

Potential effects of systematic errors in intraocular pressure measurements on screening for ocular hypertension

MJ Turner¹, SL Graham¹, AP Avolio¹ and P Mitchell²

Abstract

Aim Raised intraocular pressure (IOP) increases the risk of glaucoma. Eye-care professionals measure IOP to screen for ocular hypertension (OHT) (IOP > 21 mm Hg) and to monitor glaucoma treatment.

Tonometers commonly develop significant systematic measurement errors within months of calibration, and may not be verified often enough. There is no published evidence indicating how accurate tonometers should be. We analysed IOP measurements from a population study to estimate the sensitivity of detection of OHT to systematic errors in IOP measurements.

Methods We analysed IOP data from 3654 participants in the Blue Mountains Eye Study, Australia. An inverse cumulative distribution indicating the proportion of individuals with highest IOP > 21 mm Hg was calculated. A second-order polynomial was fitted to the distribution and used to calculate over- and under-detection of OHT that would be caused by systematic measurement errors between -4 and $+4$ mm Hg. We calculated changes in the apparent prevalence of OHT caused by systematic errors in IOP.

Results A tonometer that consistently under- or over-reads by 1 mm Hg will miss 34% of individuals with OHT, or yield 58% more positive screening tests, respectively. Tonometers with systematic errors of -4 and $+4$ mm Hg would miss 76% of individuals with OHT and would over-detect OHT by a factor of seven. Over- and under-detection of OHT are not strongly affected by cutoff IOP.

Conclusion We conclude that tonometers should be maintained and verified at intervals short enough to control systematic errors in IOP measurements to substantially less than 1 mm Hg.

Eye (2013) 27, 502–506; doi:10.1038/eye.2013.7; published online 22 February 2013

Keywords: intraocular pressure; systematic measurement error; calibration; tonometer

Introduction

Intraocular pressure (IOP) greater than 21 mm Hg, commonly classified as ocular hypertension (OHT),¹ substantially increases the risk of glaucoma.² Glaucoma, a neurodegenerative disease of the retinal ganglion cells, is the second leading cause of blindness worldwide.³ In the United States, 4 million people have glaucoma and 130 000 are legally blind from the disease.³ In the Australian Blue Mountains Eye study of adults 49 years of age and older, the prevalence of open-angle glaucoma was 3% and OHT was found in 3.7% of participants.¹ Control of IOP is important in patients with glaucoma:⁴ a decrease of 1 mm Hg may lead to a 10% reduction in the risk of progressive nerve damage.^{5,6} Reduction and control of IOP are the only effective treatments for glaucoma.

IOP is measured by eye-care professionals to screen for, diagnose, treat and monitor OHT and eye diseases such as glaucoma. The Goldmann applanation tonometer in experienced hands is recognised as the gold standard for IOP measurements.⁷ Both Goldmann applanation and non-contact tonometers are known to produce clinically significant systematic errors if inadequately maintained and calibrated,^{8–15} and current calibration intervals may be too long.^{8,9,12,14,16–18}

Systematic errors in IOP measurements are likely to lead to over- or under-detection of OHT. Over-detection of OHT during screening

¹The Australian School of Advanced Medicine, Macquarie University, Sydney, New South Wales, Australia

²Discipline of Clinical Ophthalmology and Eye Health, C24 Westmead Hospital, The University of Sydney, Sydney, New South Wales, Australia

Correspondence: MJ Turner, The Australian School of Advanced Medicine, Macquarie University, Building F10A, 2 Technology Place, Sydney, New South Wales 2109, Australia
Tel: +61 2 9812 3606;
Fax: +61 02 9812 3600.
E-mail: martin.turner@mq.edu.au

Received: 18 June 2012
Accepted in revised form: 21 January 2013
Published online: 22 February 2013

may result in excessive referrals to ophthalmologists, while under-detection may place individuals at increased risk of glaucoma. Inadequate calibration or variability of calibration¹⁹ of tonometers used for treating or monitoring patients with OHT or glaucoma may cause apparently random variations in IOP, even if monitoring is performed on a single tonometer.

