Citation

Myint J, Edgar DF, Murdoch IE, Lawrenson JG. The impact of postgraduate training on UK optometrists' clinical decision-making in glaucoma. Ophthalmic Physiol Opt. 2014 May;34(3):376-84. doi: 10.1111/opo.12126.

The impact of postgraduate training on UK optometrists' clinical decision-making in glaucoma

Myint J[1], Edgar DF[2], Murdoch IE[3], Lawrenson JG[2]

1. *Department of Vision and Hearing Sciences, Anglia Ruskin University, Cambridge, UK*
2. *Henry Wellcome Laboratories for Vision Sciences, Division of Optometry, City University London, UK*
3. *Glaucoma Unit, Moorfields Eye Hospital, London, UK*

Correspondence*: J. Myint joy.myint@anglia.ac.uk*

Keywords: glaucoma, postgraduate training, optometrist, clinical decision making

Disclosure: The authors report no conflict of interest and have no proprietary interest in any of the materials mentioned in this article

Acknowledgements: The authors would like to thank Professor David Henson, Dr Paul Artes and Dr Jonathan Denniss for use of the Discus program. Joy Myint was funded by an unrestricted grant from Pfizer Ophthalmology.

Abstract

Purpose: to investigate the impact of a postgraduate training module on optometrists' clinical decision-making in relation to the diagnosis and management of primary open-angle glaucoma .

Methods: a group of United Kingdom community optometrists (n=53) were assessed immediately before and again three months after completing a 3-day didactic postgraduate university module on the diagnosis and management of glaucoma. A smaller control cohort (n = 20), who did not receive the intervention, was recruited and completed the same assessments on two occasions, separated by approximately 3 months. The assessments comprised: knowledge of 5 key features of the optic disc in glaucoma, performance on a computer program (Discus) that assessed the ability to differentiate normal from glaucomatous discs and a clinical decision-making exercise using case-based scenarios.

Results: the scores for the knowledge of important disc features for the intervention cohort significantly increased from a median of 2/5 to 5/5 post-intervention ($P<0.001$). For the Control cohort, the difference in median scores between the two tests was not significant. Analysing the performance of the intervention cohort using the Discus program showed no significant improvement in ability to diagnose a glaucomatous disc following the intervention (mean area under the receiver operating characteristic curve pre-intervention=0.85 (95%CI: 0.76-0.91), post-intervention=0.84 (95%CI: 0.76-0.91)). Similarly, there were no statistically significant differences in mean areas under the receiver operating characteristic curves between tests for the control cohort, although both cohorts compared favourably with a previously published Discus data set from a panel of experts in disc analysis (mean area under the receiver operating characteristic curve=0.87). For the clinical decision-making exercise the median test score for the intervention cohort was unchanged pre- and post-intervention.

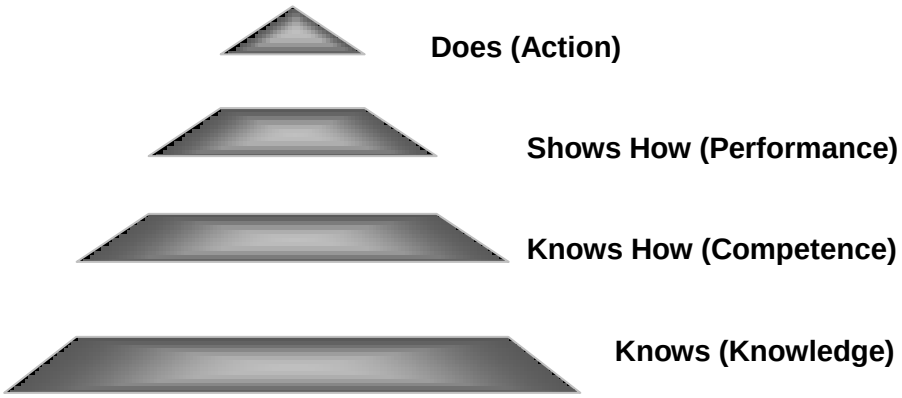
Conclusion: The results of the present study suggest that a traditional didactic approach, in isolation, is unlikely to be suited to training optometrists to achieve or develop the clinical competencies required for glaucoma detection and management. Consideration should be given to the development of specialist postgraduate training that is more practice-based, provides opportunities for active learning and includes strategies for feedback and reinforcement.

Introduction

There is currently no population-based screening programme for primary open angle glaucoma (POAG) in the UK.[1] In the absence of formal screening, community optometrists continue to play a key role in glaucoma detection with over 95% of referrals to secondary care for glaucoma originating from optometrists.[2] Detection of glaucoma and suspect glaucoma by optometrists is achieved by opportunistic case-finding and is of necessity limited to persons who attend for eye examinations. Optometrists acquire diagnostic skills for the detection and appropriate referral of glaucoma during their training, which consists of an undergraduate degree followed by a pre-registration period of supervised practice. During this period, optometrists must demonstrate that they are proficient in a number of 'core competencies' defined by the General Optical Council (GOC).[3] The term 'core competency' is used to describe the knowledge and skill elements that an optometrist must possess in order to register and practise within the UK. Although these 'entry level' competencies encompass the knowledge and skills to detect glaucoma, it is recognised that additional training is required for further specialisation.[4] The last decade has seen considerable interest in the development of postgraduate training and additional qualifications within this speciality to allow optometrists to refine referrals for suspect glaucoma and to provide care for those already diagnosed with glaucoma, suspect glaucoma or ocular hypertension.[5]

Postgraduate training programmes in glaucoma in the UK generally utilise a conventional didactic approach consisting of lectures that can either be delivered face-to-face or online, augmented by practical sessions and case-based discussions. Assessment of clinical competence is integral to the educational process, in order to help trainees learn and develop and to provide evidence of progression. A useful theoretical framework for competency-based assessment was proposed by the psychologist George Miller.[6] This framework conceptualises the essential facets of clinical competence as a pyramid (Figure 1). The base of the pyramid (knows) represents the knowledge required to perform a particular task.

