ASSESSMENT OF OPTIC DISC PHOTOGRAPHS FOR GLAUCOMA BY UK OPTOMETRISTS: THE MOORFIELDS OPTIC DISC ASSESSMENT STUDY (MODAS)

Shona E Hadwin,[1A] Tony Redmond,[1B] David F Garway-Heath,[1B] Hans G Lemij,[2] Nicolaas J Reus,[2] Gavin Ward,[1B] Roger S Anderson[1AB,3]

[1]Moorfields Eye Hospital, London, United Kingdom; [A]Optometry Department; [B]National Institute for Health Research (NIHR) Biomedical Research Centre;
[2]Glaucoma Service, Rotterdam Eye Hospital (Oogziekenhuis Rotterdam), Rotterdam, The Netherlands;
[3]School of Biomedical Sciences, University of Ulster, Coleraine, United Kingdom.

RUNNING TITLE: Performance of UK optometrists in optic disc assessment

TABLES: 0
FIGURES: 5

CORRESPONDING AUTHOR:
Shona E Hadwin,
Optometry Dept,
Moorfields Eye Hospital,
162 City Road,
London,
EC1V 2PD,
United Kingdom
Email: shonahadwin@hotmail.com
Tel: + 44 20 7566 2101

ABSTRACT

Purpose: To assess the ability of UK optometrists to accurately discriminate between stereoscopic photographs of healthy and glaucomatous optic discs.

Methods: An online survey, including questions relating to qualification, practice environment, and diagnostic methods was completed by 1256 optometrists. Based on their responses, 208 (17%) were selected to undertake an online disc assessment exercise. Optometrists evaluated the same disc images previously assessed by European ophthalmologists as part of the European Optic Disc Assessment Trial (EODAT); the task was to state if the disc appeared healthy or glaucomatous. There were 110 stereoscopic disc images, of which 40 were healthy, 48 glaucomatous, and 6 ocular hypertensive, with 16 duplicates images. Sensitivity, specificity and overall accuracy were calculated and compared between optometrist groups and with the EODAT ophthalmologists using permutation analysis.

Results: Median sensitivity was 0.92 (95% CI: 0.70, 1.00) and median specificity was 0.74 (95% CI: 0.62, 0.88). Median overall accuracy was 80% (95% CI: 67%, 88%). Agreement between optometrists was moderate (Fleiss' κ: 0.57). Optometrists with higher qualifications did not have overall higher sensitivity than those without ($p = 0.23$), but had higher specificity ($p = 0.001$) and higher overall accuracy ($p < 0.001$). Optometrists displayed higher sensitivity but lower specificity than the EODAT ophthalmologists.

Conclusion: UK optometrists displayed a high sensitivity and moderate specificity when assessing optic discs for the presence of glaucoma, in the context of this study.

INTRODUCTION

Subjective assessment of the optic disc is one of the most important examinations when investigating a patient for glaucoma. Several studies have reported the agreement within and between practitioners in optic disc assessment[1-9]. Many originate out of a desire to assess the performance of a particular cohort of practitioners within a particular practice setting, often to evaluate a training scheme[10]. Typically, the cohorts tend to be relatively small.

A recent study by Reus *et al*[7], the European Optic Disc Assessment Trial (EODAT), reported on the performance of ophthalmologists across Europe in classifying discs in stereoscopic photographs as either normal or glaucomatous, and found notable differences between professionals and moderate diagnostic accuracy when compared with imaging devices. In the UK, more than 95% of glaucoma cases referred to the hospital eye service originate in primary care optometry practice. While several previous studies have assessed the agreement, or otherwise, of optometrists' referrals for glaucoma with the ophthalmologist's opinion, there is a shortage of data on the collective performance of a large sample of UK optometrists when assessing the disc photographs of a previously well-characterised cohort of glaucoma patients and healthy subjects. In addition, the influence of various levels of experience and qualification, and different modes of practice, on optometrists' ability to classify optic discs has received little attention. This information is especially important given that, over the last 20 years, the role of optometrists in the UK has expanded, particularly in the management of stable glaucoma. A considerable number of optometrists now assess and manage patients alongside ophthalmologists in hospital-based glaucoma clinics throughout the UK. In addition, the number of successful optometry-based shared-care glaucoma schemes[11, 12] and glaucoma referral refinement pathways[13, 14] is increasing throughout the country. Optometrists are increasingly availing of higher qualifications such as the College of Optometrists' Diploma in Glaucoma, the Independent Prescribing qualification and various MSc modules in glaucoma that involve advanced training in basic theory, investigative techniques and management of patients with glaucoma.

