



Defining glaucomatous optic neuropathy using objective criteria from structural and functional testing

Jayant Venkatramani Iyer ¹, Michael V Boland,² Joan Jefferys,² Harry Quigley ³

¹Glaucoma, Singapore National Eye Centre, Singapore, Singapore

²Wilmer Eye Institute, Johns Hopkins University Department of Ophthalmology, Baltimore, Maryland, USA

³Department of Ophthalmology, Johns Hopkins University, Baltimore, Maryland, USA

Correspondence to

Harry Quigley, Department of Ophthalmology, Johns Hopkins University, Wilmer 122, Johns Hopkins Hospital, Baltimore, MD 21287, USA; hquigley@jhmi.edu

Received 4 March 2020

Revised 25 April 2020

Accepted 15 June 2020

Published Online First

22 July 2020

ABSTRACT

Background/Aims To identify objective criteria from optical coherence tomography (OCT) and perimetry that denote a useful, specific definition of glaucomatous optic neuropathy (GON) in eyes with open-angle glaucoma for comparisons among glaucoma research studies.

Methods A cross-sectional study of adult patients with glaucoma from nine centres on five continents evaluated de-identified physician diagnosis, OCT and perimetry results for 2580 eyes (1531 patients) in an online database. Each eye was graded by their glaucoma specialist as either definite, probable or not GON. Objective measures from OCT and perimetry, derived from an online consensus panel comprising 176 glaucoma specialists globally, were compared against the three diagnostic levels.

Results Diagnoses were 54% 'definite', 22% 'probable' and 24% 'not GON'. Using only OCT data or only field data had inadequate specificity (<90%). The best definitional choice for data from either the most recent or the preceding OCT/field pair had 77% sensitivity at 98% specificity and consisted of abnormal OCT superior or inferior nerve fibre layer quadrant with matching, opposite, abnormal Glaucoma Hemifield Test.

Conclusions Objective criteria to define GON are practical and may be useful for comparisons among clinical studies to supplement subjective clinical assessment.

INTRODUCTION

Initial population-based studies^{1 2} defined glaucomatous optic neuropathy (GON) by subjective disc assessment and kinetic perimetry. To identify the prevalence of glaucoma for public health planning, data from population-based surveys were analysed in 1996,³ using standard criteria from optic disc examination and visual field testing, demonstrating that glaucoma is the second leading cause of world blindness. To better specify open-angle glaucoma (OAG) prevalence,⁴ incidence,⁵ and rates of progression and blindness across population studies, an objective definition of GON was developed by a consensus panel, including objective standards for cup/disc ratio and visual field defect.⁶ This definition had been cited 1350 times by 2017, comprising 10% of all OAG publications in a 15-year period.⁷ More recently, quantitative imaging, typically by optical coherence tomography (OCT), is now applied to glaucoma diagnosis.⁸ Yet, GON continues to be defined in many research studies by the subjective, clinician-dependent description: 'characteristic glaucoma optic disc and field change', despite proven deficiencies in interobserver and intraobserver subjective grading.^{9 10}

We initiated online discussions based on the Delphi method¹¹ to discuss approaches to an objective GON definition for use in clinical glaucoma research. In the online poll, 96% of over 100 glaucoma specialists agreed that defining study patients as having GON by unspecified subjective criteria needs to be replaced by objective methods. The present investigation generated a large, internationally representative dataset to identify objective criteria from OCT and perimetry that denote a useful, specific definition of GON in eyes with OAG for comparisons among glaucoma research studies.