The effects of systematic errors in IOP measurements on over- and under-detection of OHT and treatment of glaucoma are not known and have not been studied. There is no published quantitative evidence indicating how accurate tonometers should be. In this study, we analysed IOP measurements from a population study to estimate the sensitivity of detection of OHT to systematic errors in IOP measurements.

Materials and methods

We analysed valid IOP measurements obtained in the Blue Mountain Eye Study (BMES).¹ The BMES was approved by the Western Sydney Area Health Service Human Ethics Committee, and written, informed consent was obtained from all participants. All IOP measurements were performed using a single Goldmann applanation tonometer and slit lamp (both Haag-Streit, Bern, Switzerland) after a drop of Fluress (benoxinate hydrochloride 0.4%, sodium fluorescein 0.25%, Barnes-Hind, Sunnyvale, CA, USA). A single measurement was taken and repeated if judged unreliable. All IOP measurements were performed by the same operator (a trained study orthoptist), during the 3-year course of the BMES. Any IOP measurements less than 10 or greater than 21 mm Hg were repeated by the Principal Investigator in the BMES (PM). The tonometer was calibrated regularly (at least 3 times per year during the course of the study).

We analysed all valid left, right and highest IOP measurements from all participants using a method similar to that published earlier for analysing blood pressure distributions.²⁰ IOP measurements were assumed to be rounded to the nearest mm Hg: histograms and cumulative distributions were calculated using thresholds at midpoints between integer IOP values. An inverse cumulative distribution indicating the percentage of subjects with IOPs greater than the abscissa value was calculated by summing the histograms and subtracting the sums from 100. First-, second- and third-order polynomials were fitted to the logarithm of the cumulative distribution by linear regression to smooth the distribution and facilitate interpolation. Goodness of fit of the polynomials was assessed by examining R^2 , the standard errors of the regressions and the statistical significance of the coefficients. Confidence intervals (95%) around mean predicted values were calculated.²¹

The proportion of subjects who had OHT (IOP > 21 mm Hg) was estimated from the polynomial. To estimate the effects of systematic measurement errors, the proportion was re-estimated assuming the tonometer systematically over- and under-read by 0–4 mm Hg at 0.1 mm Hg intervals. The sensitivity of over- and under-detection of OHT to the cutoff IOP was investigated by repeating the analysis for cutoff IOPs of 20 and 22 mm Hg. Data analysis was performed using Matlab 2009b (Mathworks, Natick, MA, USA).

Results

Histograms of right and left IOP measurements are shown in Figure 1. The difference between the IOP of left and right eyes was 0.09 ± 1.95 (mean \pm SD) mm Hg and the Pearson correlation coefficient was 0.78. Results for left, right and highest IOPs were similar, therefore we report results for highest IOPs only. The cumulative distribution indicating the percentage of subjects with highest IOP greater than 17.5–27.5 mm Hg is shown in Figure 2, in which the vertical scale is logarithmic. The SE of the first-, second- and third-order regressions were 0.0900, 0.0219, and 0.0233 logarithmic units, corresponding to 1.230, 1.052, and 1.055 percentage points, respectively. Respective R^2 values were 0.975, 0.999, and 0.999. The coefficients of the second-order polynomial were all significantly different from 0 at the 95% level of confidence. Therefore, the relationship between the logarithm of the inverse cumulative distribution and IOP is described best by a second-order polynomial.

Over- and under-detection of OHT (highest IOP > 21 mm Hg) caused by systematic errors of ± 2 mm Hg are indicated by the construction lines in Figure 2. Changes in the apparent prevalence of OHT that would be caused by systematic errors in IOP measurements of ± 1 , ± 2 , ± 3 and ± 4 mm Hg are