A study of the performance of optometrists, specifically in optic disc assessment, incorporating large numbers of practitioners, especially those from high-street primary care practices, is long overdue. Particularly important is performance against the confirmed status of a large number of discs from well-characterised glaucoma patients at different disease stages, rather than agreement, or otherwise, with an individual ophthalmologist.

MATERIALS AND METHODS

Participants

Participants were optometrists registered in the UK. No restrictions were placed on mode of practice, region of practice, number of additional qualifications, refractive error or binocular status. The study was divided into two stages.

Stage One: Online Survey

An online survey (Appendix 1) allowed preliminary information to be gathered about the way individual UK optometrists currently practise, to facilitate selection of representative participants (see below) and analysis by various categories in stage 2.

The survey was advertised to optometrists via email through the College of Optometrists, whose members represent approximately 95% of UK optometrists. A group was produced on the social networking site, Facebook®, advertisements were posted in optometry-related magazines and short presentations were given to local optometric committees throughout the UK. On completion of the survey, each optometrist was invited to indicate if they were willing to take part in stage 2 (optic disc assessment) of the study and advised that if they were chosen, the investigators would require their contact details. Stratified sampling was conducted for stage 2, to select optometrists from a wide range of practice environments. This involved determining the proportion of optometrists who worked for most of their working week in particular practice environments and sampling the same proportions for a sample size that approximated that of the EODAT study. The proportions in each group reflected the proportions in those environments throughout the UK.

Stage 2: Optic Disc Assessment

Optometrists invited to take part in stage 2 of the study received a pair of plano prism spectacles, with 6Δ base IN each eye, for viewing the stereo images, along with a username and a unique activation code. The optic disc assessment test was accessed from the same website as the survey. The stereo-photograph set was identical to that used in the EODAT study [7] and a study comparing the performance of imaging devices and clinical assessment by ophthalmologists [6]. Three 'calibration images' were presented initially and could be accessed at any time during the assessment. These images contained healthy optic discs: 1 small (5th percentile), 1 medium (50th percentile) and 1 large (95th percentile). These were followed by 110 randomised stereoscopic disc images; 40 were healthy, 48 glaucomatous and 6 from ocular hypertension patients, with 16 duplicates (proportions not revealed to participants). Discs were classified for the two previous studies[6,7] and the current study by 1 of 4 glaucoma ophthalmologists at Rotterdam Eye Hospital. Glaucomatous discs were required to have characteristic glaucomatous changes (e.g. notching, thinning of the neuroretinal rim, possible haemorrhage) and a corresponding visual field defect with standard automated perimetry. Patients had established glaucoma clinically and were being followed regularly and treated for the condition at Rotterdam Eye Hospital. Healthy discs were classified on the basis of a normal optic disc appearance, the absence of a visual field defect, intraocular pressure <21mmHg and a negative family history of glaucoma.

Optometrists viewed the images and registered their classification by clicking one of two buttons: 'glaucoma' or 'healthy'. They were also given the opportunity to return to previous disc images and change their classification, before submission of all responses. Optometrists who did not have binocular single vision (n = 13) were permitted to undertake the task without the spectacles. On completion of the assessment, a 'percentage correct' score was presented on the screen.

All answers were merged with the participant's survey responses by their unique activation code.

Statistical Analysis

Sensitivity, specificity and overall accuracy (number correctly identified, divided by the total number) were calculated for each optometrist.

The significance of differences in performance between groups was determined by permutation analysis, unless otherwise stated. Firstly, an observed statistic (δ_{obs}) was calculated as the difference between the mean of Group 1 and Group 2 (μ_1 - μ_2). Assuming the null hypothesis, that there is no statistically significant difference between the means of these groups, δ_{obs} would be expected to fall within the 95% confidence region of a distribution of values of δ when optometrists were randomly assigned to each group multiple times. If δ_{obs} were to fall outside this region, the difference between groups would be considered significant at the 95% confidence level. Optometrists were randomly assigned to each group 5,000 times and a distribution of δ values was plotted (δ_P). A p-value was calculated for δ_{obs} based on its position in the permutation distribution.

