



Non-physician grader reliability in measuring morphological features of the optic nerve head in stereo digital images

Victoria Addis¹ · Enny Oyeniran¹ · Ebenezer Daniel¹ · Rebecca Salowe¹ · Richard Zorger¹ · Roy Lee¹ · Maxwell Pistilli¹ · Maureen Maguire¹ · Qi Cui¹ · Eydie Miller-Ellis¹ · Joan M. O'Brien¹ · Prithvi S. Sankar¹

Received: 26 June 2018 / Revised: 9 November 2018 / Accepted: 14 December 2018 / Published online: 11 January 2019
© The Royal College of Ophthalmologists 2019

Abstract

Objective To introduce a new method of grading optic nerve stereo disc photographs and evaluate reproducibility of assessments by non-physician graders in a reading center.

Methods Three non-physician graders, experienced in grading features of the retina but not the optic nerve head (ONH), were trained by glaucoma specialists to assess digital stereo color images of the ONH. These graders assessed a total of 2554 digital stereo disc images from glaucoma cases and controls participating in the Primary Open-Angle African American Glaucoma Genetics (POAAGG) study by outlining the optic cup and disc. Inter-grader reproducibility of area, height, and width measurements was analyzed.

Results Among all images, the intraclass correlation (95% confidence interval) was 0.90 (0.89, 0.90) for the cup area using only color cues; 0.92 (0.91, 0.92) for the cup area using contour and vascular cues; and 0.99 (0.99, 0.99) for the optic disc area. The intraclass correlation for cup-to-disc ratio (CDR) was 0.61 (0.58, 0.63), as determined by the ratio of optic cup area to optic disc area (using contour and vascular cues). The CDR difference by graders for area was ≤ 0.1 in 65% of images using color/vascular cues and ≤ 0.1 in 71% of images using color cues.

Conclusions After adequate training, non-physician graders were able to measure the optic nerve CDR with high inter-grader reliability.

Introduction

Glaucoma is an optic neuropathy characterized by progressive loss of retinal ganglion cells, often manifesting as changes in the intra-papillary and papillary regions of the optic disc and retinal nerve fiber layer [1]. This disease affects 44 million individuals worldwide, with a projected prevalence of 53 million cases by 2020 and 80 million cases by 2040 [2]. The evaluation of structural changes is central to the diagnosis and management of patients with glaucoma [3]. Although the appearance of the optic nerve was first described nearly 150 years ago [4], analysis of

optic nerve head (ONH) features in glaucoma remains challenging.

A number of evaluation schemes have emerged that attempt to characterize the phenotype of the ONH, with the most common being cup-to-disc ratio (CDR). Although very widely used, CDR is limited by high degrees variability in grading among ophthalmologists [5–7]. Many researchers have developed other methods of optic nerve assessment, attempting to improve inter-observer agreement. In 1974, Shiose proposed three patterns of optic disc damage, each with six stages of structural changes [8]. Read and Spaeth, also in 1974, coupled stages of increasing CDR with worsening visual field loss [9]. The *Glaucoma and Glaucoma Suspects* grading system built upon this work by taking into account worsening CDRs, visual field changes, and neuroretinal rim asymmetry, as well as disc pallor [10, 11]. Nesterov proposed a five stage system characterized by optic disc excavation, temporal slope, depth, and CDR [12]. The *Rim to Disc (R/D) Method* shifted the focus away from cupping to the neuroretinal rim [13]. Finally, the *Disc Damage Likelihood Scale (DDLs)* proposed ten stages

These authors contributed equally: Joan M. O'Brien, Prithvi S. Sankar

✉ Joan M. O'Brien
joan.o'brien@uphs.upenn.edu

¹ Scheie Eye Institute, University of Pennsylvania, Philadelphia, PA, USA

of glaucomatous progression based on an estimation of the neuroretinal rim in any position and the circumferential extent of its absence [14].

Although each has its advantages, these grading systems remain less than optimal. The systems do not always allow an examiner to easily recognize a pathologic optic nerve, determine the severity of damage, and monitor for evidence of glaucomatous progression. Furthermore, like CDR, many of the systems yield inconsistent or variable results among examiners [6, 15–18].

