CLINICAL AND EXPERIMENTAL
OPTOMETRY

RESEARCH

Repeatability of Goldmann tonometry performed by optometry students on glaucoma patients

Clin Exp Optom 2020 DOI:10.1111/cxo.13053

Richard C Trevino* OD
Carolyn E Majcher*[§] OD
William E Sponsel*[†‡] MD
Carissa Villaflor*
Jobeth Nozicka*

*Rosenberg School of Optometry, University of the Incarnate Word, San Antonio, Texas, USA
[†]WESMDPA Glaucoma Service, Baptist Medical Center, San Antonio, Texas, USA
[‡]Department of Biomedical Engineering, The University of Texas at San Antonio, San Antonio, Texas, USA
[§]Present address: Oklahoma College of Optometry, Northeastern State University, Talequah, Oklahoma, USA
E-mail: rctrevin@uiwtx.edu

Submitted: 8 August 2019
Revised: 2 January 2020
Accepted for publication: 23 January 2020

Background: The purpose of this study was to evaluate the repeatability of masked Goldmann tonometry performed by optometry students on patients with glaucoma.
Methods: Subjects were recruited from among patients scheduled to undergo selective laser trabeculoplasty at the Rosenberg School of Optometry clinic. Each subject had masked Goldmann tonometry performed by three examiners at each office visit: two fourth professional-year optometry interns and an attending optometrist. Each examiner performed three sequential masked tonometry measurements on each eye.
Results: Twenty-eight interns and two optometrists performed masked Goldmann tonometry on 12 glaucoma patients. The co-efficient of variation was 9.1 per cent for the right eye and 12.1 per cent for the left eye for interns compared with 6.4 per cent right eye and 6.6 per cent left eye for optometrists. There was significant interaction between intern and patient on co-efficient of variation (two-factor analysis of variance, p = 0.005), indicating co-efficient of variation was influenced by both intern and patient factors. No such interaction was found for optometrist-performed measurements (p = 0.96). Mean interobserver difference for interns ranged between 0.9 and 3.1 mmHg, with 95 per cent limits of agreement that were proportional to mean intraocular pressure. Mean interobserver difference for optometrists ranged between 0.6 and 1.8 mmHg without proportionality bias. At higher pressure levels intern measurements became more variable and tended to overestimate optometrist measurements.
Conclusions: Both intraobserver and interobserver repeatability of masked tonometry was lower for interns than experienced optometrists. Intern performance differed from optometrists in that intern measurements became more variable at higher intraocular pressure levels and were significantly influenced by patient factors. The present results support the need for trainee exposure to patients with abnormally elevated intraocular pressure. Research into factors that influence trainee Goldmann tonometry repeatability is needed.

Key words: glaucoma, optometric education, repeatability, tonometry

Goldmann applanation tonometry is regarded as the gold standard for clinical assessment of intraocular pressure (IOP).[1] As such, learning how to properly perform Goldmann tonometry is part of the essential skill set for trainees in both optometry and ophthalmology.[2,3] Determining when the endpoint of Goldmann tonometry has been achieved is subjective and some judgement is required. As such, it is expected that a learning curve is associated with mastering the procedure, wherein tonometry measurements performed by novice clinicians are less precise than those performed by experienced clinicians.

Many researchers have examined the repeatability of Goldmann tonometry.[4–12] Due to the subjective nature of the measurement, most studies employ a masking procedure that eliminates examiner bias. In addition, because skill may improve with experience, most studies have utilised experienced ophthalmologists. Under such conditions, the intraobserver repeatability of Goldmann tonometry is reported to be approximately 2–3 mmHg.[4,5,13] This represents the expected variation between two sequential tonometry measurements made by the same examiner on the same eye of a given patient.

The mean interobserver reproducibility of masked Goldmann tonometry performed by experienced clinicians is reported to be < 1 mmHg with a 95% confidence interval (CI) of ±3–4 mmHg.[5,12,13] This represents the expected variation between two sequential measurements made by different examiners on the same eye of a given patient. The World Glaucoma Association has reported that under ideal conditions the intraobserver repeatability of Goldmann tonometry is 2.5 mmHg and the 95% CI for interobserver repeatability is ±4 mmHg.[6]

Little has been published regarding the repeatability of Goldmann tonometry performed by trainees. The purpose of this study was to examine the intraobserver and interobserver repeatability of Goldmann tonometry performed by fourth-year optometry interns on patients with glaucoma.

Methods

This study was approved by the Institutional Review Board of the University of the Incarnate Word, and adhered to the tenets of the Declaration of Helsinki. All data were managed in accordance with the regulations set forth in the Health Insurance Portability and Accountability Act. All subjects gave written informed consent after being informed of the nature of the study.

Procedures

Subjects were recruited from among patients scheduled to undergo selective laser trabeculoplasty at the Rosenberg School of Optometry (RSO) clinic. All patients had mild to moderate glaucoma and had discontinued all topical glaucoma medications for at least four weeks prior to being scheduled for trabeculoplasty. Both eyes of each patient were included in the study. Seven slitlamp-mounted Goldmann applanation tonometers (Haag-Streit, Bern, Switzerland) were utilised in this study. Each tonometer was calibrated and maintained in accordance with the recommendations of the manufacturer.