Optometrists working in different practice environments were assigned to groups according to whether they undertake 'any' or 'no' work within that environment, regardless of their main mode of practice.

Optometrists working in a specialist glaucoma clinic setting were asked for the number of years (<2, 2-5 or >5 years) and hours per week they undertook this work. The number of hours per year was multiplied by 1, 3.5 or 6, according to the number of years they indicated as having worked in this setting. A value of 6 was chosen for the '>5 years' category to approximate the error on the abscissa for the other categories while remaining conservative.

Each optometrist was asked to indicate, in stage 1, their confidence in optic disc assessment, on a scale from 1-7 (1: not confident at all; 7: completely confident). Performance in stage 2 was later compared between optometrists reporting different levels of confidence, using a Kruskal-Wallis test.

Statistical analysis was carried out using the freely-available open-source statistical environment, R[15] and associated packages, sp[16] and maptools[17].

We certify that all applicable institutional regulations concerning the ethical use of human volunteers were followed during this research. The protocol for this study was approved by the Moorfields Eye Hospital Research Ethics Committee. The research was conducted according to the tenets of the Declaration of Helsinki.

RESULTS

Stage 1 was completed by 1256 optometrists of working age from all regions of the UK (Figure 1(A)). At the time of the invitation, the total number of optometrists registered with the General Optical Council (GOC) in the UK was 12,761. The respondents thus represented 9.9% of GOC-registered optometrists at that time. Of this sample, 208 (17% of those surveyed, Figure 1(B)) took part in stage two. Ninety-six percent of optometrists participating in Stage 2 reported using some form of binocular ophthalmoscopy each week. The sensitivity and specificity of each optometrist are presented as single data points in Figure 2(A) and compared with European ophthalmologists in the EODAT study[7] (N = 243) in Figure 2(B). Optometrists were significantly more sensitive than ophthalmologists ($p < 0.001$) but significantly less specific ($p < 0.001$). Median sensitivity was 0.92 (95% CI: 0.70, 1.00), median specificity was 0.74 (95% CI: 0.62, 0.88) and median overall accuracy was 80% (95% CI: 67%, 88%). Marginal histograms show that accuracy and specificity values were normally distributed, but sensitivity values were not. A receiver-operator characteristic curve plotted through the mean sensitivity and specificity for optometrists and ophthalmologists appears largely symmetrical. To confirm this, the perpendicular distance from the mean performance of each group, and a diagonal line visualised from maximum performance (perpendicular) to the chance line was calculated. The distance from the mean ophthalmologists performance to this line was 0.1 and that from the mean optometrists performance to the line was 0.09. Inter-observer agreement was moderate (Fleiss' $\kappa = 0.57$). Agreement between optometrists with any hospital experience was slightly greater ($\kappa = 0.60$) than between those without ($\kappa = 0.56$). Similarly, agreement between optometrists with additional qualifications was greater ($\kappa = 0.63$)

than between those without ($\kappa = 0.55$) and agreement between optometrists working in a specialist glaucoma clinic was greater ($\kappa = 0.62$) than between those who did not ($\kappa = 0.56$). Intra-observer agreement was good (median Cohen's $\kappa = 0.71$; range: 0.08 – 1; interquartile range (IQR): 0.59 – 0.86).

Figure 3 shows the difference in mean performance between optometrists who undertake any ($n = 53$) or no ($n = 155$) work in a hospital setting, using permutation analysis. Optometrists working in a hospital have a higher specificity ($p < 0.001$) and overall accuracy ($p < 0.001$) when compared with optometrists who do not. Sensitivity was not significantly different ($p = 0.48$). Time spent in a specialised glaucoma clinic ($n = 35$) had no significant effect on sensitivity ($r^2 = 0.01$; $p = 0.76$) but a small, significant, effect on specificity ($r^2 = 0.22$; $p = 0.005$) and overall accuracy ($r^2 = 0.21$, $p = 0.005$) (Figure 4(A-C)). There was no significant association between sensitivity ($r^2 = 0.01$, $p = 0.88$), or overall accuracy ($r^2 = 0.01$, $p = 0.80$) and the number of years since professional qualification (seniority), however the association was slight, but significant for specificity ($r^2 = 0.03$, $p = 0.01$) (Figure 4(D-F)).