A variety of imaging and automated instruments such as the scanning laser polarimetry [19], scanning laser ophthalmoscope [20], and optical coherence tomography [21, 22] have also been used to evaluate the ONH in glaucoma. Despite these advances, the subjective method of grading stereo images of the ONH remains important. Automated instruments have been shown to miss both early cases of glaucomatous damage and severe cases of glaucoma; they are also more limited when facing anatomic variation or suboptimal images [23, 24]. For these reasons, European Optic Disc Assessment Trial recommended that automated devices be used to “support rather than replace skilled clinic examination” [24].

Thus, there remains a need for a reproducible method for grading and classifying optic nerve disc photos. Ideally, this methodology would be more comprehensive and precise than the existing subjective methods described above, while avoiding the pitfalls of automated instruments. Such a grading system could be used to assess progression over time or to categorize phenotypes of the optic nerve in glaucoma for research and genetic studies.

To address this need, we used digital stereo images to develop a new quantitative method of assessing the ONH, which allows computation of many features associated with glaucoma and illustrates the inter-grader reliability of measuring CDR. This method was employed on more than 2554 images from the Primary Open-Angle African American Glaucoma Genetics (POAAGG) study, which is a 5-year project investigating the genetics of primary open-angle glaucoma (POAG) in African Americans. In this paper, we describe our methodology and report concordance rates among non-physician trained graders.

Methods

Participants

The POAAGG study population consists of self-identified African Americans and individuals of African or Afro-Caribbean descent, over age 35, recruited from the Philadelphia region. Exclusion criteria have been previously

published [25]. The examining ophthalmologist recorded the vertical CDR. Fellowship-trained glaucoma specialists determined each subject’s classification as a glaucoma case, control, or suspect based on previously published criteria [25]. For this study, 30 degree stereo disc photos taken, using the Topcon TRC 50EX retinal camera (Topcon Corp. of America, Paramus, NJ), from 2554 eyes of glaucoma cases and controls were analyzed. These images were received by graders between 02/09/2016 and 01/31/2018. The University of Pennsylvania Institutional Review Board approved this study and the informed consent process, and this research followed the tenets of the Declaration of Helsinki.

Non-physician trained graders

Three non-physician graders, from the Ophthalmology Reading Center at the University of Pennsylvania, were trained by two fellowship-trained glaucoma specialists to grade digital stereo color images of the optic disc. These graders were experienced in grading the retina in other studies (such as the Comparison of Age-Related Macular Degeneration Treatments Trials) [26], but not the ONH. Prior to beginning the study, all graders were tested and found to have stereo vision of 40 s of arc. Graders were trained to use the stereo viewer (Screen-Vu stereoscope, Portland, OR). Training sessions occurred weekly for 2 h a week for 5 months. The graders were given “practice” optic nerve images to grade between sessions and these images were reviewed during the weekly meetings with the glaucoma specialists and Director of the Reading Center. The outlined parameters were each drawn while the graders were actively using the stereo viewer.