Each patient participating in this study had Goldmann tonometry performed by two fourth-year optometry interns and one optometrist at each office visit. The interns employed in this study were those interns that were assigned to the Glaucoma Service at the time of the appointment of the patient. At the time of their participation, each intern would have experienced a total of approximately 200 patient encounters, most involving use of a Goldmann tonometer. No effort was made to randomise or evenly distribute the number of measurements performed by each intern. Instead, interns were selected based on convenience. Because all fourth-year interns rotate through the RSO Glaucoma Service, our sample represents a cross-section of the entire class at this point in their training.

One of two optometrists (RCT, CEM), each with over 10 years of experience performing Goldmann tonometry, performed the third set of readings. There were four instances where a second intern was not readily available to perform tonometry and instead both optometrists performed measurements on the same patient. These measurements were separately analysed.

Following instillation of a commercially prepared topical solution containing benoxinate hydrochloride 0.4 per cent and fluorescein sodium 0.25 per cent, each clinician performed three sequential tonometry measurements on each eye; right eye first, then left eye (RLRLRL pattern). Prior to each measurement, an investigator set the tonometer dial to a random position between five and 40 mmHg (determined by lottery) and placed a mask over the dial. After each measurement the investigator recorded the measurement and set the dial to a new random position. The clinician was blinded to both the initial dial position and to his measurements. Less than one minute elapsed between sequential tonometry measurements. The procedure was then repeated with the second and third clinicians. Several minutes (≈five minutes) elapsed between clinicians. The order of the clinicians was first intern (INT1), second intern (INT2), and attending optometrist (AOD).

Statistics

Normality was assessed with the Shapiro–Wilke test. Mean and standard deviation (SD) for the three sequential measurements was calculated across patients for right and left eyes. Unpaired two-tail t-tests were used to test for differences between right and left eyes. Repeat measures two-factor analysis of variance (ANOVA) with the Bonferroni correction for multiple comparisons was performed to identify significant differences between the three sequential measurements and whether the measurements performed by interns or optometrists were significantly different. Mauchly's test of sphericity was performed and the Greenhouse–Geisser correction was applied as needed.

Measurement differences proportional to their mean is known as heteroscedasticity, and will distort measures of repeatability. Heteroscedasticity was assessed by identifying correlation between the difference of the second (M2) and third (M3) measurements and their mean using Kendall's tau.[14] Log_{10} transformation of the data was performed for significant correlations and then reassessed to confirm that correlations were no longer significant prior to performing repeatability analysis.

Both intraoperator and interoperator repeatability of tonometry measurements were assessed. For the intraoperator analysis the repeatability of the three sequential measurements performed by each examiner class (interns and optometrists) was examined. The co-efficient of variation (COV) was calculated as the ratio between the SD of the measurements and their mean. Two-factor ANOVA was performed on the COV values using intern as one factor and patient as the second factor. The intraclass correlation co-efficient (ICC) was calculated using a two-way mixed effects model for absolute agreement of single measures ($ICC_{2,1}$). ICC values ≥ 0.75 represent excellent reliability; values between 0.40 and 0.74 represent moderate reliability; and values < 0.40 indicate poor reliability.[15] The repeatability co-efficient was calculated as 2.77 times the within-subject SD (ws[SD]) for M2 and M3, with ws(SD) calculated as follows:

$$ws(SD) = \frac{1}{\sqrt{2n}} \times \sqrt{\sum (M_2 - M_3)^2} \quad [1]$$

where M2 and M3 are the second and third tonometry readings of each examiner and n is the total number of measurement pairs analysed. M2 and M3 were selected for analysis because patient anxiety and tonography are expected to cause a drop in measured IOP between successive measurements, and this change tends to be greatest between M1 and M2.[16] This approach is consistent with prior research.[17,18]

Interoperator reproducibility was assessed by analysing paired IOP measurements between interns and between each intern and the optometrist. In addition, reproducibility between optometrists in those instances, when both optometrists participated on the same patient, was analysed. Systematic bias was investigated using one-way ANOVA employing the Bonferroni correction for multiple comparisons on the measurements performed by each examiner. The interoperator $ICC_{2,1}$ was calculated by comparing each of the three sequential tonometry readings performed by one examiner with the corresponding measurement performed by another examiner. For example, the first reading by the first examiner is compared to the first reading of the second examiner, and so forth.