Figure 1: Miller's pyramid of clinical competence.



Ascending to the next level of the pyramid we reach the "knows how" region, which describes the clinician's ability to use knowledge in a particular context. An optometrist operating at this level would be using clinical reasoning and problem solving. Assessment of these skills is increasingly being carried out by presenting the trainee with a clinical scenario (paper-based or online). In the assessment, the trainee selects those procedures and management choices appropriate for the patient described in the scenario.

At the next level, the "shows how" region of the pyramid allows an assessment of the trainees ability to perform appropriately using artificial simulations or via objective structured clinical examinations (OSCE), where candidates rotate through a series of stations that test a sample of clinical skills in a range of contexts. Wide sampling and structured assessment improve reliability.[7]

The top section of the pyramid refers to actual performance in habitual practice (the "does" level). At this level, the skills being tested are those directly related to the real-life practice environment. Therefore, the assessment at this "does" level needs to be as clinically authentic as possible. This "action" or "does" component of professional behaviour is the most difficult to measure reliably and accurately.[6] Research into the performance of optometrists at this highest level of Miller's pyramid is scant.[8]

Assessment tools used in specialist training programmes in glaucoma are generally based on the first three levels of the pyramid[5]. For example, for glaucoma diagnosis,

assessments may include: the ability to perform diagnostic tests, interpret test results and integrate clinical findings to make a diagnosis. The identification of pathological changes in the optic disc is a key skill for the initial detection of glaucoma and for the identification of progression since these changes often precede visual field defects.[9-10] Previous studies have reported on levels of agreement and accuracy of eye care professionals in detecting glaucomatous disc changes.[11-16] These studies suggest that a greater consistency in disc assessment and overall diagnostic ability occurs with experience. A recent study of the diagnostic accuracy of UK optometrists in classifying optic disc photographs as healthy or glaucomatous[16] found that additional qualifications and experience in hospital glaucoma clinics improved performance.

This aim of this study is to investigate the impact of a postgraduate training module in glaucoma on optometrists' clinical decision-making in relation to POAG. In particular, the ability to identify the features of a glaucomatous optic disc and to make diagnostic decisions based on clinical case scenarios.

Methods

The educational intervention consisted of a 3-day didactic postgraduate module, 'Optometric Management of Glaucoma' that forms part of the MSc in Clinical Optometry at City University London. The module provides a series of lectures and practical sessions covering the diagnosis and management of glaucoma.

A group of UK community optometrists, referred to hereafter as the "MSc cohort" (n=53), were assessed both immediately before and again three months after completing this module. The MSc cohort comprised optometrists who wished to obtain additional training in Glaucoma by taking this Masters level module. Some were studying for a higher qualification from City University (a Diploma, Certificate or Masters in Clinical Optometry) and were using this module as one of the modules contributing towards this qualification. The remainder were taking the module in isolation to increase their knowledge of glaucoma. Optometrists based in secondary care were excluded from the study. The educational intervention consisted of a 3-day didactic postgraduate module.

A smaller cohort (the "Control cohort", n = 20) of community optometrists who had not previously attended the City University glaucoma module were used to control for a potential testing effect. The College of Optometrists agreed that the authors could recruit the control cohort through their Optometric Collaborative Research Network (OCRN), a network of community optometrists with an interest in primary care research. The only exclusion criterion was that participants must not have undertaken any form of additional training in glaucoma. The final choice of controls was made by endeavouring to ensure that the sample was representative of optometrists on the GOC register.

The Control Cohort completed the same assessment exercise as the MSc Cohort on two occasions, again separated by approximately 3 months, but without undergoing the

educational intervention. Though there was no educational intervention with the Control cohort, for convenience the two assessments for this cohort will also be referred to as "Pre-intervention" and "Post-intervention" to facilitate comparison with the MSc cohort.

There were three elements to the assessment:

- Knowledge of the key features of the optic disc in glaucoma
- Performance on a computer program (Discus)[17] that assessed the optometrists' ability to differentiate normal from glaucomatous discs
- Assessment of clinical decision-making for the detection of suspect POAG.

For the first element subjects were requested to list the five most relevant features that "*should be observed and/or considered when assessing a patient's disc for possible open angle glaucoma*". This was a paper-based exercise, with 5 being the maximum score.

An expert panel, which included those delivering the glaucoma module, established the definitive list of features for the purpose of this study. In alphabetical order, these features are; asymmetry of discs, disc haemorrhage, lamina cribrosa appearance, neuro-retinal rim appearance, retinal nerve fibre layer appearance, optic disc size, and peri-papillary atrophy. These disc features reflected the material taught during the module.

The second assessment utilised the Discus software package.[17] The program assesses clinicians' subjective judgement of the likelihood of damage in a series of discs presented on a computer monitor. Previous research using the Discus program has led to the development of a reference standard, generated by 12 glaucoma specialists (Discus Expert Panel), against which other clinicians can judge their performance.[17] The Discus Expert Panel comprised 10 ophthalmologists working in glaucoma speciality clinics and two specialist optometrists whose research interests included the optic disc in glaucoma.[17]

The optic disc images used in the Discus program were selected from patients with either diagnosed or suspected glaucoma or ocular hypertension, who attended the Optometrist-led Glaucoma Assessment (OLGA) clinics at the Royal Eye Hospital (Manchester, UK). Two groups of patients were established; those classified as visual field positive ("repeatable field loss") (n=20) and a second group who were classified as

visual field negative (“repeatable no field loss”) (n=80). The image quality of the disc images in each group was matched in an effort to eliminate any bias. For the current study the program displayed each disc image in a randomised order, for a maximum of 30 seconds, though the time allowed for making a decision was unlimited. Participants were required to rate the optic disc on a 5-point Likert scale (definitely healthy, probably healthy, not sure, probably damaged, and definitely damaged). Twenty-six images were presented twice (2 in the “damaged” group and 24 in the “healthy” group) to check the consistency of responses. Discus also records the “latency”, or the time taken to make the decision for each disc image.