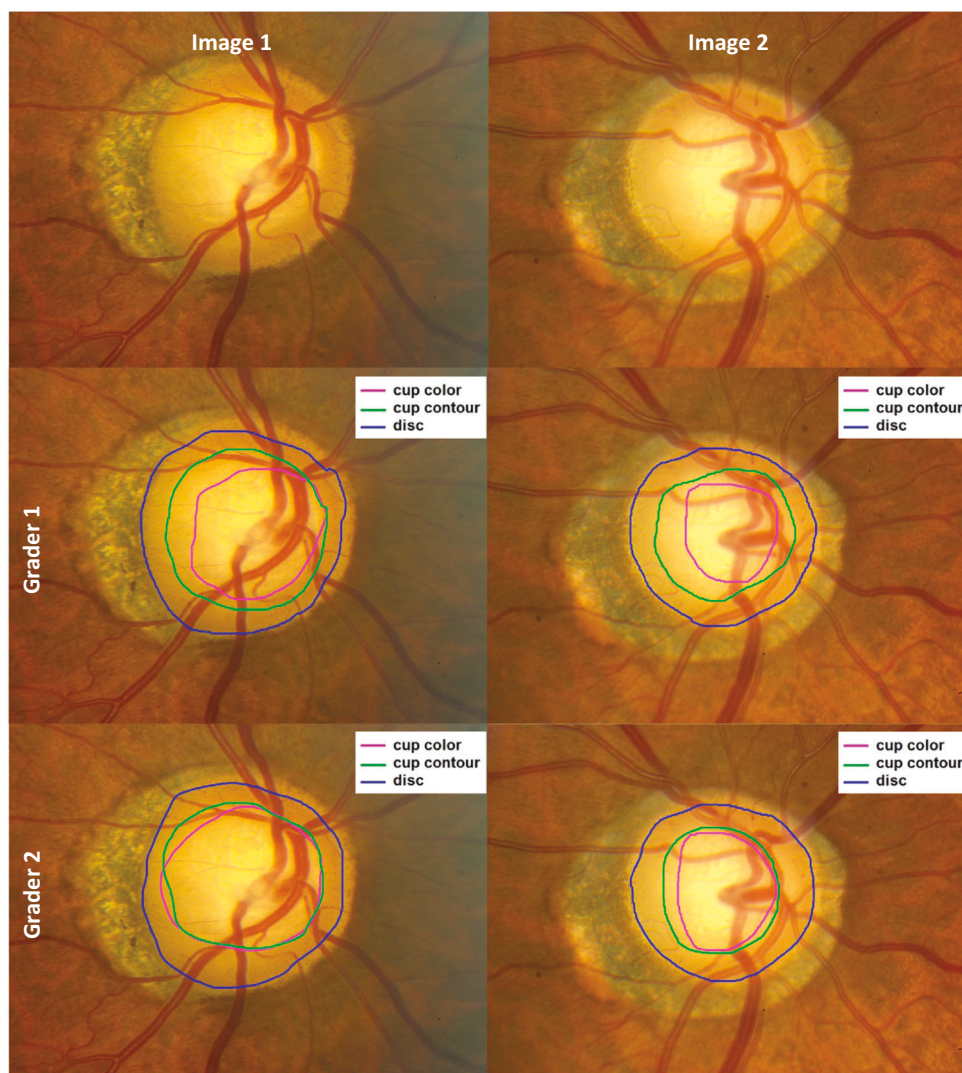
Optic disc analysis

After completing training, each digital optic nerve photograph was analyzed by two of the three trained graders. Graders were masked as to whether the image was from a glaucoma case or a control. The graders were asked to outline three structures on each optic nerve photograph using the Image J/Fiji software (available at <http://rsbweb.nih.gov/ij/>; Rasband WS, Image J, US National Institutes of Health, Bethesda, MD, 1997e2012) (Fig. 1).

Outlined structures included:

1. The optic cup using only color and pallor cues from the photograph (“color cup”).
2. The optic cup using contour and vascular cues (“contour cup”).
3. The optic disc, defined as the outer border of the nerve rim and the inner border of the scleral ring, if a scleral ring was present.

Fig. 1 Examples of stereo disc photos outlined for cup color, cup contour, and disc by two graders. For Image 1, areas were 72,070 (Grader 1) and 94,006 (Grader 2) for cup color; 108,667 (Grader 1) and 98,751 (Grader 2) for cup contour; and 174,421 (Grader 1) and 172,310 (Grader 2) for the disc. CDR was 0.41 (Grader 1) and 0.55 (Grader 2) using cup color, and 0.62 (Grader 1) and 0.57 (Grader 2) using cup contour. For Image 2, areas were 39,859 (Grader 1) and 48,728 (Grader 2) for cup color; 75,195 (Grader 1) and 63,352 (Grader 2) for cup contour; and 135,916 (Grader 1) and 133,663 (Grader 2) for the disc. CDR was 0.29 (Grader 1) and 0.36 (Grader 2) using cup color, and 0.55 (Grader 1) and 0.47 (Grader 2) using cup contour. Area units are unscaled pixels



The areas within each of these measurements, as well as the height and width of these measurements, were then calculated using the Image J/Fiji software. The software calculated the height and width based on the vertical and horizontal axes, respectively; axes were determined by the software. CDR was then calculated as the ratio of the area, height or width of the optic cup (using either the “color cup” or “contour cup”) to the area, height or width of the optic disc. The graders were not involved in any clinical diagnoses.

Statistical analysis

Intraclass correlations coefficients and 95% confidence intervals were calculated for the color cup, contour cup, disc, color CDR, and contour CDR measurements. Intraclass correlation coefficient is a measure of reliability that describes the consistency or reproducibility of quantitative measurements (i.e., optic nerve grading) made by different

observers (i.e., two graders). Reliability ranges from 0 to 1, where 1 indicates that there is no grader error (i.e., no differences in grading between graders) and 0 means that all variability across eyes is attributable to grader error. Higher reliability values indicate less grader error. These were calculated separately using the area, height, and width of the drawings, and for the area CDR calculations the square roots of the cup and disc areas were used. The CDR value used for comparison to the clinical assessment of CDR was the average of the values from two graders. All analysis was performed in SAS v9.4 (Cary, NC).