The limits of agreement (LOA) is the range for 95 per cent of paired value differences, and was calculated as 1.96 times the SD of the difference values. The LOA for heteroscedastic data were transformed back to their original scale and displayed in the Bland–Altman plots as a linear function calculated using the method described by Euser et al.[19] as follows:

$$LOA = Z \pm \frac{2\bar{x}(10^a - 1)}{(10^a + 1)} \quad [2]$$

where *Z* is the mean bias between examiners, *a* is 1.96 times the SD of the log-transformed

	n[†]	M1 (SD) mmHg	M2 (SD) mmHg	M3 (SD) mmHg	p-value[‡]
Right					
Intern	47	20.04 (4.47)	20.23 (5.31)	18.74 (4.45)	0.03
Optometrist	30	17.67 (4.36)	17.60 (3.57)	17.93 (3.81)	0.77
Left					
Intern	45	22.82 (8.93)	20.49 (5.44)	21.38 (7.64)	0.06
Optometrist	29	19.66 (4.85)	19.59 (4.87)	19.62 (6.01)	0.99

M1: first measurement, M2: second measurement, M3: third measurement.

[†]Number of examinations.

[‡]Repeat measures analysis of variance.

Table 1. Mean intraocular pressure measurements

values, and $\bar{x}$ is their mean on the original scale. COV for log-transformed values is 10^a-1.[14]

Statistical analyses were performed using SPSS Statistics version 24 (IBM Corp., Armonk, NY, USA), Excel 2016 (Microsoft, Redmond, WA, USA) and Prism 6 (Graphpad Software Inc, La Jolla, CA, USA).

Results

Over a six-month period, 28 interns (41 per cent of the entire class) performed one or more examinations on 12 glaucoma patients. Both eyes of each patient were enrolled in the study except in one instance of a monocular patient who did not have a left eye. Over the course of the study there were five protocol breaches wherein one of the following occurred: (1) no attending optometrist performed tonometry on the patient (one occurrence); or (2) both attendings performed tonometry on the patient (four occurrences). Each of these breaches are excluded from the main interobserver analysis. All tonometry measurements during breach incidents are included in the intraobserver analysis. The four occasions when both attendings performed tonometry on a single patient are analysed separately to assess interobserver repeatability between optometrists.

Intraobserver repeatability

A total of 77 examinations were performed on right eyes (OD) and 74 examinations were performed on left eyes (OS) (Table 1). Each examination consisted of three sequential masked measurements (M1, M2, M3) (Figure 1). Repeat measures two-factor ANOVA found no significant difference among the three sequential tonometry readings (OD: p = 0.16; OS: p = 0.47) or between the intern and optometrist examiner classes (OD: p = 0.11; OS: p = 0.54).

	COV[†]	Repeatability[‡]	ICC (95% CI)[§]
Right			
Interns	9.1%	9.65 mmHg	0.60 (0.45–0.74)
Optometrists	6.4%	4.32 mmHg	0.77 (0.63–0.88)
Left			
Interns	12.1%	10.95 mmHg	0.61 (0.45–0.75)
Optometrists	6.6%	5.76 mmHg	0.88 (0.78–0.94)

COV: co-efficient of variation, ICC: intraclass correlation co-efficient, 95% CI: 95% confidence interval.

[†]Ratio of mean within-examiner standard deviation across three sequential tonometry readings to the mean tonometry reading.

[‡]2.77 times the within-examiner standard deviation for second and third tonometry readings.

[§]Two-way mixed effects model for absolute agreement of single measures using three sequential readings.

Table 2. Intraobserver repeatability of Goldmann tonometry

Heteroscedasticity was not significant for either eye for each examiner class. Therefore, log transformation was not carried out for assessment of within-examiner repeatability.

The COV values indicate that the dispersion of tonometry measurements performed by interns is 50 to 100 per cent greater than readings performed by optometrists (Table 2). Two-factor ANOVA with one factor being examiner (with 28 levels) and the second factor being patient (with 12 levels) found significant interaction between examiner and patient on COV (p = 0.005) (Figure 2). This suggests that the intraobserver repeatability of any given intern is influenced by both the intern's innate variability and variability introduced by the patient, with some patients generating more variable readings than others. No such interaction between examiner and patient was found for measurements performed by optometrists (p = 0.96). Hence the influence of patient-related factors on intraobserver repeatability wanes as the examiner becomes more experienced.

The within-examiner difference between M2 and M3 plotted against their mean is presented in Figure 3. The LOA between M2 and M3 were approximately twice as broad for intern measurements compared with those performed by optometrists. Interestingly, agreement between M2 and M3 was significantly related to the level of IOP for both examiner classes for left eyes only (intern: p = 0.003, R^2 = 18.3 per cent; attending: p = 0.04, R^2 = 15.4 per cent).

Interobserver reproducibility

A total of 35 sets of tonometry measurements (18 OD, 17 OS) were available for analysis, wherein two interns and one optometrist each performed three

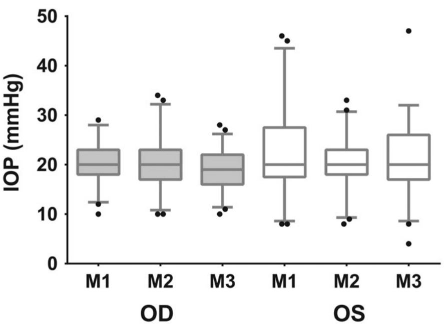

Figure 1. Box plots of Goldmann tonometry measurements performed by interns. The box represents the median and interquartile range for each of the three sequential tonometry measurements performed on each eye. The whiskers represent the fifth and 95th percentiles with outliers individually plotted. M1: first measurement, M2: second measurement, M3: third measurement, OD: right eye, OS: left eye, IOP: intraocular pressure.