Fifty-three optometrists possessed additional qualifications, including successful completion of the College of Optometrists' Diploma in Glaucoma ($n = 7$), independent prescribing qualification ($n = 22$) and successful completion of an MSc glaucoma module ($n = 36$). Thirteen optometrists had more than one of these qualifications. Optometrists with additional qualifications, compared to those without, had similar sensitivity ($p = 0.23$), but higher specificity ($p = 0.001$) and accuracy ($p < 0.001$). Results were similar when comparing optometrists with and without an independent prescribing qualification alone (sensitivity $p = 0.25$; specificity $p < 0.001$; accuracy $p < 0.001$).

Figure 5(A) shows confidence levels reported by optometrists who took part in each stage of the study. Figure 5(B-D) shows sensitivity, specificity and overall accuracy for optometrists who undertook stage 2, as a function of their previously reported confidence level. No significant difference in any performance characteristic was found between confidence levels (Kruskal-Wallis; sensitivity: $p = 0.09$; specificity: $p = 0.53$; accuracy: $p = 0.21$).

DISCUSSION

On average, UK optometrists display high sensitivity and moderate specificity when examining optic discs for glaucoma in this study. Those who undertook stage 2 of the study are likely representative of the larger sample that took part in stage 1. Figure 5(A) shows that, overall, the confidence of those optometrists was slightly lower than the average confidence of the entire cohort enrolled in stage 1, thereby avoiding, as far as possible, bias towards optometrists who felt overly confident in their ability to correctly grade an optic disc. Those optometrists with additional qualifications had, overall, more comparable confidence to that of the entire cohort enrolled in stage 1.

The higher sensitivity among optometrists and specificity among ophthalmologists likely reflects a criterion difference, rather than a difference in ability to discriminate glaucomatous discs from healthy discs. This is reflected in the similar overall accuracy between groups and the largely symmetrical receiver-operator curve drawn through the mean performance characteristics for each group in Figure 2(B). This result may not be entirely surprising when one considers the priorities of the optometrist in practice and the ophthalmologist within a hospital setting. It is also worth considering the perceived implications by either group of a 'false alarm' and 'miss' when assessing discs for glaucoma. For some optometrists, a false alarm (i.e. being over-cautions and making a false referral) may be perceived as having fewer ramifications than missing glaucoma. Conversely, for some ophthalmologists, a false alarm may lead to an inappropriate commencement of treatment, therefore it may be perceived as preferable to exercise restraint in the short-term when deciding on the presence or absence of glaucoma. The nature of the current study may introduce some bias in the results. While no indication was given beforehand about the likely proportion of glaucomatous discs in the set, optometrists are likely to suspect that glaucomatous discs represent a substantially greater proportion of the disc set than the 2% of discs in their practice. This, together with the perception that they are being examined may also have caused them to be over-cautious in their assessments. The agreement among all optometrists was greater than that among all ophthalmologists in the EODAT study (Fleiss' $\kappa = 0.54$), reflected by the reduced spread in the data for optometrists, compared to that of ophthalmologists in Figure 2(B).

The performance of optometrists with experience working in a hospital, some of which worked in glaucoma clinics, was compared to that of community optometrists. However, classifying individuals as 'hospital' or 'independent' optometrists is difficult because optometrists spend different proportions of time in various settings each week. That specificity and overall accuracy of optometrists with *any* hospital experience was significantly higher than that of optometrists without hospital experience may be a consequence of greater opportunities to compare discs that optometrists typically see when working in community practice with the discs that they observe in the hospital setting, where there is a much greater number of patients with glaucomatous discs and where visual field data are available for all discs examined. This experience may also explain the criterion shift of these individuals towards that of ophthalmologists. This finding is also supported by the fact that 24 of the 53 optometrists with hospital experience had additional qualifications. The results of this study support those of previous reports that have documented the effectiveness of professional training on the performance of optometrists at disc assessment[18, 19]. Improvement was greatest in specificity and overall accuracy, which would result in fewer false positive referrals. It is also worthy of note that many hospital-based optometrists participate in disease screening as part of clinical trials and epidemiological studies. Training in this regard may improve their ability to correctly classify disc images as glaucomatous or healthy.

