




# Application of Bayes' to the prediction of referral decisions made by specialist optometrists in relation to chronic open angle glaucoma

J. C. Gurney<sup>1,2</sup> · E. Ansari<sup>3</sup> · D. Harle<sup>4</sup> · N. O'Kane<sup>5</sup> · R. V. Sagar<sup>2</sup> · M. C. M. Dunne<sup>2</sup> 

Received: 23 June 2017 / Revised: 20 November 2017 / Accepted: 23 November 2017 / Published online: 9 February 2018  
© The Royal College of Ophthalmologists 2018

## Abstract

**Purpose** To determine the accuracy of a Bayesian learning scheme (Bayes') applied to the prediction of clinical decisions made by specialist optometrists in relation to the referral refinement of chronic open angle glaucoma.

**Methods** This cross-sectional observational study involved collection of data from the worst affected or right eyes of a consecutive sample of cases ( $n = 1,006$ ) referred into the West Kent Clinical Commissioning Group Community Ophthalmology Team (COT) by high street optometrists. Multilevel classification of each case was based on race, sex, age, family history of chronic open angle glaucoma, reason for referral, Goldmann Applanation Tonometry (intraocular pressure and interocular asymmetry), optic nerve head assessment (vertical size, cup disc ratio and interocular asymmetry), central corneal thickness and visual field analysis (Hodapp–Parrish–Anderson classification). Randomised stratified tenfold cross-validation was applied to determine the accuracy of Bayes' by comparing its output to the clinical decisions of three COT specialist optometrists; namely, the decision to discharge, follow-up or refer each case.

**Results** Outcomes of cross-validation, expressed as means and standard deviations, showed that the accuracy of Bayes' was high (95%, 2.0%) but that it falsely discharged (3.4%, 1.6%) or referred (3.1%, 1.5%) some cases.

**Conclusions** The results indicate that Bayes' has the potential to augment the decisions of specialist optometrists.

## Introduction

Optometric enhanced eye care services have been designed to overcome the burden on the Hospital Eye Service (HES) caused by glaucoma referrals that result in high first visit discharge rates [1]. The West Kent Clinical Commissioning Group Community Ophthalmology Team (COT) has provided such a service since 2006. This is led by a consultant ophthalmologist (EA) who provided hospital-based training

prior to accreditation (Optometrist with Special Interest in Ophthalmology), regular re-accreditation and direct access to advice for the specialist optometrists involved in this study. More about accreditation of optometrists can be found elsewhere [2]. The COT provides a glaucoma referral refinement service in the United Kingdom, as defined in the current Commissioning Guide, [3] as all patients undergo gonioscopy to exclude other forms of glaucoma. This service is provided for the West Kent, Medway, Dartford and Swanley areas.

Bayes' theorem predicts that high first visit discharge rates will occur for relatively rare diseases like chronic open angle glaucoma (COAG) even when the sensitivity and specificity of screening tests are high [4]. Application of Bayes' involves estimating the probability of an outcome by multiplying an initial estimate, based on prevalence of that outcome, by likelihood ratios derived from the sensitivity and specificity of each diagnostic test carried out [5, 6].

The objective of the present study was to determine how accurately Bayes' could predict clinical decisions made the specialist optometrists in the COT.

---

✉ M. C. M. Dunne  
m.c.m.dunne@aston.ac.uk

<sup>1</sup> Community Ophthalmology Team, West Kent CCG, Aylesford, UK

<sup>2</sup> Ophthalmic Research Group, Optometry School, Aston University, Birmingham, UK

<sup>3</sup> Maidstone and Tunbridge Wells NHS Trust, Maidstone, UK

<sup>4</sup> Community Ophthalmology Team, West Kent CCG, Tunbridge, UK

<sup>5</sup> Community Ophthalmology Team, West Kent CCG, Rochester, UK

**Table 1** Clinical methods that formed part of the COT SOP

Tests/method (COT SOP)
Tonometry (GAT):
IOP
ONHA (pupil dilation, slit lamp biomicroscopy, Volk lens):
VDS
VCDR
VCDR diff
CCT (ultrasound pachymetry)
VFA (SITA fast 20-4, HPA)
Gonioscopic examination of the anterior chamber (to exclude other causes of glaucoma)

*CCT* central corneal thickness, *COT* Community Ophthalmology Team, *GAT* Goldmann applanation tonometry, *HPA* Hodapp–Anderson–Parrish grade, *IOP* intraocular pressure, *IOP diff* interocular difference in IOP, *ONHA* optic nerve head analysis, *SOP* standard operating procedure, *VCDR* vertical cup-to-disc ratio, *VCDR diff* interocular difference in VCDR, *VDS* vertical disc size, *VFA* visual field analysis

## Subjects and methods

This study was approved by the Life and Health Sciences Research Ethics Committee at Aston University. It was treated as a clinical audit so that fully anonymised data collection was permitted without patient consent.