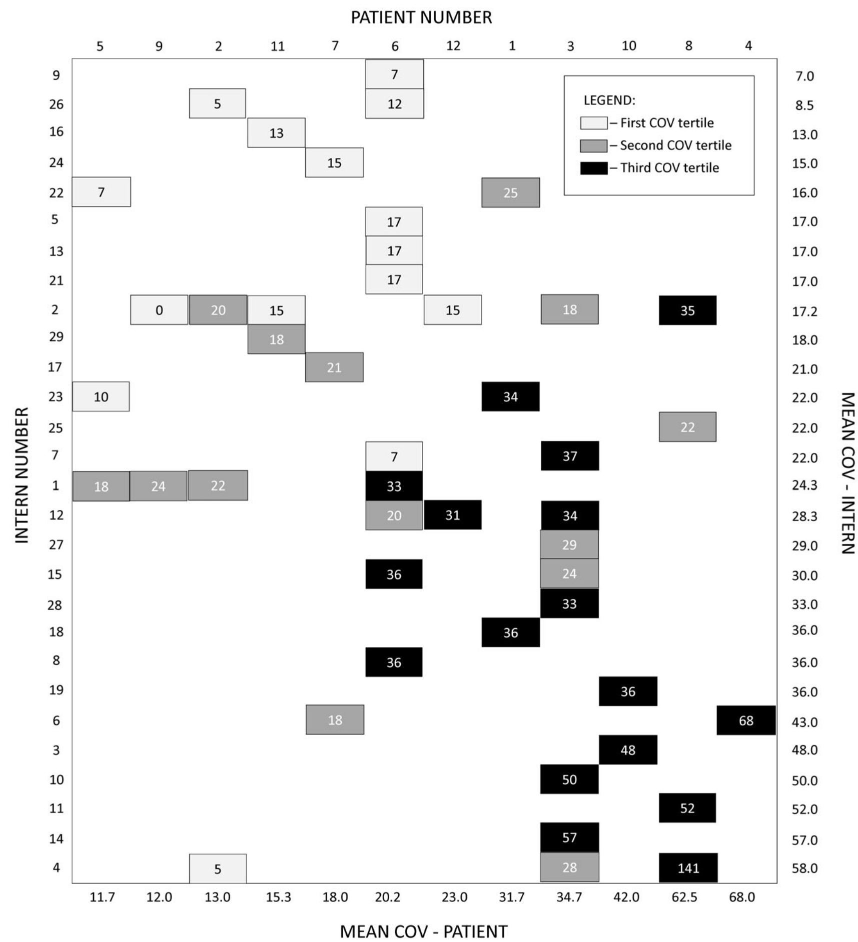

Figure 2. Scatter plot of Goldmann tonometry co-efficient of variation (COV) by intern and patient. Interns and patients are sorted by mean COV. Each datapoint represents the COV for three sequential tonometry measurements performed by an intern on a patient. The data set is divided into tertiles. Note that variance is not evenly distributed across study participants (interns and patients). Rather, there are those participants that tend to generate high variance (third COV tertile) while others are associated with low variance (first COV tertile). There is significant interaction (p = 0.005) between intern and patient indicating that both intern- and patient-specific factors govern tonometry repeatability. For example, patient 6 generates less variance when tonometry is performed by a low variance intern, but more variance when examined by a high variance intern. Similarly, intern 2 generates less variance when examining low variance patients than when examining high variance patients.

sequential tonometry measurements. There was a downward trend in mean IOP across examiners which achieved statistical significance for right eyes only (OD: p = 0.004, OS: p = 0.12; Table S1; Figure S1).

All interobserver reproducibility analyses were carried out on log-transformed values because of significant heteroscedasticity between the first and second interns (INT1-INT2) and between the first intern and the attending optometrist (INT1-AOD) for both right and left eyes. The correlation between the second intern and the optometrist (INT2-AOD) approached significance for the right eye (p = 0.06) and was not significant for the left eye (p = 0.23).

The difference between paired tonometry readings for INT1 and INT2 are plotted against their mean in Figure 4A and D. The slope of the regression line of these data points was not significantly different from zero, indicating that mean bias between interns was constant across the range of IOPs encountered in this study. The slope of the LOA indicates that the range within which 95 per cent of paired readings are expected to differ is equivalent to approximately 80 per cent of the mean IOP right eye and 100 per cent of the mean IOP left eye (Table 3).