For the final task, subjects reviewed 4 clinical case scenarios and were asked a series of clinical decision-making questions (one for each scenario) relating either to diagnosis or management. Scenarios provided all relevant clinical information, for each case, including patient history, field plots and photographs of optic discs. Answers to each question were recorded on a 5-point Likert scale. The expert panel agreed on a reference answer for each scenario. If the participant’s answer agreed with the panel reference answer they scored 2 points, or they scored 1 point if their answer differed from the reference answer but was still considered to be clinically acceptable. Incorrect answers scored zero. The maximum score for the clinical decision-making task was 8 points.

Ethical approval for these studies was granted by the City University School of Health Sciences Research and Ethics Committee and the research was carried out in compliance with the Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html).

Both parametric and non-parametric methods were used to analyse the data. For the ‘knowledge of important disc features’ data, a score was recorded for each subject, requiring a non-parametric analysis of the medians using either the Wilcoxon test for two paired samples or the Mann-Whitney test for two unpaired independent samples. The ‘clinical decision making’ data were also scores and required a similar approach to the statistical analysis. Sensitivity, specificity and latency data from the Discus programme were normally or approximately normally distributed and were analysed using parametric methods employing either the paired or unpaired (two sample) ‘t’ test.

Results

The mean scores for the knowledge of important disc features for the MSc cohort increased from 2.3 to 4.4 post-intervention (Table 1). There was a statistically significant improvement in the median score to 5 post-intervention compared with a score of 2 pre-intervention ($P<0.001$; Wilcoxon Statistic = 1308.0). For the Control cohort the mean scores on this exercise also increased, from 2.9 to 3.1 after three months but there was no statistically significant difference between median scores (Median = 3 both pre- and post-intervention).

Comparing the MSc and Control cohorts there was no statistically significant difference between the median scores pre-intervention ($p = 0.10$, $U = 663.5$, Mann-Whitney test) although the difference in median scores (3 for Controls and 5 for MSc cohort) was significant post-intervention ($p < 0.001$, $U = 869.5$).

Table 1: Mean and median number of optic disc features correctly identified by the Control cohort (n=20) and the MSc cohort (n = 53) pre- and post- the educational intervention. Scores given are out of a maximum of 5.

Cohort	**Pre**		**Post**	
	Mean	**Median**	**Mean**	**Median**
Control	**2.9**	**3**	**3.1**	**3**
MSc	**2.3**	**2**	**4.4**	**5**

For the Discus program (Table 2), in the MSc cohort the difference between the mean sensitivities ('sensitivity' is defined here as the percentage of visual field positive patients identified as having damaged discs) pre-intervention (74%) and post-intervention (81%) is statistically significant ($p = 0.0049$, $t = 2.94$, $df = 52$, Paired t-test). The difference between the mean specificities (defined as the percentage of visual field negative patients identified as having normal discs) pre-intervention (64%) and post-intervention (55%) was also statistically significant ($p = 0.0014$, $t = 3.37$, $df = 52$, Paired t-test). For the calculation of sensitivities and specificities the selection of the option "Not sure" for the optic disc appearance was interpreted as a "damaged" response. The rationale is that an optometrist who is "not sure" about the appearance of an optic disc is more likely to diagnose a patient as a 'glaucoma suspect' than not.

For the Control cohort the difference between the mean sensitivities pre-intervention (59%) and post-intervention (58%) was not statistically significant ($p = 0.78$, $t = 0.29$, $df = 19$, Paired t-test). The difference between the mean specificities pre-intervention (60%) and post-intervention (61%) was also not statistically significant ($p = 0.74$, $t = 0.34$, $df = 19$, Paired t-test).

For pre-intervention sensitivity the difference between mean sensitivities for the MSc cohort (74%) and the Control cohort (59%) was statistically significant ($p = 0.0006$, $t = 3.61$, $df = 71$, Unpaired t-test). For post-intervention sensitivity the difference between mean sensitivities for the MSc cohort (81%) and the Control cohort (58%) was also statistically significant ($p < 0.0001$, $t = 5.25$, $df = 71$, Unpaired t-test).

For pre-intervention specificity the difference between mean specificities for the MSc cohort (64%) and the Control cohort (60%) was not statistically significant ($p = 0.26$, $t = 1.14$, $df = 71$, Unpaired t-test). For post-intervention specificity the difference between mean specificities for the MSc cohort (55%) and the Control cohort (61%) was also not statistically significant ($p = 0.17$, $t = 1.38$, $df = 71$, Unpaired t-test).

Table 2: Performance in the Discus program for the Control Cohort (n=20) and the MSc Cohort (n = 53) pre- and post-intervention.

		Sensitivity %	Specificity %	Sensitivity %	Specificity %
		Pre	**Pre**	**Post**	**Post**
Control	**Mean**	**59**	**60**	**58**	**61**
MSc	**Mean**	**74**	**64**	**81**	**55**

The sensitivity and specificity data allowed composite Receiver Operating Characteristics (ROC) curves to be generated for both cohorts pre- and post-intervention using Medcalc software (http://www.medcalc.org/) (Figures 2 & 3). The areas under the ROC (AUROC) curves were:

MSc Pre-intervention = 0.85 (95% CI 0.76 – 0.91)
MSc Post-intervention = 0.84 (95% CI 0.76 – 0.91)
Controls Pre-intervention = 0.84 (95% CI 0.76 – 0.91)
Controls Post-intervention = 0.91(95% CI 0.83 – 0.96)

There were no statistically significant differences between any of the AUROCs either within or between cohorts pre- or post-intervention.