METHODS

The research conformed to the Declaration of Helsinki and was approved by the Johns Hopkins School of Medicine Institutional Review Board including its lay member who represents the public, and by the review boards of the participating institutions that provided de-identified data. No patients were involved in the design or analysis. Subject consent was obtained or waived according to data collection method and instructions of the research ethics board of each participating institution. An online, expert-derived consensus process with participation by 176 experts from the American, European, Asia-Pacific, Japanese, Korean, Australian—New Zealand and Latin American Glaucoma Societies suggested possible criteria for inclusion in an objective characterisation of GON.¹² A large dataset of representative eyes was collected from participating international sites. Subjects were patients attending clinics and were similar to those who would typically be recruited in clinical research studies. Data were entered by each centre into an online data management system (REDCap).¹³ Centre clinicians clinically categorised each of their patients' eyes separately as 'definite', 'probable' or 'not GON', based on their assessment of all available data. Data were included for patients older than 21 years, for whom at least 2 reliable OCT scans and visual field tests were available, the most recent of each within 12 months, and all tests performed within a 2-year period. Eyes were excluded that had any secondary cause for GON, such as uveitic or neovascular disease. Eyes with non-OAG diseases that could cause abnormality in the structural and functional data were also excluded, as were angle closure eyes. The contributing investigators were instructed to exclude eyes whose anatomical variations could cause unreliable testing or produce abnormalities in imaging or perimetry that could be confused with GON. Eyes with substantial myopic degeneration were in this



© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Iyer JV, Boland MV, Jefferys J, et al. *Br J Ophthalmol* 2021;**105**:789–793.

excluded category. The intent was to develop a definition of GON for OAG—including primary OAG, pseudoexfoliation and pigment dispersion glaucoma—in particular, not ‘glaucoma’. Intraocular pressure was not included as a criterion for GON. GON was defined separately for each eye.

While it is the purpose of the investigation to provide an objective set of findings that replaces subjective assessment, the only available gold standard for GON is clinician diagnosis, based on available clinical and test information. Since interobserver diagnostic choices may vary, we compared how a variety of clinicians categorise eyes with similar objective findings.

The following demographic information was solicited for each patient or eye: age, sex and race/ethnicity. Structural data that was documented per eye included: OCT analyses from two tests per eye, instrument type, signal strength/quality, average retinal nerve fibre layer (NFL) thickness, rim area, disc area, vertical cup/disc ratio, retinal NFL quadrant statistical grading provided by instrument software (abnormal, borderline, normal), superior temporal two clock hour NFL statistical grading, inferior temporal two clock hour NFL statistical grading and macular OCT ganglion cell/inner plexiform thickness with number of zones superior and inferior statistically abnormal. Functional data documented included perimetric data from two 24-2 or 30-2 tests per eye with test dates, false positive %, quality grading (acceptable or not acceptable, as determined by clinician), Glaucoma Hemifield Test (GHT) result, GHT abnormal region (upper, lower, both, undetermined—as assessed by at least 3 test points abnormal at $p < 5\%$ on the pattern deviation plot), mean deviation (MD) and probability, pattern SD (PSD) and probability, number of abnormal ($p < 5\%$) non-edge points in the upper and/or lower field of the pattern deviation plot, and similar perimetric data from two 10-2 field tests, if performed. Eyes without OCT NFL and HVF 24-2 or 30-2 data were excluded. Refractive error data was not collected.

The online expert discussion pointed to the need to identify whether the upper or the lower hemifield in perimetry was abnormal and matched the abnormal position in OCT testing. All field tests were performed with Humphrey Field Analysers (HFAs, Carl Zeiss Meditec, Dublin, California, USA).

The primary outcome variables were the concordance between the clinical diagnosis of GON and the objective data from OCT and field findings, either separately or combined. We also evaluated the degree to which the clinical assessment and objective criteria differed by study centre and by physician within a centre. We excluded eyes with significant missing data and data from 4 centres with fewer than 15 qualifying eyes in ‘definite’ or ‘not GON’ groups (preventing calculation of a centre-specific sensitivity and specificity). Since both eyes of some persons were included in the overall database, we used generalised estimating equation (GEE) models with binomial distribution function and logit link function to provide estimates and 95% CIs for sensitivity and specificity that account for the correlation between two eyes of a patient.