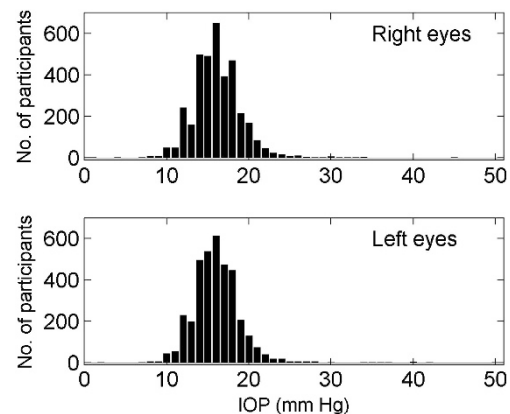


Figure 1 Histograms of all baseline intraocular pressures measured in the Blue Mountains Eye Study.¹

Table 1 Apparent prevalence (%) of ocular hypertension (highest IOP > 21 mm Hg) in the Blue Mountains Eye Study¹ associated with systematic errors in IOP measurements

| Systematic error (mm Hg) | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|-------------------|---------------------|---------------------|
| Mean prevalence (95% CI) (%) | 1.4 (1.3–1.5) | 1.9 (1.8–2.0) | 2.6 (2.5–2.7) | 3.8 (3.6–4.0) | 5.7 (5.4–6.0) | 9.0 (8.6–9.4) | 14.7 (14–15.5) | 25.3 (23.5–27.2) | 45.1 (40.6–50.2) |

Abbreviations: CI, confidence interval; Highest IOP, highest of left and right IOP of each participant; IOP, intraocular pressure. Percentage values are rounded to nearest 0.1%.

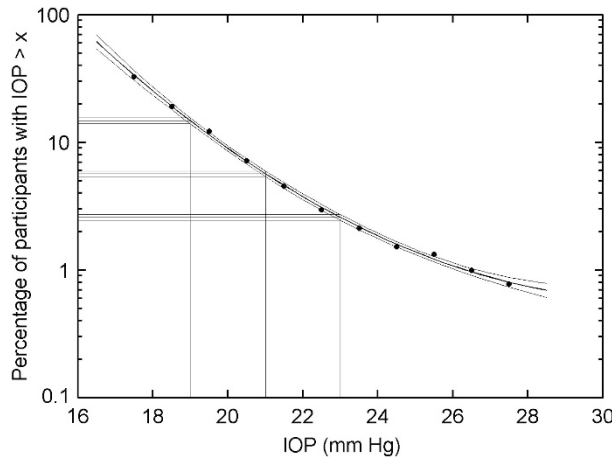


Figure 2 Inverse cumulative distributions showing proportions (%) of participants with highest IOP greater than the abscissa value. Lines above and below the regression line indicate 95% confidence interval (CI) of predicted mean value. Vertical and horizontal lines indicate percentages of participants (with corresponding 95% CI) in whom ocular hypertension (IOP > 21 mm Hg) would be detected when the tonometer has a systematic error of -2, 0, and +2 mm Hg, respectively.

shown in Table 1. Figure 3a and b show the effects of systematic errors in IOP on the detection of OHT over the continuous ranges 0 to +4 and -4 to 0 mm Hg, respectively, using cutoff IOPs of 20, 21, and 22 mm Hg.

Discussion

In this study, we show that detection of OHT is very sensitive to systematic errors in IOP measurements. When a population is screened for OHT, a tonometer that over-reads by 1 mm Hg may result in the number of positive results increasing by ~58% (Figure 3a) or ~3.3% of the population (Table 1). Conversely, a tonometer that under-reads by 1 mm Hg may result in ~34% of individuals with OHT (1.9% of the population) being missed (Figure 3b). Systematic errors of +1.5 mm Hg and -1.7 mm Hg may result in the number of individuals with detected OHT doubling or halving, respectively. Larger systematic errors of ± 4 mm Hg may result in the number of positive screening results increasing by a factor of seven (Figure 3a), or 76% of individuals with OHT being missed (Figure 3b). Some