Figure 2: Composite ROC curves for MSc cohort pre- and post-intervention.

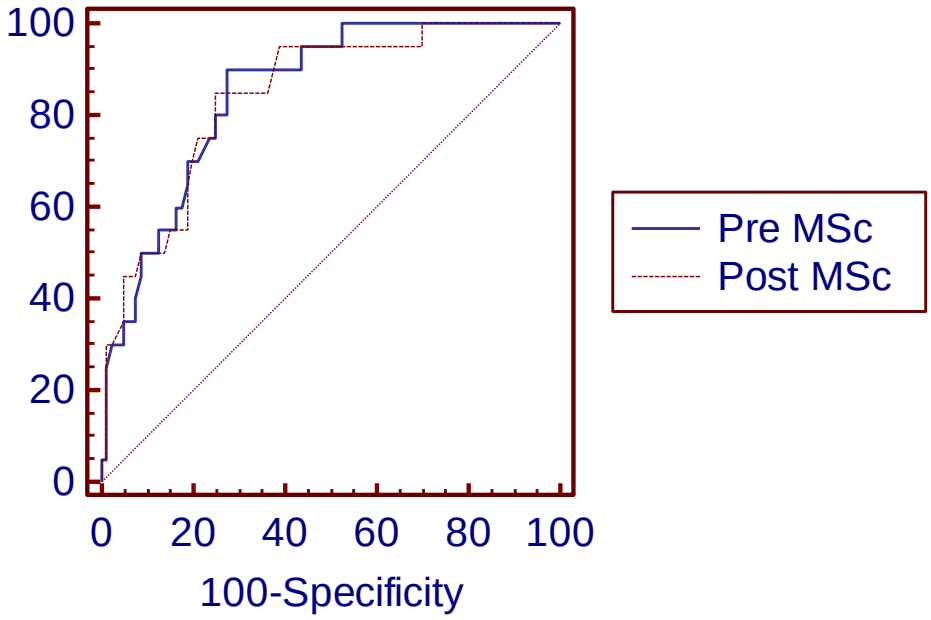
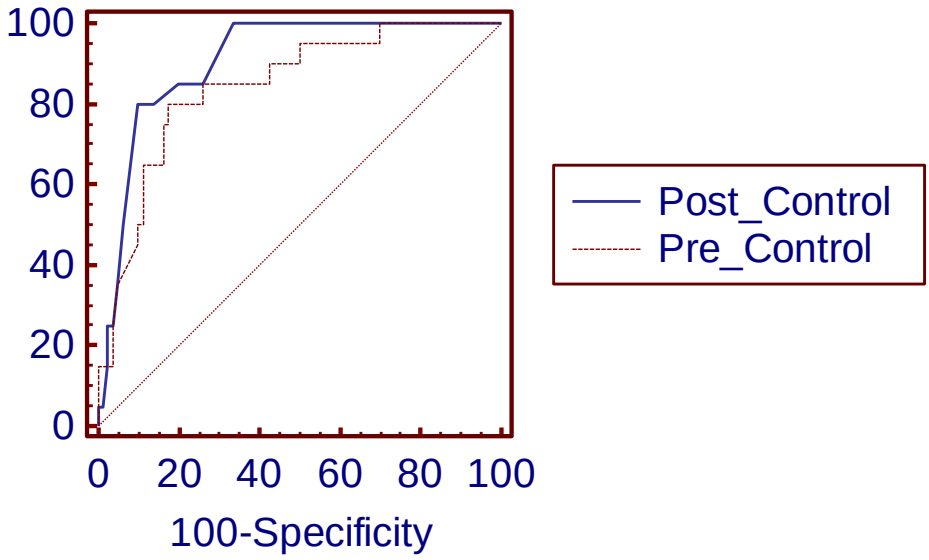


Figure 3: Composite ROC curves for the Control cohort pre- and post-intervention.



The repeatability of responses was analysed for the MSc cohort for both the pre-intervention and post-intervention data by taking the difference between the first score for each repeated image (where 5 = definitely damaged and 1 = definitely healthy) and the second score. Agreement (zero difference) between the first and second scores

occurred in 58% of repeats both pre-intervention and post-intervention. Discrepancies of at least one category occurred in 42% of repeats both pre- and post-intervention. For the pre-intervention data the distribution of the 42% of discrepancies was almost perfectly symmetrical between discrepancies in the positive ("healthier" disc on repeat) and negative directions. The 42% comprised 31% with one category difference on repeat (15% a negative difference, and 16% positive), 8% with two categories difference (4% positive and 4% negative), and 2% with three categories difference (1% positive and 1% negative). Two subjects obtained the maximum difference of 4 categories (one positive and one negative) although the numbers are so low that these registered as zero in percentage terms. For the post-intervention data, the distribution of the repeats was slightly skewed in the positive direction (healthier discs) on repeat. The 42% comprised 28% with one category difference on repeat (15% positive and 14% negative, 10% with two categories difference (6% positive and 4% negative), 2% three categories difference (equally split between positive and negative), and 1% (9 repeats) which had the maximum possible 4 categories difference. All these 9 discs that had four categories of difference were in the positive direction i.e. discs that were rated 5 (definitely damaged) on first presentation but were rated 1 (definitely healthy) on the repeat.

Repeatability was higher for the Controls, with agreement (zero difference) between the first and second scores occurring in 68% of repeats pre-intervention and 71% post-intervention. The distribution between positive and negative differences on repeat presentation was almost perfectly symmetrical both pre- and post-intervention, and there were no discs with four categories of difference.