Intern-optometrist bias was substantially greater for INT1 than INT2 for both left and right eyes (Table 3, Figure S2). This is attributable to a tonography effect, which generated a downward trend in mean IOP across examiners (Figure S1). No trend favouring INT1 or INT2 was found in other measures of repeatability, such as ICC and COV. The difference in paired tonometry readings between each intern and the optometrist are plotted against their mean in Figure 4B–F. The slope of the regression line of these data points was not significantly different from zero for the right eye but was for the left eye, indicating that for left eyes only, mean bias between the intern and the optometrist was not constant across the range of IOPs encountered in this study. For both interns, the measurements of the interns tended to be less than that of the optometrists' measurements at low IOPs and greater than that of the optometrists at high IOPs for left eyes only.

Between-observer repeatability of the optometrists was analysed using the four instances where a protocol breach resulted in the two optometrists performing tonometry on the same patient. These data provide a context for intern

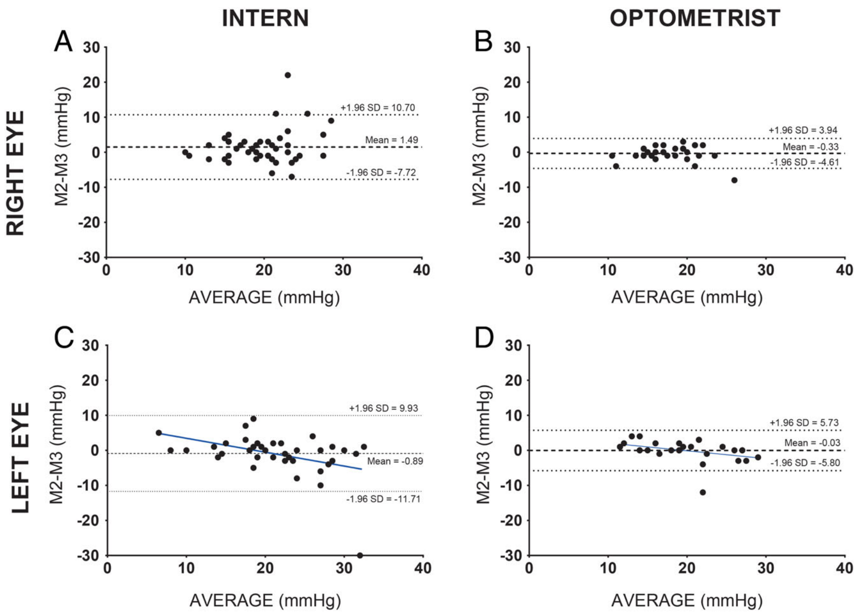

Figure 3. Bland–Altman plots of intraobserver Goldmann tonometry measurements. A: Right eye measurements performed by interns. B: Right eye measurements performed by optometrists. C: Left eye measurements performed by interns. D: Left eye measurements performed by optometrists. The solid lines in C and D are where linear regression of the data points yield a line in which the slope is significantly different from zero. M2: second measurement, M3: third measurement.

performance. The data were not significantly heteroscedastic, therefore log transformation was not performed. Repeatability statistics are reported in Table S2 and the Bland–Altman plots are presented in Figure S3. The slope of the regression line of these data points was not significantly different from zero.

Discussion

This study documents that masked tonometry performed by fourth professional-year optometry interns on glaucoma patients was approximately twice as variable as measurements performed by experienced optometrists. Because tonometer masking may have a detrimental effect on repeatability,[5,20] it would be appropriate to interpret the repeatability values reported herein as 'worst case' values.

Intraobserver repeatability

Prior studies reporting the intraobserver repeatability co-efficient for masked Goldmann tonometry performed by experienced clinicians have reported values ranging between 2.2 and 2.9 mmHg.[4,5,13] This is