It was expected that participants' level of experience and confidence in disc assessment would influence their decision-making. However, despite participants having a wide range of reported confidence levels, there was no effect on performance. Nevertheless, it is interesting to note the wide range of performance at each level, particularly for those who indicated a confidence level of 4 or 5. Seniority had little influence on performance, but the degree of variance in the data shown in Figure 4 (D, E) is noteworthy. Interestingly, overall accuracy was consistent for all optometrists, therefore the variance is largely accounted for by the diagnostic criterion.

In conclusion, the current study provides important information about the performance of UK optometrists in their ability to classify optic discs and provides evidence for the effectiveness of additional qualifications and experience in hospital glaucoma clinics in enhancing performance.

ACKNOWLEDGEMENTS

Funding for the study was provided by the College of Optometrist's iPRO small grants scheme for practice-based research (SEH).

The authors acknowledge a proportion of their financial support from the Department of Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology for a Biomedical Research Centre for Ophthalmology. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health.

The authors thank all optometrists who participated in each stage of this study, and J Hadwin for proofreading the paper.

This study was presented in part as a poster at the Association for Research in Vision & Ophthalmology (ARVO) annual meeting, Fort Lauderdale, USA in 2012.

CONFLICTS OF INTEREST

None

REFERENCES

1. Marks JR, Harding AK, Harper RA, Williams E, Haque S, Spencer AF, et al. Agreement between specially trained and accredited optometrists and glaucoma specialist consultant ophthalmologists in their management of glaucoma patients. Eye (Lond). 2012;26(6):853-61. Epub 2012/04/14.
2. Varma R, Steinmann WC, Scott IU. Expert agreement in evaluating the optic disc for glaucoma. Ophthalmology. 1992;99(2):215-21. Epub 1992/02/01.
3. 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(10):1662-7. Epub 1994/10/01.
4. 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(5):342-50. Epub 2001/08/03.
5. Harper R, Reeves B, Smith G. Observer variability in optic disc assessment: implications for glaucoma shared care. Ophthalmic Physiol Opt. 2000;20(4):265-73. Epub 2000/08/30.
6. Reus NJ, de Graaf M, Lemij HG. Accuracy of GDx VCC, HRT I, and clinical assessment of stereoscopic optic nerve head photographs for diagnosing glaucoma. Br J Ophthalmol. 2007;91(3):313-8. Epub 2006/10/13.
7. Reus NJ, Lemij HG, Garway-Heath DF, Airaksinen PJ, Anton A, Bron AM, et al. Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European Optic Disc Assessment Trial. Ophthalmology. 2010;117(4):717-23. Epub 2010/01/05.
8. Theodossiades J, Murdoch I. What optic disc parameters are most accurately assessed using the direct ophthalmoscope? Eye (Lond). 2001;15(Pt 3):283-7. Epub 2001/07/14.
9. 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(4):742-6. Epub 2010/11/09.
10. Denniss J, Echendu D, Henson DB, Artes PH. Discus: investigating subjective judgment of optic disc damage. Optom Vis Sci. 2011;88(1):E93-101. Epub 2010/11/03.