Results

A total of 2554 digital stereo images were analyzed for this study, including 1984 images from POAG cases and 570 images from controls. The intraclass correlation (95% confidence interval) for agreement between the grading

value of CDR height and the clinical value of vertical CDR was 0.72 (0.63, 0.79) by color pallor cues and 0.71 (0.63, 0.79) by contour.

Among all images, the intraclass correlation (95% confidence interval) for agreement between graders was 0.90 (0.89, 0.90) for the cup area using only color/pallor cues; 0.92 (0.91, 0.92) for the cup area using contour and vascular cues; and 0.99 (0.99, 0.99) for the disc area (Table 1). Intraclass correlations among all subjects were highest when using height (versus area or width). The intraclass correlation for the area of the CDR was 0.74 (0.73, 0.76) when using “color cup” and 0.61 (0.58, 0.63) when using “contour cup.”

The intraclass correlations were stratified by whether the stereo disc images were from glaucoma cases versus controls (Table 1). The intraclass correlations were slightly higher among cases than controls for the color cup, the contour cup, and the optic disc. The cup-to-disc area ratio using the contour cup also had a slightly higher intraclass correlation among cases (0.57 [0.54, 0.60]) than controls (0.35 [0.27, 0.42]). This lower concordance among controls was also seen when height and width measurements were used in the CDR measurements.

The difference between CDR values assigned by the two graders was calculated using the color cup and contour cup for area, height, and width measurements (Table 2 and Fig. 2). Using color cups, the CDR difference by area between graders was ≤ 0.1 in 71% of images and ≤ 0.2 in 94% of images; using contour cups, the difference by area was ≤ 0.1 in 65% of images and ≤ 0.2 in 92% of images. Again, using height measurements cups yielded the highest reproducibility (compared with area or width), with difference between graders of ≤ 0.1 in 75% of images for color cups and ≤ 0.1 in 68% of images for contour cups. Reproducibility was higher among cases when using contour cups and higher among controls when using color cups (Table 3).

Discussion

Our study introduced a new quantitative method of assessing the ONH and deriving CDR using digital stereo images. Non-physician graders trained by glaucoma specialists were capable of grading and measuring ONH parameters with high inter-grader reliability. Graders had very low discordance

Table 1 Intraclass correlations between trained graders based on area, height, and width of measurements

	Number of photos	Drawing	AREA Intraclass correlation	HEIGHT Intraclass correlation	WIDTH Intraclass correlation
Glaucoma cases	1984	Color cup	0.88 (0.87, 0.89)	0.91 (0.90, 0.92)	0.85 (0.84, 0.86)
		Contour cup	0.91 (0.90, 0.92)	0.92 (0.91, 0.92)	0.86 (0.85, 0.87)
		Disc	0.99 (0.99, 0.99)	0.98 (0.98, 0.99)	0.99 (0.99, 0.99)
		CDR (using color cup)	0.68 (0.66, 0.70)	0.73 (0.71, 0.75)	0.60 (0.57, 0.63)
		CDR (using contour cup)	0.57 (0.54, 0.60)	0.63 (0.60, 0.65)	0.44 (0.41, 0.48)
Controls	570	Color cup	0.84 (0.81, 0.86)	0.84 (0.81, 0.86)	0.78 (0.74, 0.81)
		Contour cup	0.82 (0.79, 0.84)	0.79 (0.75, 0.82)	0.69 (0.65, 0.73)
		Disc	0.99 (0.99, 0.99)	0.97 (0.97, 0.98)	0.97 (0.97, 0.98)
		CDR (using color cup)	0.62 (0.56, 0.67)	0.69 (0.65, 0.73)	0.50 (0.44, 0.56)
		CDR (using contour cup)	0.35 (0.27, 0.42)	0.43 (0.36, 0.49)	0.20 (0.12, 0.28)
All optic nerve images	2554	Color cup	0.90 (0.89, 0.90)	0.93 (0.92, 0.93)	0.87 (0.86, 0.88)
		Contour cup	0.92 (0.91, 0.92)	0.92 (0.92, 0.93)	0.87 (0.86, 0.88)
		Disc	0.99 (0.99, 0.99)	0.98 (0.98, 0.99)	0.99 (0.99, 0.99)
		CDR (using color cup)	0.74 (0.73, 0.76)	0.79 (0.78, 0.81)	0.65 (0.63, 0.68)
		CDR (using contour cup)	0.61 (0.58, 0.63)	0.67 (0.65, 0.69)	0.47 (0.44, 0.50)