A consecutive sample of all referrals for suspected COAG seen by three COT optometrists (JCG, DH and NOK) over a period of 1 year (October 2014 to 2015) was included in this study. This amounted to 1,006 new cases referred into the COT from high street optometrists and data were only taken from the worst affected or right eyes.

A summary is provided of the clinical methods (Table 1) and multilevel groups (Table 2) used in this study. The use of multilevel groups has been advocated by others [5, 6]. The clinical methods used were part of a standard operating procedure (SOP) adopted by the COT that was informed by The National Institute for Health and Care Excellence (NICE) guideline CG85 [7].

Intraocular pressure was measured using a Haag-Streit AT-900 Model T Goldmann applanation tonometer (GAT IOP) and multilevel groups followed NICE guidelines. [7] Groups for interocular differences in pressure (GAT IOP diff) were based on a previous study [8].

Dilated stereoscopic slit lamp biomicroscopy (Haag-Streit BQ900 or Topcon PS30) with a Volk lens (66D or Digital 1×) was used for optic nerve head assessment (ONHA). Previous studies informed multilevel groups for vertical optic disc size (VDS), [9] vertical cup-to-disc ratio (VCDR) [10, 11] and interocular differences in VCDR (VCDR diff) [12].

Central corneal thickness (CCT) was measured using handheld ultrasound pachymetry (Accutome Pachpen,

Pachmate or Pachmate 2) and multilevel groups were, again, based on NICE guidelines. [7]

The SITA fast (Swedish interactive thresholding algorithm) 24-2 testing strategy is recommended by NICE guidelines for visual field assessment (VFA) [7]. The Zeiss Humphrey Visual Field Analyser (model 720 or 720i) was used and multilevel groups followed the Hodapp–Anderson–Parrish (HAP) grading system [13].

Management decisions of COT optometrists were (a) discharge, (b) follow-up in the COT for suspected COAG or (c) referral to the HES for COAG diagnosis.

Table 3 shows the equations used for Bayes' [5]. Decision matrices were constructed for the 41 multilevel groups shown in Table 2 and the three COT management decisions (discharge, follow-up or refer); ( $41 \times 3 =$ ) 123 decision matrices in total. Each decision matrix contained frequencies of true and false positives and negatives. This gave rise to 41 sets of likelihood ratios for positive (TEST+) and negative (TEST−) test outcomes. Calculation of the probability of any of the three COT management decisions then involved determining the product of all 41 positive or negative likelihood ratios, depending on each test outcome. The final step was to select the COT management decision with the highest probability.

A well-known problem with decision matrices is the occurrence of zero frequency counts which give rise to likelihood ratios of zero [14]. It follows that the product of a number of likelihood ratios, including one of zero value, would also be zero. This would absolutely rule out a COT management decision when, in reality, no statistical model is perfect enough to do this. The solution is usually to make a Laplacian correction by adding 1 to the counts in each cell of a diagnostic matrix [14], but this also leads to small artificial alterations to calculated likelihood ratios. A Laplacian correction of 0.001 was used in the present study to ensure that such alterations were minimised.

Decisions made by Bayes' were compared to those of the specialist optometrists. The simplest way to evaluate Bayes' could have been to calculate likelihood ratios based on all 1,006 cases (the training phase) and then to test how well these predicted optometrists' decisions on the same 1,006 cases (the testing phase). The problem with this approach was that training and testing would then have been carried out on the same cases, leading to a very optimistic assessment of accuracy.

Randomised stratified tenfold cross-validation was used instead [14]. Here, the dataset of 1,006 cases was divided into tenfolds of about 100 cases each. The cases in each fold were selected randomly and stratification ensured that each COT decision was equally represented. Each fold, in turn, was used in the testing phase with all other folds used in the training phase. Treating the data in this way delivered the most realistic estimate of accuracy [14].