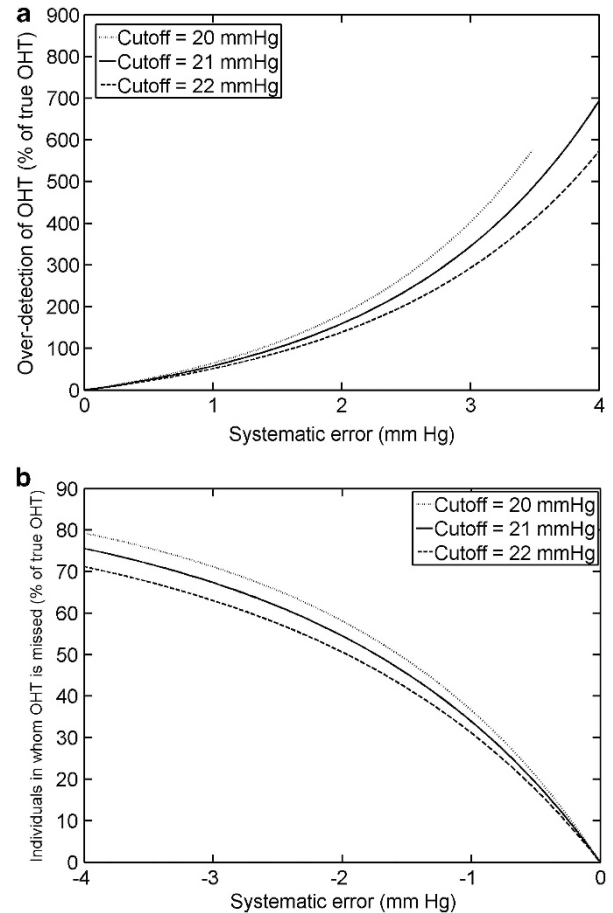


Figure 3 Over-detection of ocular hypertension caused by positive systematic measurement errors (a), and proportion of individuals with ocular hypertension who would be missed due to negative systematic measurement error (b). The dotted and broken lines indicate the effects of variations in cutoff IOP on over- and under-detection.

patients whose IOP is screened with an inadequately calibrated tonometer may be referred to an ophthalmologist unnecessarily and, some patients with OHT may not be referred and be at unnecessarily increased risk of glaucoma.

Figure 3a and b shows that under- and over-detection of OHT does not depend strongly on the IOP cutoff used to define OHT. For example, at a systematic error of +3 mm Hg, a change of cutoff from 21–22 mm Hg

changes over-detection from 345 to 293 (Figure 3a), a change of only $(293 - 345)/345 \times 100 = -15\%$. An earlier study that showed when the logarithm of the inverse cumulative distribution is linear, variations in the cutoff have no effect on over- and under-detection.²⁰ The weak non-linearity shown in Figure 2 is associated with small changes in under- and over-detection of OHT with changes in IOP cutoff.

Systematic error or bias may be common in both non-contact and applanation tonometers in regular clinical use. Three non-contact tonometers evaluated against a calibrated Goldmann applanation tonometer exhibited mean errors of 0.5, 2.9, and 1.8 mm Hg after 20, 9, and 24 months of operation, respectively.⁸ Choudhari *et al*⁹ checked the calibration of 132 Goldmann applanation tonometers, using the standard weight bar supplied by the manufacturer, 6 months after they had been calibrated by hospital biomedical engineers, and found calibration errors ranging from -8 to $+20$ mm Hg. Intra- and inter-observer limits of agreement for duplicate calibrations were up to 4 mm Hg.⁹ At 20 mm Hg, 55%, 10%, and 5% of tonometers had errors greater than or equal to 2, 4, and 6 mm Hg respectively.⁹ Only 4% of tonometers complied with the manufacturer's tolerance of ± 0.5 mm Hg. Sandhu *et al*¹² found that 6 months after 34 Goldmann tonometers had been calibrated by a specially designated technician, 41% exhibited calibration errors of 3 mm Hg or greater at a nominal 20 mm Hg. Four months later, 52% exhibited calibration errors of 3 mm Hg or greater.¹² George *et al*²² assessed *in vivo* agreement between two calibrated Goldmann applanation tonometers from different manufacturers and found limits of agreement of -2.47 to $+6.16$ mm Hg. Systematic measurement errors may also be caused by distortions of the applanating surface of disposable tonometer heads.²³