Sensitivity for a set of criteria was estimated as the percent of eyes with ‘definite GON’ who met those criteria. Specificity for a set of criteria was estimated as percent of eyes with ‘no GON’ who did not meet those criteria. For the comparison of four qualification criteria, the model accounted for correlations within a single patient when both eyes were included. The correlation matrix specified for models was unstructured. The Tukey-Kramer method was used to adjust significance levels for multiple pairwise comparisons. Estimates and 95% CIs for sensitivity and specificity were generated from GEE models with binomial distribution function and logit link function. The sample sizes to

differentiate definite GON from not GON using one criterion compared to a different criterion were calculated to be 300 eyes per group. A similar sample size was calculated for comparisons of OCT instruments used in criterion comparisons.

All statistical analyses were performed using SAS 9.2 (SAS Institute, Cary, North Carolina, USA).

RESULTS

Demographic data

Data are presented for 1531 persons, 54% (830) women, mean age (\pm SD) of 66.2 ± 12.2 years, of whom 49% (754) were European-derived, 24% (369) Asian-derived (including Chinese, Korean and Japanese), 11% (175) African-American and 7% (100) Hispanic. Pseudophakic eyes represented 29% (820) eyes. The nine glaucoma centres providing data were from North and South America, Europe, China, Korea, Japan, Australia and New Zealand. Of OCT instruments, data came from Zeiss Cirrus (75%), Heidelberg Spectralis (16%), Nidek (6%) and Topcon (2%). HFA perimetry was from 24-2 (95%) and 30-2 (5%) test patterns, with 91% SITA Standard and 9% SITA Fast. Data from macular OCT and 10-2 field tests were entered in $< 1\%$ of eyes and these were not included in further analysis.

Characteristics of eyes based on clinician diagnosis

The clinical diagnosis for 2835 eyes was ‘definite GON’ in 54% (1539), ‘probable GON’ in 22% (615) and ‘not GON’ in 24% (681). ‘Definite GON’ diagnosed eyes had median MD of -5.55 dB and median PSD of 6.43 dB, indicating that half had mild damage by published categorisation.¹⁴ There were no significant age or sex differences between the definite GON and Not GON groups in the sample. Women represented 52% of 1539 persons among ‘definite GON’ and 56% of 681 ‘not GON’ ($p = 0.26$, χ^2 with Bonferroni adjustment). Mean age (SD) of ‘definite GON’ persons was 66.6 (12.0) years and 65.3 (12.8) years for ‘not GON’ ($p = 0.31$, Analysis of variance (ANOVA) with Bonferroni adjustment).

OCT NFL features

In the more recent of OCT tests submitted, we evaluated the superior and inferior quadrant OCT data, since the nasal and temporal NFL have lower signal/noise ratios¹⁵ and poor sensitivity/specificity in our analysis and others.¹⁶ OCT quadrant abnormality was more common inferiorly than superiorly in ‘definite GON’ (table 1); however, 9% of ‘not GON’ eyes were statistically abnormal in one or both vertical OCT quadrants.

Data for OCT NFL thickness abnormality by clock hour was evaluated as number of abnormal (red) zones in two superior temporal clock hours (11, 12 o’clock, right eyes; 12, 1 o’clock, left eyes) and in two inferior temporal clock hours (6, 7 o’clock, right eyes; 5, 6 o’clock, left eyes; table 2). Among ‘definite GON’ eyes, inferior NFL had 2 clock hours abnormal 3 times more

Table 1 Abnormality of nerve fibre layer circumpapillary quadrants on optical coherence tomography, by glaucoma diagnosis

Clinician diagnosis	Number of eyes	Abnormal superior quadrant	Abnormal inferior quadrant	Abnormal superior, inferior or both quadrants
Definite GON, n (%)	1539 (54)	863 (56)	1063 (69)	1270 (83)
Probable GON, n (%)	615 (22)	191 (31)	174 (28)	265 (43)
Not GON, n (%)	681 (24)	50 (7)	24 (4)	64 (9)

GON, glaucomatous optic neuropathy.