11. Gray SF, Spry PG, Brookes ST, Peters TJ, Spencer IC, Baker IA, et al. The Bristol shared care glaucoma study: outcome at follow up at 2 years. Br J Ophthalmol. 2000;84(5):456-63. Epub 2000/04/27.
12. Mandalos A, Bourne R, French K, Newsom W, Chang L. Shared care of patients with ocular hypertension in the Community and Hospital Allied Network Glaucoma Evaluation Scheme (CHANGES). Eye (Lond). 2012. Epub 2012/01/10.
13. Bourne RR, French KA, Chang L, Borman AD, Hingorani M, Newsom WD. Can a community optometrist-based referral refinement scheme reduce false-positive glaucoma hospital referrals without compromising quality of care? The community and hospital allied network glaucoma evaluation scheme (CHANGES). Eye (Lond). 2010;24(5):881-7. Epub 2009/08/04.
14. Devarajan N, Williams GS, Hopes M, O'Sullivan D, Jones D. The Carmarthenshire Glaucoma Referral Refinement Scheme, a safe and efficient screening service. Eye (Lond). 2011;25(1):43-9. Epub 2010/10/23.
15. R Development Core Team. R. A Language and Environment for Statistical Computing2004.
16. Pebesma EJ, Bivand RS. Classes and methods for spatial data in R. R News [Internet]. 2005; 5(2):[http://cran.r-project.org/doc/Rnews/ pp.]. Available from: http://cran.r-project.org/doc/Rnews/.
17. Lewin-Koh NJ, Bivand R. maptools: Tools for reading and handling spatial objects. R package 2012;version 0.8-14:http://CRAN.R-project.org/package=maptools.
18. Theodossiades J, Murdoch I, Cousens S. Glaucoma case finding: a cluster-randomised intervention trial. Eye (Lond). 2004;18(5):483-90. Epub 2004/05/08.
19. Patel UD, Murdoch IE, Theodossiades J. Glaucoma detection in the community: does ongoing training of optometrists have a lasting effect? Eye (Lond). 2006;20(5):591-4. Epub 2005/07/16.

FIGURE LEGENDS

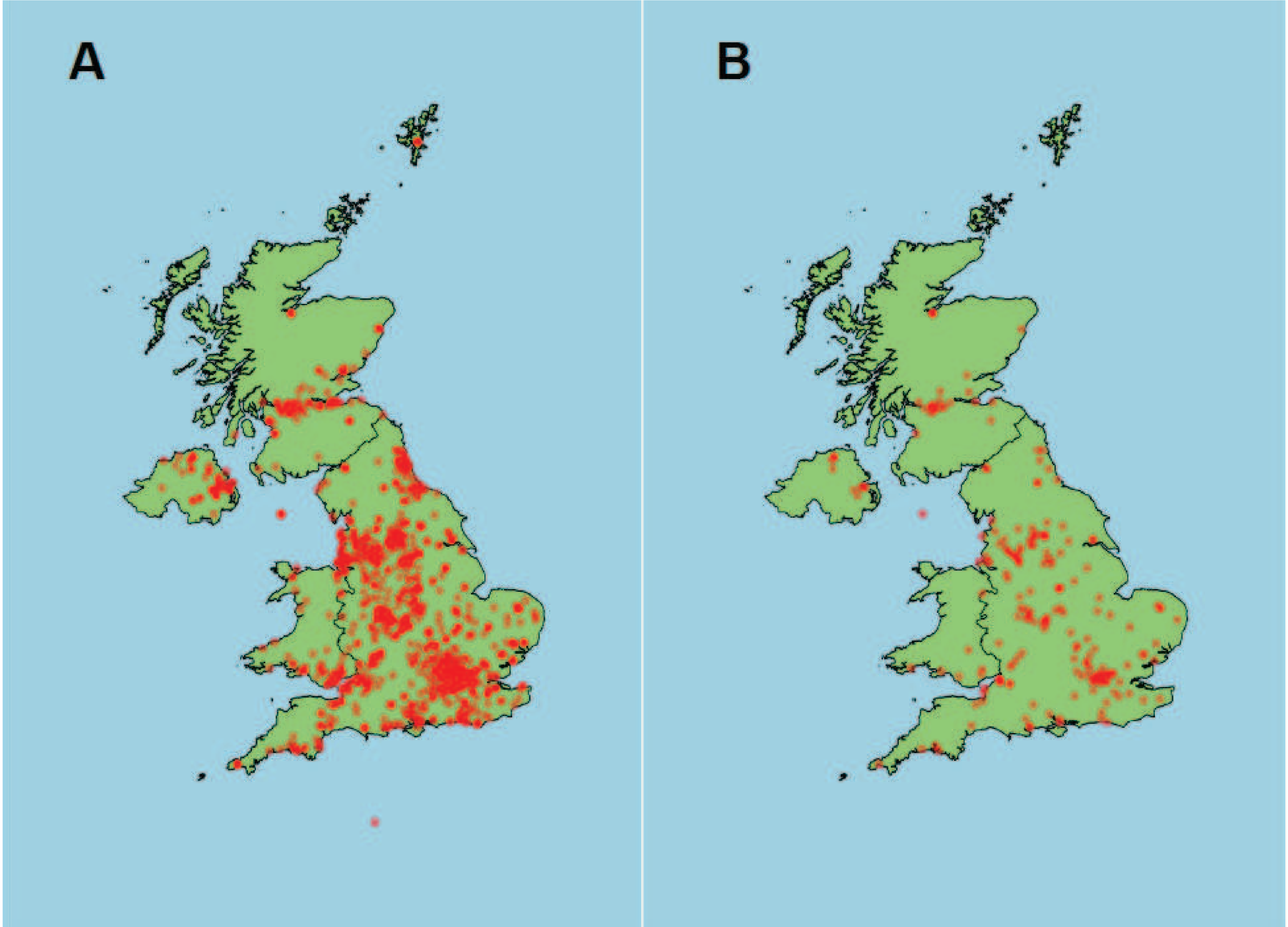
<u>Figure 1</u>: Geographic distribution of optometrists who completed stage 1 (A) and stage 2 (B).