In a survey of 155 ophthalmology units in the United Kingdom, Kumar *et al*¹⁷ found that 39% of units checked the calibration of their Goldmann tonometers either never or in a random manner. In a separate UK survey of ophthalmology residents, 85% never checked tonometers and 70% felt that calibration checks were not part of their responsibility.¹⁶ Only 7% checked the tonometer at the start of each clinical session and 8% would only check the tonometer if they had suspicious or unexpected measurements.¹⁶

IOP is not the only variable used by ophthalmologists to estimate risk of loss of vision. Glaucoma occurs in many patients whose IOP is below 21 mm Hg. Clinical optic disc examination and other quantitative measurements²⁴ contribute to the detection, treatment and management of glaucoma. IOP, however, is an important risk factor for glaucoma²⁵ measured to screen for OHT, and it is important that the measurement is reliable.

Intra-individual variation of IOP can be several mm Hg,^{26,27} often substantially larger than calibration errors, possibly leading to an assumption that small systematic errors are not important. Systematic errors are approximately constant and create a bias at every IOP reading in every individual. The effect is similar to a shift in the population distribution of IOP. Decisions regarding treatment or follow-up are made, explicitly or implicitly, for every patient examined. On average, over a large number of patients, a small positive systematic error will result in more patients being treated or referred for further investigation, and a small negative systematic error will result in some patients not being treated or referred, who otherwise would be. This study quantifies those effects and shows that the influence of small systematic errors may be surprisingly large.

Important limitations of this study include the intra-individual variability of IOP of the participants and the possibility of undetected calibration errors in the tonometers used in the Blue Mountains Eye Study. Intra-individual variability and random calibration errors would increase the inter-individual IOP variance, broaden the IOP distribution and hence reduce the slope of the inverse cumulative distribution of IOP. This slope reduction would reduce the apparent sensitivity of OHT detection to tonometer calibration errors. Hence, this study may under-estimate the sensitivity.

We conclude that tonometers should be maintained and properly calibrated at intervals short enough to control systematic errors in IOP measurements to substantially less than 1 mm Hg.

Summary

What was known before

- Intraocular pressure (IOP) variation of 1 mm Hg may affect the risk of progressive nerve damage in glaucoma patients by 10%.
- Goldmann applanation tonometers may indicate IOP with errors of 4 mm Hg or greater within months of calibration.
- Many tonometers are not calibrated often enough.

What this study adds

- Systematic IOP errors of ± 1 mm Hg may cause an apparent increase in the prevalence of ocular hypertension (IOP > 21 mm Hg) of 58%, or 34% of individuals with ocular hypertension to be missed.
- Systematic IOP errors of ± 4 mm Hg may cause an apparent increase in the prevalence of ocular hypertension (IOP > 21 mm Hg) of 700%, or 76% of individuals with ocular hypertension to be missed.
- Tonometers should be maintained and calibrated to control systematic errors in IOP to substantially less than 1 mm Hg.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We gratefully acknowledge the assistance of George Burlutsky and Elena Rochtchina at Clinical Ophthalmology and Eye Health, Westmead Clinical School, C24 —Westmead Hospital, The University of Sydney. MJT was supported by an Innovation Fellowship from Macquarie University.