For the MSc cohort the difference between the mean latencies (time to reach a decision on the disc image) pre-intervention (7.4s) and post-intervention (11.0s) was statistically significant ($p < 0.0001$, $t = 6.32$, $df = 52$, Paired t-test). For the Control cohort the difference between the mean latencies pre-intervention (13.6s) and post-intervention (13.1s) was not statistically significant ($p = 0.70$, $t = 0.40$, $df = 19$, Paired t-test).

For pre-intervention latency the difference between mean latencies for the MSc cohort (7.4s) and the Control cohort (13.6s) was statistically significant ($p < 0.0001$, $t = 6.69$, $df = 71$, Unpaired t-test). For post-intervention latency the difference between mean latencies for the MSc cohort (11.0s) and the Control cohort (13.1s) was not statistically

significant (p = 0.15, t = 1.46, df = 71, Unpaired t-test).

Analysis of the distribution of the mean scores pre-intervention for each disc image for the MSc and Control cohorts demonstrated a difference between the two distributions, with the Control scores tightly bunched around the median of 2.6 and no mean scores above 3.5 or below 1.9. The MSc cohort means have a similar median score of 2.5 but the mean scores are much more evenly distributed between 4.5 and 1.4. The distributions of the mean scores pre- and post-intervention in the Control cohort reveal little change in the range of mean scores post-intervention (median = 2.6, and no mean scores above 3.7 or below 1.9 (Figures 4 and 5).

Figure 4: Box and whisker plots of mean scores for each of 100 images for the pre-intervention Control cohort and pre-intervention MSc cohort. Each circle represents the mean score for one image. The y-axis scale represents the mean score for the cohort for each image on a scale from 1 to 5. The median score is shown by the horizontal green line inside the box and the top and bottom of the box are the upper and lower quartiles respectively.

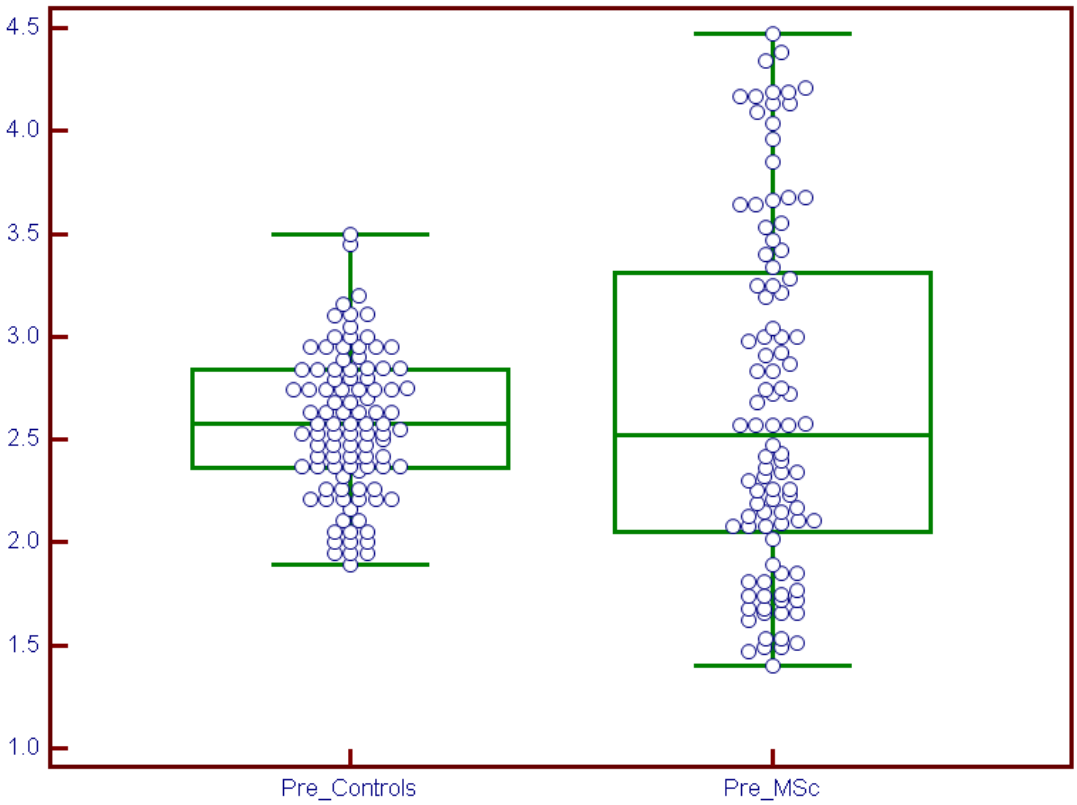
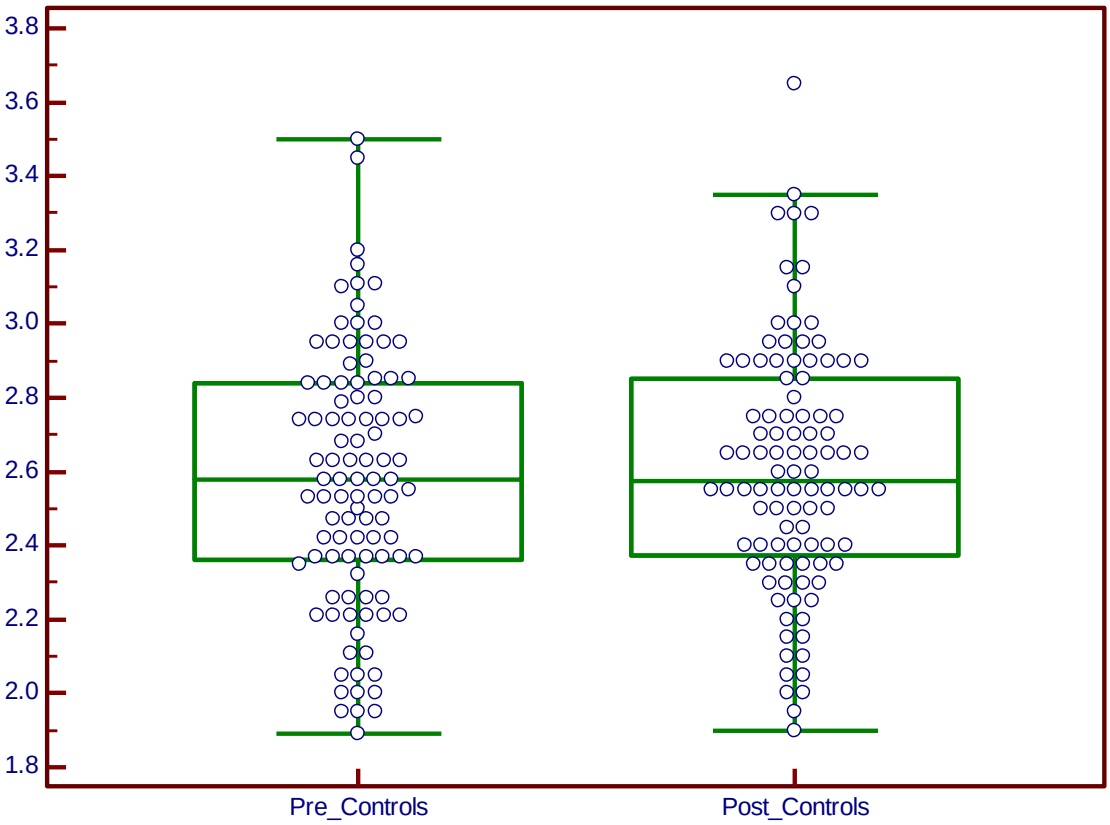