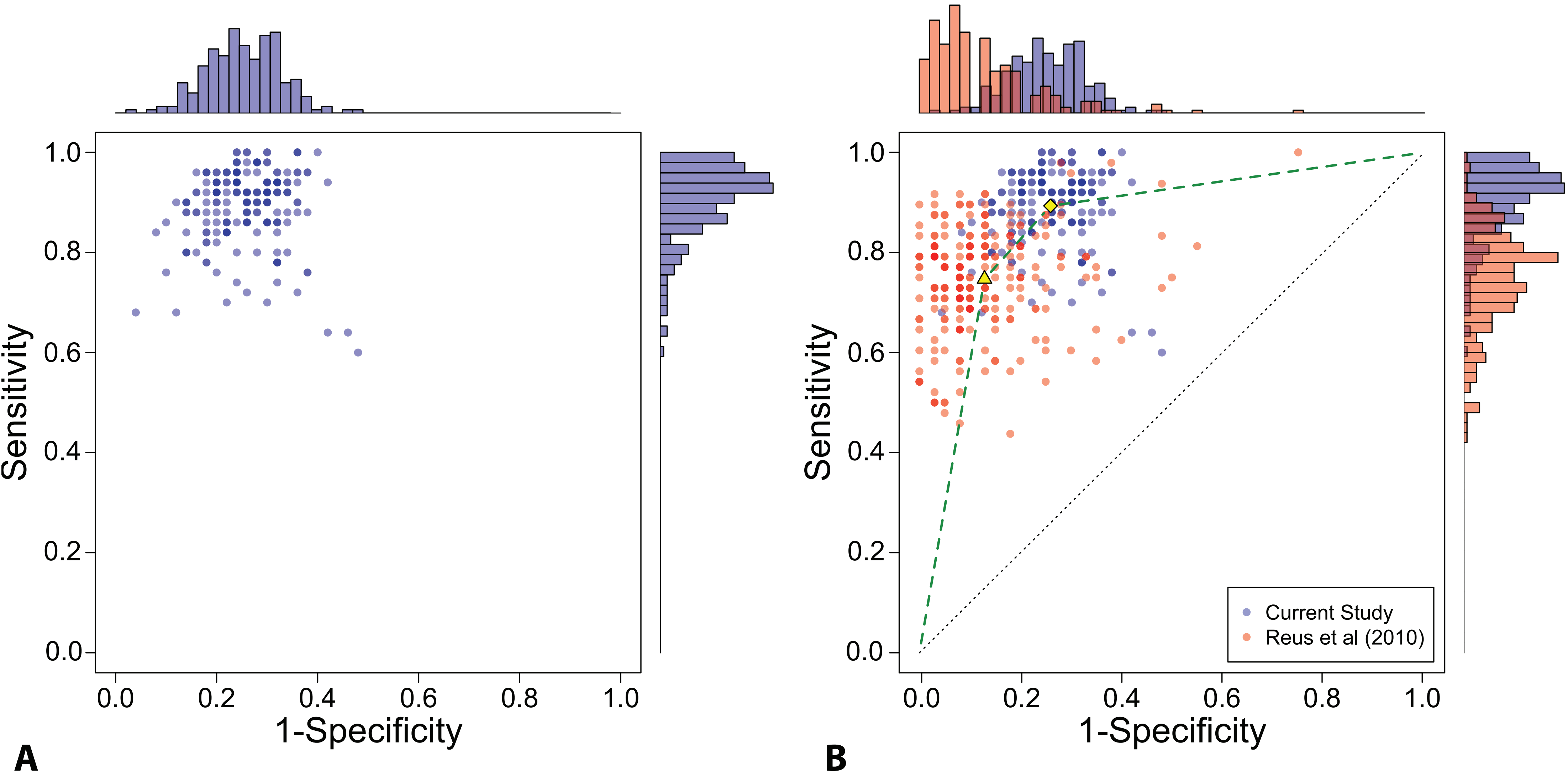
<u>Figure 2</u>: (A) Sensitivity and specificity plots showing the performance characteristics of UK optometrists in the current study. (B) The same data, plotted together with those of the EODAT study. The yellow triangle and diamond represent the mean performance of ophthalmologists and optometrists respectively.