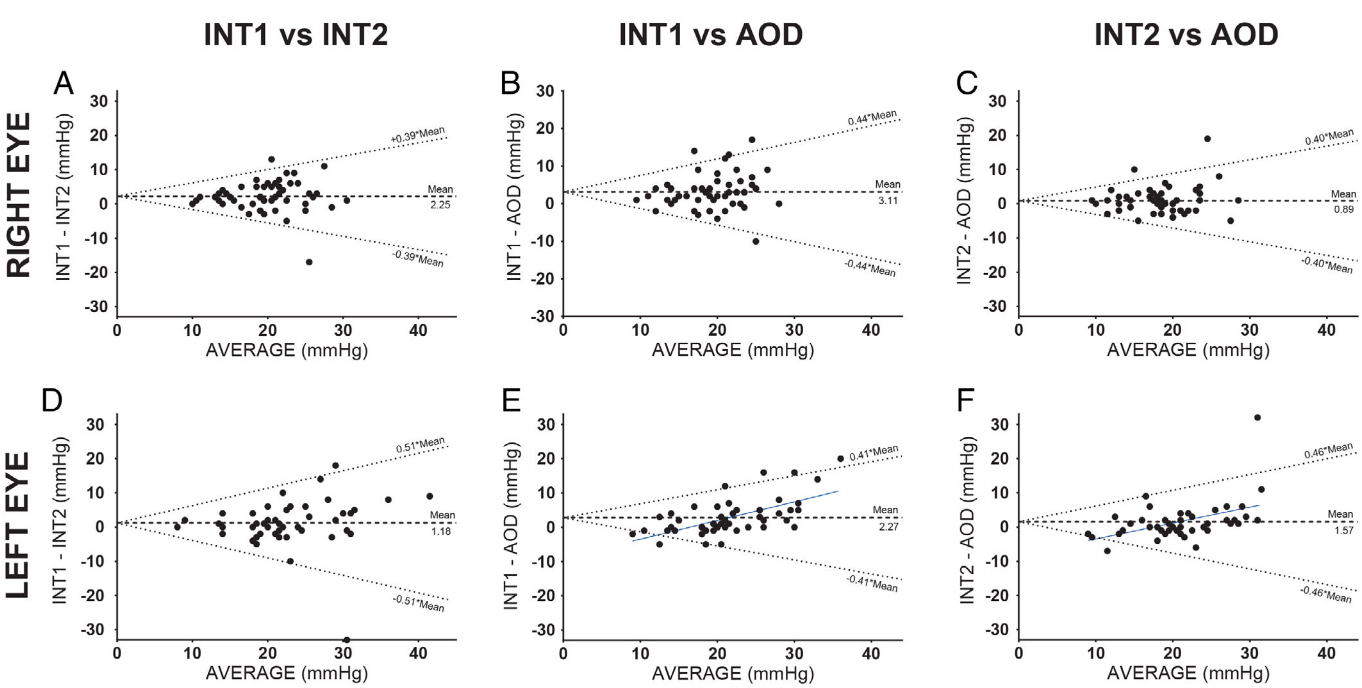

Figure 4. Bland–Altman plots of interobserver Goldmann tonometry measurements. A: Right eye measurements performed by intern 1 and intern 2. B: Right eye measurements performed by intern 1 and optometrist. C: Right eye measurements performed by intern 2 and optometrist. D: Left eye measurements performed by intern 1 and intern 2. E: Left eye measurements performed by intern 1 and optometrist. F: Left eye measurements performed by intern 2 and optometrist. The solid lines in E and F are where linear regression of the data points yield a line in which the slope is significantly different from zero. INT1: first intern, INT2: second intern, AOD: attending optometrist.

substantially more repeatable than interns or optometrists in this study. While lack of experience may account for the higher variability of interns, the research protocol may have affected the performance of optometrists. They were the last of three examiners to perform tonometry on each patient, and the quality of the tear film, integrity of the corneal epithelium and patient co-operation may all have been suboptimal at that point.

Prior studies reporting the intraobserver COV for masked Goldmann tonometry performed by experienced clinicians have reported values ranging between 3.1 and 9.7 per cent.[10–12] These values compare favourably to the results for both interns (9.1–12.1 per cent) and optometrists (6.4–6.6 per cent) in the current study. Finally, studies reporting intraobserver ICC values for masked Goldmann tonometry performed by experienced clinicians have reported values ranging between 0.79 and 0.99.[11,12] Intern performance fell well below this range (0.60–0.61) while it was consistent with that of the optometrists (0.77–0.88).

While numerous studies have investigated the repeatability of Goldmann tonometry, this is the first to involve optometry interns. One published study has evaluated the repeatability of masked Goldmann tonometry performed by a single medical student.[8] The repeatability co-efficient for their student (2.2 mmHg) is on the low end of published values and is superior to the performance of the interns in the present study (9.7–11.0 mmHg). This difference is possibly related to differences in randomisation of the dial starting position. Tonnu et al.[8] set the dial to a random starting position between five and 10 mmHg, while the present study set the dial to a random starting position between five and 40 mmHg.

There was a surprising difference in intraobserver repeatability on right and left eyes. There was a trend of lower repeatability for left eye compared to right eye measurements for both examiner classes. For example, intern COV was 9.1 per cent OD and 12.1 per cent OS (Table 2). A similar but smaller interocular difference was recorded by optometrists. Because a RLRLRL measurement pattern was followed, it seems unlikely that the sequence of measurements introduced any systematic bias. This difference may be attributable to hand and/or eye dominance of the examiner. For example, there may be greater precision when the adjustment is made with the dominant rather than the nondominant hand. The authors are not aware of any prior published reports describing interocular differences of Goldmann tonometry repeatability.[21]

The current study found that Goldmann tonometry performed by interns is influenced both by intrinsic and extrinsic factors, indicating that intern performance may vary from patient to patient. The authors believe this finding has important clinical and educational implications, as discussed below.

Interobserver reproducibility

The tonography effect resulted in a downward trend in mean IOP across examiners, and hence greater mean intern-optometrist bias for INT1 compared with INT2. Other measures of reproducibility (95% LOA, ICC COV) did not favour one intern over the other (Table 3). Hence, intern reproducibility can be analysed independent of intern sequence.

Prior studies have reported the mean interobserver difference of masked Goldmann tonometry performed by experienced clinicians to range between 0.5 to 0.8 mmHg with a 95% CI of ±3.0–4.0 mmHg.[5,12,13] The current study found that interobserver differences involving interns were significantly related to IOP level, whereas differences between the optometrists were not significantly heteroscedastic. Significant heteroscedasticity of interobserver masked Goldmann tonometry has not been previously reported and may represent a feature of novice performance.