References

- 1 Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology* 1996; **103**(10): 1661–1669.
- 2 Kwon YH, Fingert JH, Kuehn MH, Alward WL. Primary open-angle glaucoma. *N Engl J Med* 2009; **360**(11): 1113–1124.
- 3 Noecker RJ. The management of glaucoma and intraocular hypertension: current approaches and recent advances. *Ther Clin Risk Manag* 2006; **2**(2): 193–206.
- 4 Chihara E. Assessment of true intraocular pressure: the gap between theory and practical data. *Surv Ophthalmol* 2008; **53**(3): 203–218.
- 5 Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B, Hussein M. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* 2002; **120**(10): 1268–1279.
- 6 Bengtsson B, Leske MC, Hyman L, Heijl A. Fluctuation of intraocular pressure and glaucoma progression in the early manifest glaucoma trial. *Ophthalmology* 2007; **114**(2): 205–209.
- 7 Sheehan W, Adams D, Wells C, Booth A. Does Goldmann applanation tonometry performed by community optometrists reduce referrals? A pilot study. *Br J Ophthalmol* 2011; **95**(2): 295.
- 8 Atkinson PL, Wishart PK, James JN, Vernon SA, Reid F. Deterioration in the accuracy of the Pulsair non-contact tonometer with use: need for regular calibration. *Eye (Lond)* 1992; **6**(Pt 5): 530–534.
- 9 Choudhari NS, George R, Baskaran M, Vijaya L, Dudeja N. Measurement of Goldmann applanation tonometer calibration error. *Ophthalmology* 2009; **116**(1): 3–8.
- 10 Kumar N, Jivan S, Batterbury M. Tonometer calibration. *Ophthalmology* 2009; **116**(12): 2480–2481.
- 11 Chuo JY, Mikelberg FS. Calibration errors of Goldmann tonometers in a tertiary eye care centre. *Can J Ophthalmol* 2007; **42**(5): 712–714.
- 12 Sandhu SS, Chattopadhyay S, Birch MK, Ray-Chaudhuri N. Frequency of goldmann applanation tonometer calibration error checks. *J Glaucoma* 2005; **14**(3): 215–218.
- 13 Wessels IF, Oh Y. Tonometer utilization, accuracy, and calibration under field conditions. *Arch Ophthalmol* 1990; **108**(12): 1709–1712.
- 14 Tattersall C, Vernon S, Wood C. Improving the accuracy of intraocular pressure readings. *Nurs Times* 2006; **102**(27): 36–38.
- 15 Brandt JD. The myth of clinical precision. *Ophthalmology* 2009; **116**(1): 1–2 e1.
- 16 Kumar N, Jivan S. Goldmann applanation tonometer calibration error checks: current practice in the UK. *Eye (Lond)* 2007; **21**(6): 733–734.
- 17 Kumar N, Hillier RJ. Goldmann tonometer calibration: a national survey. *Eye (Lond)* 2009; **23**(2): 413–415.
- 18 Ng J, Kashani S, Qureshi K, Ferguson V. Intraocular pressure measurement and goldmann calibration. An eye opening review of practice in United Kingdom. *Eye (Lond)* 2007; **21**(6): 851–853.
- 19 Choudhari NS, Jadhav V, George R, Vijaya L. Variability in the calibration error of the goldmann applanation tonometer. *J Glaucoma* 2011; **20**(8): 492–496.
- 20 Turner MJ, Baker AB, Kam PC. Effects of systematic errors in blood pressure measurements on the diagnosis of hypertension. *Blood Press Monit* 2004; **9**(5): 249–253.
- 21 Draper NR, Smith H. *Applied Regression Analysis*. 3rd ed. Wiley-Interscience: New York, 1998.
- 22 George R, Arvind H, Baskaran M, Ramesh SV, Raju P, Vijaya L. Agreement between two Goldmann type applanation tonometers. *Indian J Ophthalmol* 2008; **56**(6): 516–517.
- 23 Goel S, Chua C, Dong B, Butcher M, Ahfat F, Hindi SK *et al*. Comparison between standard Goldmann applanation prism and disposable applanation prism in tonometry. *Eye (Lond)* 2004; **18**(2): 175–178.
- 24 Quigley HA. Glaucoma. *Lancet* 2011; **377**(9774): 1367–1377.
- 25 Medeiros FA, Weinreb RN. Estimating the risk of developing glaucoma. *Open Ophthalmol J* 2009; **3**: 50–53.
- 26 Wilensky JT. Diurnal variations in intraocular pressure. *Trans Am Ophthalmol Soc* 1991; **89**: 757–790.
- 27 Wilensky JT, Gieser DK, Dietsche ML, Mori MT, Zeimer R. Individual variability in the diurnal intraocular pressure curve. *Ophthalmology* 1993; **100**(6): 940–944.