Table 2 Abnormality of nerve fibre layer circumpapillary clock hours on optical coherence tomography, by glaucoma diagnosis

Clinician diagnosis	Number of eyes	Abnormal superior		Abnormal inferior		Abnormal either superior or inferior		
		2 hours	1 hour	2 hours	1 hour	2 hours	1 hour	Total 2+1 hour
Definite GON, n (%)	1271	143 (11)	397 (31)	497 (39)	381 (30)	555 (44)	470 (37)	1025 (81)
Probable GON, n (%)	516	6 (1)	92 (18)	24 (5)	115 (22)	30 (6)	165 (32)	195 (38)
Not GON, n (%)	589	3 (1)	28 (5)	1 (0)	34 (6)	4 (1)	54 (9)	58 (10)

Total 2+1 hour=number of eyes with either 1 or 2 clock hours abnormal either superior or inferior. Clock hour data not available for some imaging devices.
GON, glaucomatous optic neuropathy.

often than superiorly (39% vs 11%; 497 vs 143 eyes). The 'not GON' eyes had at least one clock hour abnormal superior or inferior in 10% (58 eyes), while 81% (1025) of definite eyes met this criterion. One percent of 'not GON' eyes had 2 clock hours abnormal, so this criterion had good specificity, but its sensitivity for definite GON was only 44%.

Perimetry features

In the more recent visual field test provided, 85% of 'definite GON' eyes had GHT 'outside normal limits' with 3 points abnormal ($p < 5\%$) in the abnormal hemifield (our standard criterion), but 15% of 'not GON' also had that abnormal GHT outcome (specificity=85%; table 3). 'Definite GON' eyes with either upper or lower GHT abnormality had nearly twice the rate of defect in the upper field (29%) as in the lower field (17%), matching the corresponding greater OCT inferior defect rate.

In the number of visual field test points abnormal ($p < 5\%$), 95% of 'definite GON' eyes (1457) had ≥ 3 abnormal points in either upper or lower field, but so did 49% of 'not GON' eyes (333). The mean MD and PSD by diagnosis group had expected differences, but neither their values nor their probability of abnormality clearly separated definite from not GON eyes (table 4).

OCT NFL and perimetry combination criteria in defining GON

The best sensitivity and specificity to separate definite from 'not GON' eyes derived from the combination of OCT NFL quadrant and GHT data from each eye.

When either the most recent or the preceding pair of tests had an abnormal OCT quadrant with matching superior/inferior GHT abnormality (criterion 1), the sensitivity was 77% (1185/1539;

95% CI 75, 79%) and specificity was 98% (667/681; 95% CI 96, 99%, specificity=percentage of not GON not meeting the criterion).

Criterion 2 was tested allowing either abnormal GHT or, when GHT was normal, PSD probability $\leq 2\%$ with OCT quadrant defect in either the most recent or the preceding pair of tests. Its sensitivity was 75% (1149/1533; 95% CI 73, 77%), specificity 98% (666/679; 95% CI 96, 99%).

Criterion 3 was defined as having matching abnormal OCT quadrant and GHT abnormalities at the most recent test pair: sensitivity =73% (95% CI 70, 75%), specificity =98% (95% CI 97, 99%). Criterion 1 was significantly more sensitive than criteria 2 ($p=0.001$) and 3 ($p < 0.0001$). Only 6% (96/1539) of definite GON had normal OCT quadrants and normal GHT.

Criterion 4 was an abnormal OCT quadrant with matching GHT abnormality on both the most recent pair and preceding pair of tests: sensitivity =65% (997/1539; 95% CI 62, 67%), specificity =99% (674/681; 9% CI 98, 100%).

GEE models accounting for inclusion of both eyes had the same sensitivity, specificity and CI for all the criteria.

The specific positions of OCT/field abnormality and their concordance are shown (table 5) for the most recent test pair (criterion 3). 'Definite GON' eyes failed the criterion due to normal OCT quadrants (18%), normal GHT (8%) and non-matching GHT position (2%). Among 'probable GON' eyes, only 13% met criterion 3; those that did so had more perimetric damage than those that did not (MD=-4.2 dB vs -1.7 dB, $p < 0.0001$). The instrument used for OCT did not significantly affect sensitivity or specificity in criterion 1; sensitivity with Cirrus =78% (1126 eyes), all other OCT instruments combined =75% (413 eyes, $p=0.33$).