**Table 2** Multilevel groups adopted and their frequency of occurrence in 1,006 cases

Tests	Groups (frequency)—likelihood ratios calculated for each [5]
Age (years)	<40 (60), 40–49 (135), 50–59 (242), 60–69 (275), 70–79 (218), 80+ (78)
Sex	Female (524), male (482)
Race	Caucasian (965), Asian (17), Afro-Caribbean (12), Hispanic (9), African (3)
FGH	Mother (143), father (110), sibling (62)
GAT IOP (mm Hg)	<21 (773), 21–25 (181), >25–32 (43), >32 (9)
GAT IOP diff (mm Hg)	<3 (926), 3–6 (53), >6 (27)
ONHA VDS (mm)	<1.4 (87), 1.4–1.7 (439), 1.8+ (480)
ONHA VCDR (%)	<50 (496), 50–70 (449), >70 (71)
ONHA VCDR diff (%)	<20 (865), 20–30 (121), >30 (20)
CCT (µm)	<555 (443), 555–590 (374), >590 (189)
VFA HPA	Mild (91), moderate (44), severe (29)

CCT central corneal thickness, FGH family glaucoma history, GAT Goldmann applanation tonometry, HPA Hodapp–Anderson–Parrish grade, IOP intraocular pressure, IOP diff interocular difference in IOP, ONHA optic nerve head analysis, SOP standard operating procedure, VCDR vertical cup-to-disc ratio, VCDR diff interocular difference in VCDR, VDS vertical disc size, VFA visual field analysis

**Table 3** Equations for Bayes’ based on decision matrices containing the frequency of TP, FP, TN and FN for every multilevel group (shown in Table 2) and COT management decision

	DECISION+	DECISION–
TEST +	TP	FP
TEST–	FN	TN
Prevalence = (TP + FN)/(TP + FN + TN + FN)		
Pre-test odds = prevalence/(1 – prevalence)		
Sensitivity = TP/(TP + FN)		
Specificity = TN/(TN + FP)		
Likelihood ratio (for TEST +) = sensitivity/(1 – specificity)		
Likelihood ratio (for TEST–) = (1 – sensitivity)/specificity		
Post-test odds = pre-test odds x product of likelihood ratios for all tests		
Post-test probability = post-test odds/(1 + post-test odds)		

COT Community Ophthalmology Team, FP false positives, FN false negatives, TN true negatives, TP true positives

Results were initially expressed in the form of ten separate confusion matrices [14], one for each cross-validation run. These matrices simplified side-by-side comparisons of the Bayes’ and specialist optometrists’ decisions to discharge, follow-up or refer. Accuracy was expressed as the percentage of cases for which Bayes’ matched decisions made by specialist optometrists. A weighted accuracy was calculated for each confusion matrix, being a single quantity that simultaneously expressed the accuracy for all three COT decisions [14]. Confusion matrices also simplified calculation of percentage false discharge and false referral rates that would have arisen, theoretically, had Bayes’ decisions replaced those of the specialist optometrists. Averages and standard deviations were calculated from the ten confusion matrices for

**Table 4** Confusion matrix comparing Bayes’ management decisions to those of the specialist optometrists

	Bayes’			
	Discharge	FUP	Refer	
Specialist optometrists	Discharge	<b>78.3 (0.8)</b>	0.2 (0.4)	0.8 (0.4)
	FUP	3.0 (1.6)	<b>1.7 (1.3)</b>	2.3 (1.5)
	Refer	0.4 (0.5)	1.1 (0.9)	<b>12.2 (1.2)</b>

Percentages shown were averaged from ten cross-validation runs. Standard deviations are shown in brackets. The bold figures shown diagonally are those in which agreement occurred. Use of Bayes’ would have led to some false discharges and false referrals (italics figures). Management decisions included discharge, follow-up in COT for suspected COAG (FUP) or referral to the HES to confirm COAG (Refer)

COAG chronic open angle glaucoma, COT Community Ophthalmology Team, FUP follow-up in COT for suspected COAG, HES Hospital Eye Service

weighted accuracy, false discharge and false referral rates. These average values are shown, for brevity, in a single confusion matrix (Table 4).

## Results

Table 4 summarises the findings of this study. Summing the percentages shown in rows gives the total percentage of each management decision made by the specialist optometrists. Summing the percentages shown in columns gives the same for Bayes’.

The average weighted accuracy of Bayes’ was 95.4% (standard deviation 1.6%). Note that this does not equal the sum of the percentages shown in bold in Table 4, which, instead, show that the management decisions of Bayes’

matched the specialist optometrists 92.2% of the time. Replacing the decisions of specialist optometrists with Bayes' would have resulted in an average false discharge rate (Table 4) of 3.4% (standard deviation 1.6%) and an average false referral rate (Table 4) of 3.1% (standard deviation 1.5%).

## Discussion

As far as we are aware, this is the first study to have reported the accuracy of a Bayesian learning scheme applied to the prediction of clinical decisions made by specialist optometrists relating to referral refinement of COAG.