4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Pre_Controls
Pre_MSc

Figure 5: Box and whisker plots of mean scores for each of 100 images for the pre- and post-intervention Control cohort. Each circle represents the mean score for one image. The y-axis scale represents the mean score for the cohort for each image on a scale from 1 to 5. The median score is shown by the horizontal green line inside the box and the top and bottom of the box are the upper and lower quartiles respectively.



For the clinical decision-making exercise the mean scores for the MSc cohort increased from 5.5 pre-intervention to 5.9 post-intervention (Table 3). There was no statistically significant improvement in median score, which was 6 both pre- and post-intervention (P = 0.123; Wilcoxon Statistic = 575.5). For the Control group the mean score (5.5) did not change pre- and post-intervention and was identical to the baseline mean for the MSc cohort. There was no statistically significant difference in median score, which was 5 both pre- and post-intervention.

Comparing the MSc and Control cohorts there was no statistically significant difference between the medians of the two cohorts pre-intervention (p = 0.61, U = 572.0, Mann-Whitney test) or post-intervention (p = 0.09, U = 669.0).

Table 3: Performance in the four clinical decision making scenarios for the Control cohort (n=20) and the MSc Cohort (n = 53) pre- and post-intervention. Scores given are out of a maximum of 8.

Cohort	Pre		Post	
	Mean	Median	Mean	Median
Control	5.5	5	5.5	5
MSc	5.5	6	5.9	6

Discussion

This study demonstrated that the educational intervention was associated with an increased awareness of glaucomatous disc features, with a statistically significant increase in median scores for the MSc cohort. For the Control cohort there was a marginal increase in mean scores post-intervention but no statistically significant difference between medians. This result for the Control cohort is not surprising and is supportive evidence for the validity of the study design. Overall, these findings support the value of the educational intervention for the acquisition of knowledge. This was, however, a desktop-based exercise rather than one which reflects the application of knowledge to a clinical practice setting. In Miller's pyramid of clinical competence the 'features of the optic disc' exercise is firmly rooted in the 'knows' section, consisting of factual knowledge, which lies at the base of the pyramid.[6] Nevertheless, this method of evaluation demonstrated that, not surprisingly, qualified optometrists retain the ability to memorise and recall factual information. The didactic, taught lecture component of the glaucoma module was high (approximately 70%) and therefore the improvement in scores for the MSc cohort may reflect this.

Previous studies have used optic disc images to assess the ability of optometrists and ophthalmologists to detect glaucomatous damage.[11-16] The current study used the Discus software package, which presents a series of monoscopic disc images on a computer screen and uses a 5-point Likert scale to record the probability of damage.[17] It should be noted that the calculation of sensitivity and specificity for the Discus element of this study is a somewhat unorthodox use of these values, which are more commonly used to indicate the validity of a medical diagnostic test, rather than the outcome of an educational intervention.[18,19] However, a similar approach was used previously by the developers of the Discus program.[17] Comparing pre- and post–intervention data for the MSc cohort, there was a significant increase in mean sensitivity from 74% to post-intervention 81%. This was at the price of reduced specificity, which fell from 64% to 55%, a reduction that was also statistically significant. The intervention, although improving the correct identification of damaged discs, could therefore result in an increased number of false positive referrals if undamaged discs are being incorrectly identified as damaged. A similar analysis for the Control cohort revealed minimal differences in both mean sensitivity and mean specificity. Considering the pre-intervention results, there is evidence to suggest that there were differences between the two cohorts.. The pre-intervention mean

sensitivities were significantly higher in the MSc cohort (74%) compared with the Controls (59%), differences that were even greater post-intervention (81% versus 58%). Interestingly, the MSc cohort also had a higher mean specificity pre-intervention than the Controls (64% versus 60%) but this was reversed post-intervention with the MSc mean specificity falling to 55% compared with 61% for the Controls, with neither difference being statistically significant.