CDR cup-to-disc ratio

Table 2 Differences between CDR values assigned by two trained graders using color and contour cup for area, height, and width measurements, for all patients

Difference	Color			Contour		
	Area	Vertical	Horizontal	Area	Vertical	Horizontal
≤ 0.1	1826 (71%)	1912 (75%)	1632 (64%)	1655 (65%)	1740 (68%)	1393 (55%)
$>0.1 - \leq 0.2$	587 (23%)	510 (20%)	680 (27%)	686 (27%)	610 (24%)	747 (29%)
$>0.2 - \leq 0.3$	104 (4%)	98 (4%)	178 (7%)	171 (7%)	154 (6%)	318 (12%)
>0.3	37 (1%)	34 (1%)	64 (3%)	42 (2%)	50 (2%)	96 (4%)

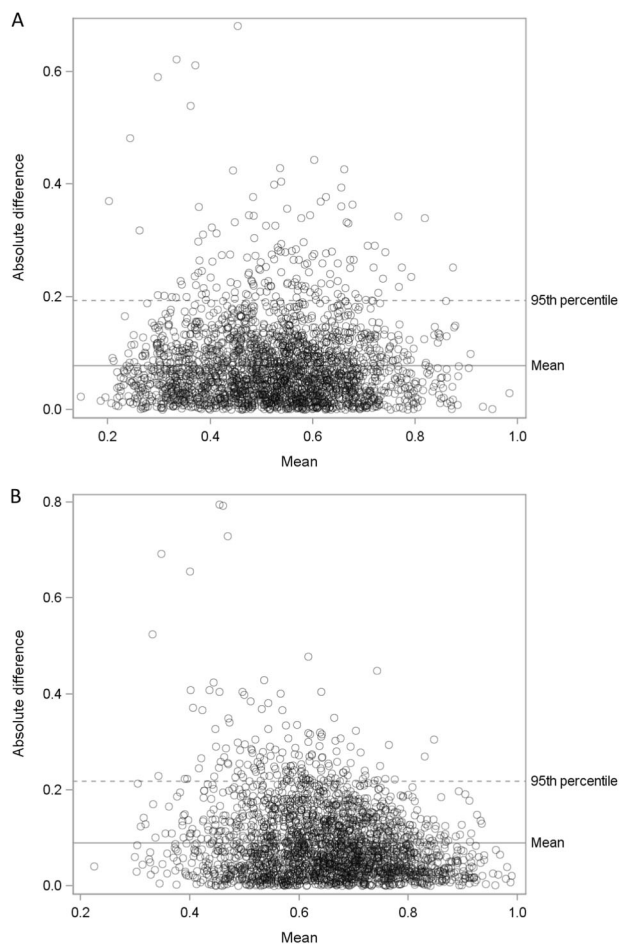


Fig. 2 Bland-Altman plots for cup-to-disc ratio for color and contour grading. **a** Bland-Altman plot of the absolute value of the difference between two graders of the cup-to-disc ratio by the mean of the two values when using the area of the cup as defined by color and the area of the disc. The mean absolute difference was 0.078, and the 95th percentile of the differences is 0.193. **b** Bland-Altman plot of the absolute value of the difference between two graders of the cup-to-disc ratio by the mean of the two values when using the area of the cup as defined by contour and the area of the disc. The mean absolute difference was 0.089, and the 95th percentile of the differences is 0.218

when outlining the optic cup and disc using color or vascular cues and deriving CDR from these measurements.

When outlining the optic cup, both color/pallor and contour/vascular cues yielded high intraclass correlations between graders. We used area, height, and width to make these measurements, finding that height yielded the highest intraclass correlations. Vertical CDR, which was calculated as the ratio of the optic cup (both by color and vascular cues) and optic disc, maintained strong intraclass correlations among graders. More than 94% of images (using color cups) and 92% of images (using contour cups) had a CDR difference of ≤ 0.2 between graders. This CDR difference (≤ 0.2) was previously defined as “concordant” by the Ocular Hypertension Treatment Study [27]. The lower intraclass correlation for CDR in controls when compared to cases could be due to discrepancies in smaller-sized cup ratios.