Table 3 Visual field Glaucoma Hemifield Test (GHT) result, most recent test, by glaucoma diagnosis

Diagnosis	N	GHT=outside normal limits				GHT borderline	GHT normal	GHT other
		Total	Upper only	Lower only	Both			
Definite GON, n (%)	1539	1315 (85)	444 (29)	255 (17)	601 (39)	84 (5)	122 (8)	18 (1)
Probable GON, n (%)	615	189 (31)	55 (9)	57 (9)	51 (8)	101 (16)	303 (49)	22 (4)
Not GON, n (%)	681	99 (15)	30 (4)	25 (4)	20 (3)	102 (15)	455 (67)	25 (4)

GON, glaucomatous optic neuropathy.

Position of GHT abnormality was not submitted in 15 definite, 26 probable and 24 not GON eyes (1-4%).

Table 4 Mean deviation and pattern SD values for the most recent perimetry, by glaucoma diagnosis

Clinical diagnosis	Mean deviation (dB)*			Pattern SD (dB)				
	N	Mean (SD)	Median	N	Mean (SD)	Median	PSD $\leq 0.5\%$	PSD $\leq 2\%$
Definite GON	1536	-7.41 (6.44)	-5.55	1539	6.94 (4.01)	6.43	1085 (70)	1273 (83)
Probable GON	614	-2.03 (2.75)	-1.42	615	2.49 (1.60)	1.94	85 (14)	187 (30)
Not GON	676	-0.96 (2.18)	-0.59	681	1.92 (1.09)	1.69	48 (7)	90 (13)

*Eyes with mean deviation $> +2$ dB (decibel) are excluded.

GON, Glaucoma Hemifield Test; PSD pattern SD.

Table 5 Abnormality of nerve fibre layer quadrant and Glaucoma Hemifield Test (GHT) by eye, by diagnosis for the most recent tests

Definite GON			
Matching GHT 633 (41%)	Superior OCT abnormal AND Abnormal GHT non-match 147 (10%)	GHT=normal or other 77 (5%)	Normal OCT superior 676 (44%)
Matching GHT 875 (57%)	Inferior OCT abnormal AND Abnormal GHT non-match 86 (6%)	GHT=normal or other 99 (6%)	Normal OCT inferior 476 (31%)
Matching GHT 1115 (73%)	Either superior or inferior OCT abnormal Abnormal GHT non-match 21 (2%)	GHT=normal or other 128 (8%)	Normal OCT superior & inferior 269 (18%)
Not GON			
Matching GHT 11 (2%)	Either superior or inferior OCT abnormal Abnormal GHT non-match 4 (1%)	GHT=normal or other 47 (7%)	Normal OCT superior & inferior 617 (91%)
Probable GON			
Matching GHT 76 (13%)	Either superior or inferior OCT abnormal Abnormal GHT non-match 3 (0%)	GHT=normal or other 174 (29%)	Normal OCT superior & inferior 350 (58%)

GON, glaucomatous optic neuropathy; OCT, optical coherence tomography.

Other=any GHT finding except outside normal limits. The position of GHT abnormality was undetermined in 6 definite GON, 2 not GON and 12 probable GON eyes.

'Definite GON' eyes meeting criterion 3 had more field loss than definite eyes with an abnormal OCT quadrant, but normal GHT (-9.0 vs -2.1 dB, $p < 0.0001$). When definite GON eyes were divided into quartiles by severity of field MD, the sensitivity was 96% for those from -30.1 to -6.6 dB (633 eyes), 75% for -6.6 to -2.8 dB (360 eyes), 55% for -2.8 to -0.8 dB (153 eyes) and 31% for -0.8 to $+2.0$ dB (37 eyes).