Based on their performance on the Discus program, it is debateable whether the MSc cohort benefitted from the intervention. Glaucoma is a disease with low prevalence, and it can be argued that the clinician would need to have a markedly increased sensitivity post-intervention if their specificity is to be reduced, as happened on average to the MSc cohort. However, it must be borne in mind that this was a difficult sample of disc images to interpret. The sample was highly selective and included a large proportion of discs from patients in the glaucoma clinic who were considered to be glaucoma suspects but had normal visual fields. It is therefore likely that the proportion of unequivocally healthy discs was under-represented compared to an unselected sample.[17] Nonetheless, the ROC analysis revealed an impressive composite performance by both cohorts when considered in isolation and also when compared with the results from the Discus Expert Panel.[17] There was no significant difference between the AUROCs for the two cohorts pre-intervention (MSc 0.85 and Control 0.84) and both AUROCs were close to that achieved by the experts (0.87). The AUROC of the MSc cohort was essentially unchanged post-intervention (0.84) with the improvement in sensitivity being offset by the reduction in specificity.

The repeatability of the MSc subjects' responses was moderate, with 42% of repeats showing a difference of at least one category, and 9 of the 1378 repeats post-intervention revealing a discrepancy of 4 categories. However, assessment of discs is a challenging clinical task. Interestingly, when repeatability was assessed in the same way as in this paper by the Discus Expert Panel, agreement was again moderate; "on average, discrepancies of one category were seen in 44% of [the] 26 repeated images".[17] This figure is similar to that obtained for the MSc cohort (42%). Repeatability was higher for the Control cohort, with around 30% of repeats showing a difference of at least one category.

There is evidence from the latency data to suggest that, post-intervention, the members of the MSc cohort may have been adopting a more critical approach to disc

interpretation as there was a statistically significant increase in mean latency post-intervention (11.4s) compared with pre-intervention (7.4s). Assuming that this extra time was spent analysing each image, it may reflect a more intense scrutiny of the images for more subtle indications of glaucoma. The equivalent data for the Discus Expert Panel were an average of 7 seconds, similar to the pre-intervention results for the MSc cohort.[17] The Control cohort took significantly longer on average to respond to the presented images pre-intervention (13.6s) compared with the MSc cohort, but the longer latencies of the MSc cohort post-intervention resulted in the difference between them and the Controls (13.1s) failing to reach statistical significance.

Although both the MSc and Control cohort have almost identical AUROCs pre-intervention they are very different in their approach to grading the Discus images. The MSc pre-intervention subjects were much more prepared to use the full range of the 5-point scale, while the Controls were much more reluctant to use the 'definitely normal' and 'definitely abnormal' grades. Yet the ROC curves indicate that both cohorts graded the images with equal facility overall. This different approach to grading is further supported by the relatively moderate repeatability of the MSc subjects' responses, compared with the higher repeatability of the control cohort. Though the Control cohort were less confident in their grading abilities than the MSc cohort they were equally good at grading the images.

In addition to assessing the impact of the intervention on disc assessment, the study also evaluated its impact on the ability to make clinical decisions on 'virtual' glaucoma patients using case scenarios. The four scenarios covered a range of possible diagnoses and management options, featuring cases in which the "patients" were of differing ethnic origin. Discs and fields ranged from the probably normal to the almost certainly damaged and featured asymmetries between right and left eyes. Although the mean scores on this assessment increased for the MSc cohort post-intervention, there was no significant difference in median scores. For the Control cohort there was, as could be expected, no change in mean scores pre- and post-intervention and no significant difference in median scores. There were no significant differences between the MSc and Control cohorts' performance on this exercise either pre- or post-intervention. It is clear that any improvement in the MSc group at this task was marginal, and their overall performance was little better than that of the Control cohort.

A clinical scenario-based approach in the assessment of these decision-making skills is regularly used in the core training of optometrists and in continuing professional education for registered optometrists. According to Miller's pyramid, this task belongs in the "knows how" region, one level up from the "knows" region in which the disc features exercise resides. The "knows how" level describes the ability of the clinician to use their knowledge in a particular context. Based on the current study, the results of the "knows how" exercise were rather disappointing, suggesting that the intervention did not significantly improve the students' performance at these tasks. These results suggest that the Glaucoma module may have had too little focus on developing the "knows how" skills of participants.

In common with many other postgraduate Masters modules offered by UK optometry departments, the City University London 'Optometric Management of Glaucoma' module that was used in this study had a high proportion of didactic lecture content. The finding that the module did not appear to improve clinical decision-making is consistent with the findings of previous studies with respect to the impact of didactic interventions for continuing medical education.[20,21] Systematic reviews of educational interventions in Primary Care found that combinations of interventions were more effective than single interventions, particularly if the educational activity was related to the clinicians' actual practice).[22] Significantly, previous reports of successful interventions to improve optometrist case-finding, refine referrals or co-management diagnosed glaucoma utilised multi-component training programmes, with lectures being augmented by, for example, training sessions in the glaucoma clinic, case-based discussions and/or targeted feedback on referrals by ophthalmologists.[23-25]

The study had several limitations that should be borne in mind when considering the generalisability of the results. The scenario-based clinical decision making assessment used had not been previously validated and therefore its sensitivity to detect meaningful changes in decision-making ability has not been determined. Furthermore, it can be argued that neither cohort comprised a representative sample of UK optometrists. For the MSc cohort, all subjects were attending the module through choice and were likely to have a particular interest in glaucoma. The Control cohort may also not be representative, since these subjects were prepared to volunteer for the study, and may therefore be more confident of their glaucoma diagnostic skills than the average UK community optometrist. Consequently, it is possible that the performance of both groups overestimates that of UK optometry as a whole. Equally, the "high

baseline" ability of the MSc cohort may have masked the overall impact of the intervention.