Our results demonstrate the importance of training graders in optic disc evaluation; we show that with training, very high concordance rates can be obtained between graders. Studies that included graders (even glaucoma specialists) that were not trained on standard images show lower rates of concordance. For example, Jampel et al. [15] asked glaucoma specialists to evaluate disc changes over time in a cohort of patients with visual field loss. Despite analyzing the qualities of cup enlargement, focal rim thinning, cup depth, and optic disc hemorrhages concurrently, inter-observer agreement was poor ($k = 0.2$). Likewise, Azuara-Blanco et al. [28] reported an inter-observer k of 0.34 to 0.68 (and intraobserver k of 0.55 to 0.78) among specialists assessing whether optic discs are compatible with glaucomatous changes. It is likely that even a small amount of standardized training among graders may have increased concordance in these studies. Inter-observer agreement among non-experts in detecting disc changes has been shown to increase after only one training session [29].

Our results also highlight how detailed and specific measurements can lead to high concordance rates among graders. In the Ocular Hypertension Treatment Study,

Table 3 Differences between CDR values assigned by two trained graders using color and contour cup for area, height, and width measurements, for cases and controls

	Difference	Color			Contour		
		Area	Vertical	Horizontal	Area	Vertical	Horizontal
Glaucoma cases	≤ 0.1	1407 (71%)	1477 (74%)	1253 (63%)	1341 (68%)	1405 (71%)	1141 (58%)
	$>0.1 - \leq 0.2$	461 (23%)	398 (20%)	526 (27%)	499 (25%)	440 (22%)	545 (27%)
	$>0.2 - \leq 0.3$	85 (4%)	79 (4%)	140 (7%)	115 (6%)	106 (5%)	228 (11%)
	>0.3	31 (2%)	30 (2%)	55 (3%)	29 (1%)	33 (2%)	70 (4%)
Controls	≤ 0.1	419 (74%)	435 (76%)	379 (66%)	314 (55%)	335 (59%)	252 (44%)
	$>0.1 - \leq 0.2$	126 (22%)	112 (20%)	144 (25%)	187 (33%)	170 (30%)	202 (35%)
	$>0.2 - \leq 0.3$	19 (3%)	19 (3%)	38 (7%)	56 (10%)	48 (8%)	90 (16%)
	>0.3	6 (1%)	4 (1%)	9 (2%)	13 (2%)	17 (3%)	26 (5%)

trained graders judged optic discs as deteriorated or not based on thinning of the neuroretinal rim, yielding agreement (κ) ranging from 0.65 to 0.83 over 5 years [30]. Likewise, the European Optic Disc Assessment Trial, which prompted 243 ophthalmologists to grade stereoscopic optic disc photos of healthy and glaucomatous eyes, found an overall diagnostic accuracy of 80.5% and intraobserver agreement average of 0.7 [24]. These results are considered an overestimate, as the study was conducted in ideal conditions, with graders having unlimited time to grade and only eyes with a definitive diagnosis used. These studies demonstrate the difficulty in classifying optic nerve damage solely based on subjective disc features, even with established protocols. We believe that using detailed measures to calculate CDR or grade the optic nerve, as in our study, can increase the accuracy of measurements. Meanwhile, this methodology also avoids the pitfalls of automated instruments, such as limitations when faced with anatomic variations or suboptimal images.

The time spent training the readers was extensive; we do not know whether the same degree of concordance would have been obtained with less training. It is also possible that the graders were more effective learners, given their prior familiarity with examining digital images. Additionally, there remains no gold-standard for analysis of the ONH and thus no “right answer” to compare against for measurements in this study. Nonetheless, the moderate agreement of the vertical CDR assessed clinically, with an intraclass correlation of 0.72 for assessment by color cues or by contour, indicates that the grading assessments are generally consistent with clinical judgment.

