.0	90.0	60.0
range	93 – 100	91 – 100	69 – 100	3 – 85