With respect to the diagnosis of glaucomatous discs, the performance of both cohorts on the Discus program could have been affected by a number of confounders. The images were monoscopic and did not allow the appreciation of the optic disc in three dimensions. Furthermore, only one disc was shown per 'patient', which prevented the grader from identifying disc asymmetry. It has been suggested that the lower specificity of community optometrists when assessing discs for glaucoma compared to ophthalmologists could be a consequence of the perceived ramifications of misdiagnosing glaucoma, compared to making a false positive referral. [16] In the current study, this over-cautious approach may have been further confounded by the perception that both cohorts were being examined, despite the reassurance that their data was being collected anonymously.

Based on the discriminatory power of the assessments used in the current study, a predominantly didactic educational intervention did not improve clinical decision-making using a scenario-based assessment nor improve performance in disc assessment as determined by the Discus computer program. Nonetheless, UK optometrists performed creditably on this task in comparison with an expert panel.

The results suggest that the use of a traditional didactic approach in isolation may not be suited to training optometrists to achieve or develop the clinical competencies required for glaucoma detection and management. Consideration should be given to the development of specialist postgraduate training that is more practice-based, provides opportunities for active learning and includes strategies for feedback and reinforcement.

References

1: Mowatt G, Burr JM, Cook JA et al. Screening Tests for Detecting Open-Angle Glaucoma: Systematic Review and Meta-analysis. *Invest Ophthalmol Vis Sci* 2008; 49: 5373-85

2: Bowling B, Chen SD & Salmon JF. Outcomes of referrals by community optometrists to a hospital glaucoma service. *Br J Ophthalmol* 2005; 89: 1102-4.

3: General Optical Council 2011 optometry stage 2 core competencies http://www.optical.org/en/Standards/Standards_in_competence.cfm (accessed 9.10.13)

4: Myint J, Edgar DF, Kotecha A, Crabb DP & Lawrenson JG. Development of a competency framework for optometrists with a specialist interest in glaucoma. *Eye* 2010; 24: 1509-14

5: Harper RA, Lawrenson JG, Vernon SA & Spry PGD Post-graduate specialist glaucoma training and accreditation in optometry. *Optom Pract* 2013 (in Press)

6: Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; 65: S63-7.

7: Wass V, Van der Vleuten C, Shatzer J & Jones R. Assessment of clinical competence. *Lancet* 2001; 357: 945-9.

8: Shah R, Edgar DF & Evans BJ. A comparison of standardised patients, record abstraction and clinical vignettes for the purpose of measuring clinical practice. *Ophthalmic Physiol Opt* 2008; 30: 209-24.

9: Keltner JL, Johnson CA, Anderson DR et al. The association between glaucomatous visual fields and optic nerve head features in the Ocular Hypertension Treatment Study. *Ophthalmology* 2006; 113: 1603–12.

10: Miglior S, Pfeiffer N, Torri V, Zeyen T, Cunha-Vaz J & Adamsons I. Predictive factors for open-angle glaucoma among patients with ocular hypertension in the European Glaucoma Prevention Study. *Ophthalmology* 2007; 114: 3–9.

11: 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.

12: Harper R, Radi N, Reeves BC, Fenerty C, Spencer AF & Batterbury M. Agreement between ophthalmologists and optometrists in optic disc assessment: training implications for glaucoma co-management. *Graefes Arch Clin Exp Ophthalmol* 2001; 239: 342-50.

13: Reus NJ, Lemij HG, Garway-Heath DF et al. Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European Optic Disc Assessment Trial. *Ophthalmology* 2010; 117: 717-23.

14: 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.

15: Kong YXG, Coote MA, O'Neill EC et al; Optic Nerve Study Group. Optic disc evaluation in optic neuropathies: the optic disc assessment project. *Ophthalmology* 2011; 118: 964-70.

16: Hadwin SE, Redmond T, Garway-Heath DF et al. Assessment of optic disc photographs for glaucoma by UK optometrists: the Moorfields Optic Disc Assessment Study (MODAS). *Ophthalmic Physiol Opt* 2013; 33: 618-24.

17: Denniss J, Echendu D, Henson DB et al Artes PH. Discus: investigating subjective judgment of optic disc damage. *Optom Vis Sci* 2011; 88: E93-101.

18: Altman, D. G. & Bland, J. M. Diagnostic tests 1: sensitivity and specificity. *BMJ* 1994; 308: 1552.

19: Harper R & Reeves B. The sensitivity and specificity of direct ophthalmoscopic optic disc assessment in screening for glaucoma: a multivariate analysis. *Graefes Arch Clin Exp Ophthalmol* Graefes 2000; 238: 949-55.

20: Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P & Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behaviour or health care outcomes? *JAMA* 1999; 282: 867-74.

21: Bloom BS. Effects of continuing medical education on improving physician clinical care and patient health: a review of systematic reviews. *Int J Technol Assess Health Care* 2005; 21: 380-5.

22: Patel UD, Murdoch IE & Theodossiades J. Glaucoma detection in the community: does ongoing training of optometrists have a lasting effect? *Eye* 2006; 20: 591-4.

23: Devarajan N, Williams GS, Hopes M, O'Sullivan D & Jones D. The Carmarthenshire Glaucoma Referral Refinement Scheme, a safe and efficient screening service. *Eye* 2011; 25: 43-9.

24: Azuara-Blanco A, Burr J, Thomas R, Maclennan G & McPherson S. The accuracy of accredited glaucoma optometrists in the diagnosis and treatment recommendation for glaucoma. *Br J Ophthalmol* 2007; 91:1639-43.

25. Spencer IC, Spry PG, Gray SF, Baker IA, Menage MJ, Easty DL, Sparrow JM. The Bristol Shared Care Glaucoma Study: study design. *Ophthalmic Physiol Opt* 1995; 15: 391-4.