In conclusion, this study introduced a new method of quantitatively grading the ONH on stereo disc images, which resulted in high inter-grader reliability. We envision this system being useful for detecting and measuring small differences in the ONH for large-volume glaucoma studies. Going forward, we plan to train the graders to assess other parameters of digital photographs, such as alpha and beta peripapillary atrophy. In addition, we hope to analyze disc asymmetry between eyes and to examine POAG progression by evaluating future stereo disc images of the same patients over time.

Study highlights

What was known before

- Optic nerve grading for glaucoma has historically been challenging, with multiple evaluation schemes emerging in recent years. These systems are often subjective and not fully reproducible among graders.
- Automated instruments, such as optical coherence tomography, have also been used to grade the optic

nerve. These instruments, though useful for mass grading, have been shown to miss both early and severe cases of glaucoma.

- Thus, there is a need for a reproducible method for grading and classifying optic nerve disc photos. Ideally, this methodology would be more comprehensive and precise than existing subjective methods, while avoiding the pitfalls of automated instruments.

What this study adds

- We propose a novel quantitative method for assessing optic nerve stereo disc photographs.
- This method utilizes color and vascular cues, as well as area, height, and width, to evaluate the optic nerve and determine cup-to-disc ratio, an essential tool for evaluating glaucoma progression.
- We show that non-physician graders, who assessed 2554 images from African American patients, achieve high concordance rates with this method after receiving training.
- We believe that our method can be useful for assessing glaucoma progression or categorizing phenotypes for glaucoma research studies.

Funding This work was supported by the National Eye Institute, Bethesda, Maryland (grant #1RO1EY023557-01) and the Department of Ophthalmology at the Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA. Funds also come from the Vision Research Core Grant (P30 EY001583), F.M. Kirby Foundation, Research to Prevent Blindness, The UPenn Hospital Board of Women Visitors, The Paul and Evanina Bell Mackall Foundation Trust, and the National Eye Institute, National Institutes of Health, Department of Health and Human Services, under eyeGENETM and contract Nos. HHSN260220700001C and HHSN263201200001C. The sponsor or funding organization had no role in the design or conduct of this research.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Gandhi M, Dubey S. Evaluation of the optic nerve head in glaucoma. *J Curr Glaucoma Pract.* 2013;7:106–14.
2. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology.* 2014;121:2081–90.