The predictive power of criterion 3 varied among centres, with sensitivities for definitive GON from 53% to 83%, though specificities varied minimally, from 94% to 100%. Five centres with predominantly European-derived patients had mean sensitivity for criterion 3 = 69% (individual centres = 53%, 61%, 70%, 78%, and 83%). With criterion 3, centres with Asian patients had mean sensitivity = 73% (individual centres = 66% (Chinese), 74% (Japanese) and 78% (Korean)). In three centres with multiple physicians providing clinical diagnoses, sensitivity and specificity did not significantly differ by physician (sensitivity $p = 0.26$, 0.06, 1.0, respectively; Freeman-Halton exact test and Fisher's exact test).

DISCUSSION

We identified criteria from OCT and visual field test results that can define GON objectively, facilitating standardisation of outcome comparisons across clinical glaucoma research studies. The large number of eyes studied and their worldwide distribution support the validity and generalisability of our findings. When two pairs of OCT/perimetry tests were examined after exclusion of non-glaucoma conditions that could confound test results, 77% of 'definite GON' eyes had matching OCT quadrant/field GHT abnormality in either the most recent or preceding pair (criterion 1), with 98% specificity. This criterion of diagnosis has a high positive likelihood ratio—likelihood of a person with the disease being classified as positive compared to a person without the disease—of 38.5 and a low negative likelihood ratio—likelihood of a person with the disease being classified as negative compared to a person without the disease—of 0.23. Using the most recent test pair (criterion 3), sensitivity was 73% and specificity was 98%, with positive and negative likelihood ratios of 36.5 and 0.28, respectively. Use of either OCT data alone or perimetric data alone had 80–90% sensitivity, but specificities

much lower than the combined OCT/perimetry criteria. Individual OCT NFL clock hour abnormality or individual test point abnormality in perimetry were also relatively non-specific. Sensitivity for the objective criteria varied among centres, suggesting that physicians differ in their subjective diagnosis of GON, even when objective data are the same. These variabilities support the need for an objective definition.

Published studies have used methods of artificial intelligence to separate GON from normal eyes.¹⁶ In these studies too, clinical diagnosis was the gold standard. Medeiros and colleagues¹⁷ used deep learning analysis of NFL circle scan data in 612 OAG eyes, reporting 81% sensitivity at 95% specificity for a clinical diagnosis of GON, outperforming either global or sectoral NFL thickness values. All their subjects had an abnormal GHT. At 98% specificity, their sensitivity was below that of our criteria 1 or 2. Deep learning was applied to OCT macular and NFL data in 86 OAG eyes classified clinically by disc and NFL exam alone, achieving 95% sensitivity/100% specificity.¹⁸ While such approaches are promising, the precise criteria separating groups are often not determined and may be relatively population-specific.

Those clinically diagnosed as definite GON who met one of our criteria had worse mean MD than definite GON eyes that did not. Sensitivity was 96% for MD from -6 to -30 dB, and 75% for MD from -2 to -6 . Sensitivity was lower with MD in the normal range and 6% (96 eyes) diagnosed as definite GON had normal OCT and GHT. It is inevitable that some early GON eyes will be difficult to specify and further study of criteria to identify it are being sought. An approach that adds macular OCT and central visual field tests to NFL scans and 24-2 field tests has suggested good predictive power in 53 early GON eyes.¹⁹ While the combined OCT and perimetry criteria presented here may fail to include some eyes considered to have early glaucoma damage, their strength is objectivity, high specificity, simple implementation and widespread availability.

The patients recruited for this study are typical for those seen in glaucoma clinics, but are not necessarily representative of a population-based sample of those with OAG. Our purpose was to specify possible definitions for use in clinical research. Patients in glaucoma clinics, as included here, are typical for those

recruited for such investigations. The eyes coded as not GON were often glaucoma suspects due to family history of glaucoma, history of ocular hypertension or suspect optic disc findings. They are not population-based normals, but since they have some potential risk factors for OAG, the high specificity of our criteria indicates that the predictive power of the approach might be even greater among population normals.