There was a pattern of interns underestimating the measurements of optometrists at low pressures and over-estimating them at high pressures, but only for measurements performed on left eyes. An opposite pattern of bias was reported Tonnu et al.[8] wherein tonometry was first performed by a medical student and then by an experienced clinician. Only one eye of each patient was examined by Tonnu et al.[8] so it is not possible to determine if this effect was influenced by laterality in their study. The intraobserver findings reported herein revealed lower repeatability of measurements performed on left eyes, so the observed proportionality bias may reflect a pattern of lower precision of measurements performed on left eyes.

Educational considerations

The present study found that two sequential Goldmann tonometry measurements by the same intern on the same eye of a given patient will vary by up to 10 mmHg. When an intern and optometrist both perform tonometry on the same patient, the reading of the intern will fall within 4 mmHg of that of the

	n[†]	Mean bias (mmHg)	95% LOA (mmHg)	ICC (95% CI)[‡]	COV
Right					
INT1-INT2	54	2.25	± 0.39·Z	0.57 (0.30–0.74)	48%
INT1-AOD	54	3.11	± 0.44·Z	0.45 (0.12–0.67)	57%
INT2-AOD	54	0.89	± 0.40·Z	0.64 (0.45–0.77)	49%
Left					
INT1-INT2	51	1.18	± 0.51·Z	0.58 (0.37–0.74)	67%
INT1-AOD	51	2.75	± 0.41·Z	0.62 (0.36–0.78)	52%
INT2-AOD	51	1.57	± 0.46·Z	0.58 (0.37–0.74)	59%

AOD: attending optometrist, CI: confidence interval for ICC, COV: co-efficient of variation, ICC: intraclass correlation co-efficient, INT1: first intern, INT2: second intern, LOA: limits of agreement, Z: mean intraocular pressure.

[†]Number of paired comparisons.

[‡]Two-way mixed effects model for absolute agreement of single measures.

Table 3. Interobserver repeatability of Goldmann tonometry

optometrist when mean IOP is 10 mmHg and be within 8 mmHg when mean IOP is 20 mmHg. These findings have potentially important patient care ramifications.

Preceptor verification of tonometry measurements on patients with elevated IOP is warranted, while pressure measurements within the normal range are more trustworthy. When learning to perform the procedure, trainees need exposure to glaucoma patients with abnormal IOPs to facilitate the transfer of technical skills to a clinical setting.[22] While intern exposure to patients with abnormal IOP requires greater oversight, it provides interns with experience necessary to master this important skill.

A right/left laterality difference in measurement repeatability was found for both interns and experienced practitioners, suggesting a poorer prognosis for remediation of this phenomena through educational interventions.

Examiner- and patient-specific factors independently contributed to measurement repeatability for interns, but not for optometrists. This suggests that with experience clinicians learn how to cope with patient factors that adversely impact repeatability. The authors are unaware of any prior research that has attempted to identify such factors. The present study found that interns perform more poorly on patients with abnormally elevated IOP. Prior research has identified many factors that influence Goldmann tonometry precision,[16] including some that may pose a greater challenge to the novice, such as minimising patient anxiety. Educational interventions would benefit from the identification of those factors that pose the greatest challenge to trainees.

The present study has limitations that may influence interpretation of its findings. First, there was a large turnover of both interns and patients over the course of the study with some interns performing more measurements than others and some patients participating more than others. This makes interpretation of intern-patient interactions difficult. Second, many tonometry measurements were performed on each eye within a relatively short period of time. This introduced a tonography effect and may have adversely impacted the reliability of later measurements. Third, the large range for tonometer dial randomisation may have adversely impacted repeatability by prolonging each tonometry measurement.

The present study also has important strengths that should be taken into consideration. A relatively large number of measurements were collected over a wide range of pressures on patients with established glaucoma, enhancing the ability to evaluate intern performance on such patients. In addition, the study enabled the evaluation of the relative performance of interns and clinicians under identical experimental conditions thereby revealing the influence of experience.

In summary, intraobserver and interobserver repeatability of masked Goldmann tonometry performed by interns was approximately twice as variable as experienced optometrists. For both examiner classes, the repeatability of measurements performed on right eyes was greater than measurements performed on left eyes. Intern measurements became more variable at higher pressure levels. In addition, intraobserver repeatability was influenced both by intern and patient-related factors. Studies with larger numbers of interns and patients are needed to confirm our findings and to identify those factors that adversely influence repeatability. Training strategies could then be developed addressing those factors.