3. Spaeth GL, Lopes JF, Junk AK, Grigorian AP, Henderer J. Systems for staging the amount of optic nerve damage in glaucoma: a critical review and new material. *Surv Ophthalmol.* 2006;51:293–315.
4. Hitchings RA, Spaeth GL. The optic disc in glaucoma. I: Classification. *Br J Ophthalmol.* 1976;60:778–85.
5. Lichter PR. Variability of expert observers in evaluating the optic disc. *Trans Am Ophthalmol Soc.* 1976;74:532–72.
6. Varma R, Steinmann WC, Scott IU. Expert agreement in evaluating the optic disc for glaucoma. *Ophthalmology.* 1992;99:215–21.
7. Abrams LS, Scott IU, Spaeth GL, Quigley HA, Varma R. Agreement among optometrists, ophthalmologists, and residents in evaluating the optic disc for glaucoma. *Ophthalmology.* 1994;101:1662–7.
8. Shiose Y. Quantitative analysis of “optic cup” and its clinical application. III. A new diagnostic criterion for glaucoma using “quantitative disc pattern” (Shiose) (author’s transl). *Nippon Ganka Gakkai Zasshi.* 1975;79:445–61.
9. Read RM, Spaeth GL. The practical clinical appraisal of the optic disc in glaucoma: the natural history of cup progression and some specific disc-field correlations. *Trans Am Acad Ophthalmol Otolaryngol.* 1974;78:OP255–74.
10. Schwartz B. Cupping and pallor of the optic disc. *Arch Ophthalmol.* 1973;89:272–7.
11. Richardson KT. Glaucoma and glaucoma suspects. *Glaucoma: Conceptions of a disease, pathogenesis, diagnosis, therapy.* 1978; 2–6.
12. Nesterov AP, Listopadova NA. Classification of physiological and glaucomatous extraction of the optic disk. *Vestn Oftalmol.* 1981;2:17–22.
13. Jonas JB, Gusek GC, Naumann GO. Optic disc morphometry in chronic primary open-angle glaucoma. II. Correlation of the intrapapillary morphometric data to visual field indices. *Graefes Arch Clin Exp Ophthalmol.* 1988;226:531–8.
14. Spaeth GL, Henderer J, Liu C, Kesen M, Altangerel U, Bayer A, et al. The disc damage likelihood scale: reproducibility of a new method of estimating the amount of optic nerve damage caused by glaucoma. *Trans Am Ophthalmol Soc.* 2002;100:181–5.
15. Jampel HD, Friedman D, Quigley H, Vitale S, Miller R, Knezevich F, et al. Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients. *Am J Ophthalmol.* 2009;147:39–44.e1.
16. Montgomery DM, Craig JP. Optic disc interpretation in glaucoma: is confidence misplaced? *Ophthalmic Physiol Opt.* 1993;13:383–6.
17. Harper R, Reeves B, Smith G. Observer variability in optic disc assessment: implications for glaucoma shared care. *Ophthalmic Physiol Opt.* 2000;20:265–73.
18. Gaasterland DE, Blackwell B, Dally LG, Caprioli J, Katz LJ, Ederer F, et al. The Advanced Glaucoma Intervention Study (AGIS): 10. Variability among academic glaucoma subspecialists in assessing optic disc notching. *Trans Am Ophthalmol Soc.* 2001; 99:177–84.
19. Vermeer KA, Vos FM, Lemij HG, Vossepoel AM. Detecting glaucomatous wedge shaped defects in polarimetric images. *Med Image Anal.* 2003;7:503–11.
20. Correnti AJ, Wollstein G, Price LL, Schuman JS. Comparison of optic nerve head assessment with a digital stereoscopic camera (discam), scanning laser ophthalmoscopy, and stereophotography. *Ophthalmology.* 2003;110:1499–505.
21. Schuman JS, Wollstein G, Farra T, Hertzmark E, Aydin A, Fujimoto JG, et al. Comparison of optic nerve head measurements obtained by optical coherence tomography and confocal scanning laser ophthalmoscopy. *Am J Ophthalmol.* 2003;135:504–12.
22. Zangwill LM, Weinreb RN, Beiser JA, Berry CC, Cioffi GA, Coleman AL, et al. Baseline topographic optic disc measurements are associated with the development of primary open-angle glaucoma: the Confocal Scanning Laser Ophthalmoscopy Ancillary Study to the Ocular Hypertension Treatment Study. *Arch Ophthalmol.* 2005;123:1188–97.
23. Banister K, Boachie C, Bourne R, Cook J, Burr JM, Ramsay C, et al. Can automated imaging for optic disc and retinal nerve fiber layer analysis aid glaucoma detection? *Ophthalmology.* 2016; 123:930–8.
24. Reus NJ, Lemij HG, Garway-Heath DF, Airaksinen PJ, Anton A, Bron AM, et al. Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European Optic Disc Assessment Trial. *Ophthalmology.* 2010;117:717–23.
25. Charlson ES, Sankar PS, Miller-Ellis E, Regina M, Fertig R, Salinas J, et al. The primary open-angle african american glaucoma genetics study: baseline demographics. *Ophthalmology.* 2015;122:711–20.
26. CATT Research Group, Martin DF, Maguire MG, Ying GS, Grunwald JE, Fine SL, et al. Ranibizumab and bevacizumab for neovascular age-related macular degeneration. *N Engl J Med.* 2011;364:1897–908.
27. Feuer WJ, Parrish RK 2nd, Schiffman JC, Anderson DR, Budenz DL, Wells MC, et al. The Ocular Hypertension Treatment Study: reproducibility of cup/disk ratio measurements over time at an optic disc reading center. *Am J Ophthalmol.* 2002;133: 19–28.
28. Azuara-Blanco A, Katz LJ, Spaeth GL, Vernon SA, Spencer F, Lanzl IM. Clinical agreement among glaucoma experts in the detection of glaucomatous changes of the optic disk using simultaneous stereoscopic photographs. *Am J Ophthalmol.* 2003; 136:949–50.
29. Breusegem C, Fieuws S, Stalmans I, Zeyen T. Agreement and accuracy of non-expert ophthalmologists in assessing glaucomatous changes in serial stereo optic disc photographs. *Ophthalmology.* 2011;118:742–6.
30. Parrish RK 2nd, Schiffman JC, Feuer WJ, Anderson DR, Budenz DL, Wells-Albornoz MC, et al. Test-retest reproducibility of optic disk deterioration detected from stereophotographs by masked graders. *Am J Ophthalmol.* 2005;140:762–4.