There are issues unresolved by the present investigation. Field testing was conducted on the 24-2 and 30-2 programmes of HFA instruments. It will be important to test whether field testing on other instruments or using central perimetric algorithms would complement the criteria. While use of different OCT instruments may affect implementation of objective GON criteria, our analysis found no significant difference in the sensitivity for identifying GON by structural features between >1100 eyes tested with the Cirrus device and over 400 eyes tested with three other devices. As technology advances in hardware and software, such interinstrument comparisons will continue to be important. Interocular asymmetry of optic nerve imaging or perimetry might be useful, but presently lacks normative values. Our subjects were generally experienced at perimetry and may have higher reliability than new OAG subjects; however, those selected for glaucoma research are typically screened to have reliable testing. The normative databases in OCT and visual field instruments exclude many persons who have unusual discs or high refractive errors. While these subjects are also often excluded in glaucoma research studies, specific normative databases for myopes and by derivation group are now being constructed. This research is not intended to be relevant to medical insurance reimbursement or treatment decisions.

This study aimed to determine criteria that could best be used to diagnose GON in OAG eyes at a single point of time. Progressive change in OCT or field is an important feature of GON, but occurs slowly in the majority of eyes.²⁰ Few centres have extended, standardised progression data on those eligible to participate in research studies.²¹ It will be important to compare the objective features identified here as denoting GON as classified by clinician-scientists to the eyes with documented progressive change in one or more of their parameters. There is no consensus on which single or combination of algorithms best denotes progressive change,²² which would be a prerequisite for such comparison.

In summary, this large international database provides useful, objective criteria that identify three-fourth of those identified as GON by experts across a range of derivation groups, while effectively excluding those without GON. In ongoing research,²³ multiple experts will grade each eye as GON on a 0–100 likelihood scale, after reviewing colour fundus photographs, NFL and macular OCT, and standard and central visual field testing. These approaches will help to determine objective data that can provide standards that would be available to compare research outcomes among glaucoma investigations.

Correction notice This paper has been corrected since it was published online. Table 5 and the acknowledgements have been updated.

Acknowledgements We are grateful for the generous participation of the following international glaucoma specialists and collaborators in this study: Ramanjit Sihota and Dewang Angmo, All India Institute of Medical Sciences, New Delhi, India; Don Hood and Gustavo de Moraes, Columbia University, New York, USA; Poemen Chan, Clement C Tham, Annie Ling and Marco C Pang, Chinese University of Hong Kong, Hong Kong, China; Jayme Vianna and Michael West, Dalhousie University, Halifax, Canada; Makoto Araie, Kazuhisa Sugiyama, Tomomi Higashide, Ryo Asaoka, Atsuya Miki, Sachiko Udagawa and Aiko Iwase, Japan Glaucoma Society, Japan; Harry Quigley, Jayant V Iyer and Michael V Boland, Wilmer Eye Institute, Baltimore, USA; Ravi Thomas, Queensland Eye Institute, Queensland, Australia; Ahnul Ha and Ki Ho Park, Seoul National University Hospital, Seoul, South Korea; Helen Danesh-Meyer, Elkie Wong, Linda Rodriguez and Hannah Kersten, University of Auckland, Auckland, New Zealand; Diane Ngo, Joseph

Caprioli, Alessandro Rabiolo and Esteban Morales, University of California Los Angeles, USA; Dario Romano, Emanuele Maggiolo, Luca Rossetti, Roberta Farci and Giovanni Esposito, University of Milano, Milan, Italy; Fernando Gomez, New Granada Military University, Bogota, Colombia; Cassandra Barnhart, University of North Carolina, Chapel Hill, USA; Jeffrey Chou and Murray Fingeret, VA New York, USA; and Joel Schuman, Gadi Wollstein, Katie Lucy and Rachel Anderson, New York University, New York, USA. We are also grateful to Dr Htoon Hla Myint from Singapore Eye Research Institute, Singapore, for his assistance in provisional analysis of the initial data.