ACKNOWLEDGEMENTS

This research was presented, in part, at the annual meeting of the American Academy of Optometry, October 2018. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Okafor KC, Brandt JD. Measuring intraocular pressure. *Curr Opin Ophthalmol* 2015; 26: 103–109.
2. Cioffi GA, Durcan FJ, Girkin CA et al. Glaucoma. In: Skuta GL, Cantor LB, Weiss JS, eds. *2018–2019 Basic and Clinical Science Course*, Vol. 10. San Francisco, CA: American Academy of Ophthalmology, 2017.
3. National Board of Examiners in Optometry. *Clinical Skills Exam: Candidate Guide 2018–2019*. 2018.
4. Dielemans I, Vingerling JR, Hofman A et al. Reliability of intraocular pressure measurement with the Goldmann applanation tonometer in epidemiological studies. *Graefes Arch Clin Exp Ophthalmol* 1994; 232: 141–144.
5. Kotecha A, White E, Schlottmann PG et al. Intraocular pressure measurement precision with the Goldmann applanation, dynamic contour, and ocular response analyser tonometers. *Ophthalmology* 2010; 117: 730–737.
6. Weinreb RN, Brandt JD, Garway-Heath DF. Measurement of intraocular pressure. In: Weinreb RN, Brandt JD, Garway-Heath DF et al., eds. *Intraocular Pressure: Consensus Series*, Vol. 4. Ft. Lauderdale, FL: Kugler Publications, 2007. pp. 17–58.
7. Thorburn W. The accuracy of clinical applanation tonometry. *Acta Ophthalmol* 1978; 56: 1–5.
8. Tonnu P-AA, Ho T, Sharma K et al. A comparison of four methods of tonometry: method agreement and interobserver variability. *Br J Ophthalmol* 2005; 89: 847–850.
9. Schweier C, Hanson JVM, Funk J et al. Repeatability of intraocular pressure measurements with Icare PRO rebound, Tono-Pen AVIA, and Goldmann tonometers in sitting and reclining positions. *BMC Ophthalmol* 2013; 13: 44.
10. Salim S, Du H, Wan J. Comparison of intraocular pressure measurements and assessment of intraobserver and interobserver reproducibility with the portable icare rebound tonometer and goldmann applanation tonometer in glaucoma patients. *J Glaucoma* 2013; 22: 325–329.
11. Salvetat ML, Zeppieri M, Tosoni C et al. Repeatability and accuracy of applanation resonance tonometry in healthy subjects and patients with glaucoma. *Acta Ophthalmol* 2014; 92: 66–73.
12. Wang AS, Alencar LM, Weinreb RN et al. Repeatability and reproducibility of Goldmann applanation, dynamic contour, and ocular response analyzer tonometry. *J Glaucoma* 2013; 22: 127–132.
13. Cheng J, Salam T, Russell PJ et al. Dynamic contour tonometer and goldmann applanation tonometer performance in a developing world setting: intraocular pressure measurement acquisition and precision. *J Glaucoma* 2013; 22: 736–739.
14. Bland JM, Altman DG. Measurement error proportional to the mean. *BMJ* 1996; 313: 106.
15. FJL. *Design and Analysis of Clinical Experiments*. Hoboken, NJ: John Wiley & Sons, 1986.
16. Whitacre MM, Stein R. Sources of error with use of Goldmann-type tonometers. *Surv Ophthalmol* 1993; 38: 1–30.
17. Pandav SS, Sharma A, Gupta A et al. Reliability of Pro-Ton and Goldmann applanation tonometers in normal and postkeratoplasty eyes. *Ophthalmology* 2002; 109: 979–984.
18. Vernon SA. Reproducibility with the keeler pulsair 2000 non-contact tonometer. *Br J Ophthalmol* 1995; 79: 554–557.
19. Euser AM, Dekker FW, le Cessie S. A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *J Clin Epidemiol* 2008; 61: 978–982.
20. Kotecha A, White ET, Shewry JM et al. The relative effects of corneal thickness and age on Goldmann applanation tonometry and dynamic contour tonometry. *Br J Ophthalmol* 2005; 89: 1572–1575.
21. Buchan JC, Macleod D, Hickman W et al. Systematic bias in real-world tonometry readings based on laterality? *Eye (Lond)* 2019; 34: 360–365.
22. Wong YL. Utilizing the principles of Gagne's nine events of instruction in the teaching of Goldmann Applanation Tonometry. *Adv Med Educ Pract* 2018; 9: 45–51.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Mean (SD) intraocular pressure by examiner.

Table S2. Interobserver repeatability of Goldmann tonometry performed by optometrists.

Figure S1. Mean intraocular pressure recorded by each of three sequential examiners. Error bars represent standard deviation. INT1: first intern, INT2: second intern, AOD: attending optometrist, OD: right eye, OS: left eye, IOP: intraocular pressure.

Figure S2. Scatter plots of Goldmann tonometry measurements performed by interns and attending optometrists. Points that lie on the diagonal line represent identical readings by both clinicians. A: Agreement between first intern and optometrist for measurements on right eyes. B: Agreement between first intern and optometrist for measurements on left eyes. C: Agreement between second intern and optometrist for measurements on right eyes. D: Agreement between second intern and optometrist for measurements on left eyes. INT1: first intern, INT2: second intern, AOD: attending optometrist.

Figure S3. Bland–Altman plots of Goldmann tonometry measurements performed by attending optometrists. A: Right eye; B: left eye. AOD1: optometrist 1, AOD2: optometrist 2.