Although the accuracy of Bayes' was high, it still gave rise to some false discharges and referrals. False discharges risk avoidable vision loss while false referrals risk avoidable NHS burden. We explored different methods of making Bayes' cost sensitive. As Bayes' works by choosing the management decision that has the highest probability, the simplest method of adding cost sensitivity [14] was to weight one or more of the management decision probability values in order to move false discharges and referrals to follow-up. Various weightings were trialled and all successfully removed false discharges and referrals but at the cost of a dramatic increase in follow-ups.

There were two limitations to this study. The first of these was that its conclusions have been based on application of the simplest form of Bayes' theorem. Although this sort of learning scheme may perform just as well as more sophisticated machine learning methods [14], no attempt was made in this study to confirm this. The second limitation relates to the use of multilevel groups based on NICE guidelines and previous literature. No attempt was made to discover whether better groupings may have improved the accuracy of Bayes'.

The assumption was that the COT always made the correct diagnosis. In the longer term, it would be useful to see how COT decision compare to the Bayes' model that has arisen from this study.

To conclude, the findings of this study indicate that this simple form of Bayes' has the potential to augment rather than replace the decisions of specialist optometrists. This facility may be useful when borderline cases are encountered. Further research on more sophisticated learning schemes or improved multilevel groupings may lead to increased accuracy in the future.

## Summary

### What was known before

- Specialist optometrists working in ophthalmologist-led Glaucoma Referral Refinement centres already play a useful role in reducing unnecessary referrals to the Hospital Eye Service.
- Bayes' theorem predicts that high numbers of unnecessary referrals inevitably arise for rare conditions such as glaucoma.

### What this study adds

- Artificial intelligence based on the simplest form of Bayes' theorem can match referral decisions of specialist optometrists with remarkable accuracy.
- However, further research is needed before this form of artificial intelligence is accurate enough to replace specialist optometrists without adding to the risk of false discharges and referrals.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no competing interests.

## References

1. Baker H, Ratnarajan G, Harper RA, Edgar DF, Lawrenson JG. Effectiveness of UK optometric enhanced eye care services: a realist review of the literature. *Ophthalmic Physiol Opt.* 2016;36:545–57.
2. Harper RA, Vernon S, Spry P. Postgraduate specialist glaucoma training and accreditation in optometry. *Optom Pract.* 2013;14:125–36.
3. Commissioning Guide: Glaucoma (Long Version). June 2016. <https://www.rcophth.ac.uk/wp-content/uploads/2016/06/Glaucoma-Commissioning-Guide-Long-June-2016-Final.pdf>. Accessed 22 June 2017.
4. Thomas R, Mengersen K, Parikh RS, Muliyl J. Enter the reverend: introduction to and application of Bayes' theorem in clinical ophthalmology. *Clin Exp Ophthalmol.* 2011;39:865–70.
5. Parikh R, Parikh S, Arun E, Thomas R. Likelihood ratios: clinical application in day-to-day practice. *Indian J Ophthalmol.* 2009;57:217–21.
6. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine.* Boston, USA: Little, Brown & Co; 1991.
7. Glaucoma: Diagnosis and Management. NICE Guidelines [CG85]. April 2009. <https://www.nice.org.uk/guidance/cg85>. Accessed 22 June 2017.
8. Williams AL, Gatla S, Leiby BE, Fahmy I, Biswas A, de Barros DM, et al. The value of intraocular pressure asymmetry in diagnosing glaucoma. *J Glaucoma.* 2013;22:215–8.
9. Crowston JG, Hopley CR, Healey PR, Lee A, Mitchell P. The effect of optic disc diameter on vertical cup to disc ratio percentiles in a population based cohort: the Blue Mountains Eye Study. *Br J Ophthalmol.* 2004;88:766–70.
10. Garway-Heath DF, Ruben ST, Viswanathan A, Hitchings RA. Vertical cup/disc ratio in relation to optic disc size: its value in the assessment of the glaucoma suspect. *Br J Ophthalmol.* 1998;82:1118–24.

11. Gloster J. Quantitative relationship between cupping of the optic disc and visual field loss in chronic simple glaucoma. *Br J Ophthalmol.* 1978;62:665–9.
12. Ong LS, Mitchell P, Healey PR, Cumming RG. Asymmetry in optic disc parameters: the Blue Mountains eye study. *Invest Ophthalmol Vis Sci.* 1999;40:849–57.
13. Anderson D. Automated static perimetry. St Louis, USA: Mosby-Year Book Inc.; 1992.
14. Witten IH, Frank E, Hall MA. Data mining. Practical machine learning tools and techniques. 3rd ed. Oxford, UK: Elsevier; 2011.