Contributors JVI and HQ were involved in the conception, study design, data acquisition and analysis, and manuscript preparation. MVB was involved in study design, data acquisition and manuscript preparation. JJ was involved in data analysis and in manuscript preparation.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs

Jayant Venkatramani Iyer <http://orcid.org/0000-0002-2397-7100>
Harry Quigley <http://orcid.org/0000-0003-0095-8756>

REFERENCES

- Hollings FC, Graham PA. Intra-ocular pressure, glaucoma, and glaucoma suspects in a defined population. *Br J Ophthalmol* 1966;50:570–86.
- Tielsch JM, Sommer A, Katz J, et al. Racial variations in the prevalence of primary open angle glaucoma: the Baltimore Eye Survey. *JAMA* 1991;266:369–74.
- Quigley HA. The number of people with glaucoma worldwide. *Br J Ophthalmol* 1996;80:389–93.
- Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology* 2014;121:2081–90.
- Quigley HA, Vitale S. Models of open-angle glaucoma prevalence and incidence in the United States. *Invest Ophthalmol Vis Sci* 1997;38:83–91.
- Foster PJ, Buhrmann R, Quigley HA, et al. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol* 2002;86:238–42.
- Google scholar data. (accessed 8 Aug 2017).
- Schuman JS, Hee MR, Arya AV, et al. Optical coherence tomography: a new tool for glaucoma diagnosis. *Curr Opin Ophthalmol* 1995;6:89–95.
- Coleman AL, Sommer A, Enger C, et al. Interobserver and intraobserver variability in the detection of glaucomatous progression of the optic disc. *J Glaucoma* 1996;5:384–9.
- Werner EB, Bishop KI, Koelle J, et al. A comparison of experienced clinical observers and statistical tests in detection of progressive visual field loss in glaucoma using automated perimetry. *Arch Ophthalmol* 1988;106:619–23.
- Available <https://www.rand.org/topics/delphi-method.html>
- Iyer JV, Vianna JR, Chauhan BC, et al. Toward a new definition of glaucomatous optic neuropathy for clinical research. *Curr Opin Ophthalmol* 2020;31:85–90.
- Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
- Mills RP, Budenz DL, Lee PP, et al. Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. *Am J Ophthalmol* 2006;141:24–30.
- Quigley HA. Examination of the retinal nerve fiber layer in the recognition of early glaucoma damage. *Trans Am Ophthalmol Soc* 1986;84:920–66.
- Maetschke S, Antony B, Ishikawa H, et al. A feature agnostic approach for glaucoma detection in OCT volumes. *PLoS One* 2019;14:e0219126.
- Thompson AC, Lammal AA, Berchuck SI, et al. Assessment of a segmentation-free deep learning algorithm for diagnosing glaucoma from optical coherence tomography scans. *JAMA Ophthalmol* 2020;138:333.
- Lee J, Kim YK, Park KH, et al. Diagnosing glaucoma with spectral-domain optical coherence tomography using deep learning classifier. *J Glaucoma* 2020;29:287–94.
- Hood DC, Tsamis E, Bommakanti NK, et al. Structure-function agreement is better than commonly thought in eyes with early glaucoma. *Invest Ophthalmol Vis Sci* 2019;60:4241–8.
- Chauhan BC, Malik R, Shuba LM, et al. Rates of glaucomatous visual field change in a large clinical population. *Invest Ophthalmol Vis Sci* 2014;55:4135–43.
- Ohnell HM, Bengtsson B, Heijl A. Making a correct diagnosis of glaucoma: data from the EMGT. *J Glaucoma* 2019;28:859–64.
- Vianna JR, Chauhan BC. How to detect progression in glaucoma. *Prog Brain Res* 2015;221:135–58.
- Vianna JR, Chauhan BC. Pilot analysis of the web-application and inter-evaluator variability for the crowd-sourced glaucoma study. abstract, ARVO 2019 annual meeting. *Invest Ophthalmol Vis Sci* 2019;60:6152.