<u>Figure 3</u>: Distributions of permutations of δ in the analysis of the effect of hospital experience on performance. The graphs show the median of the distribution (blue dotted line), the 95% confidence level (orange line) and the test statistic for the observed difference (δ_{obs}; red square).

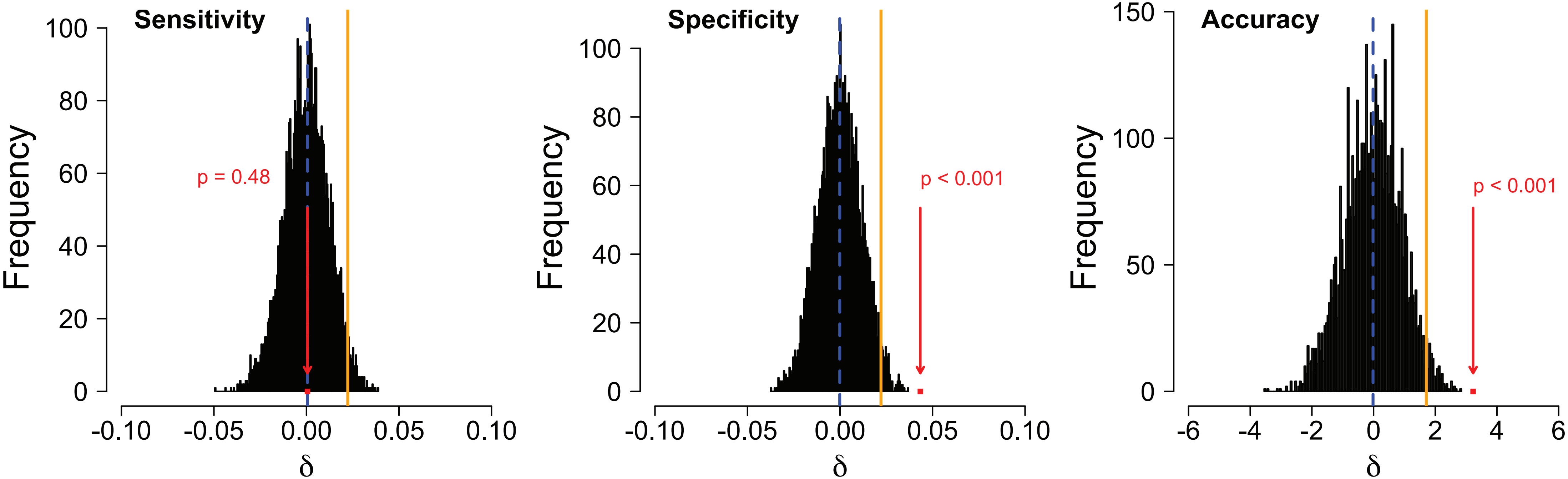
<u>Figure 4</u>: (A-C) The association between time spent in specialist glaucoma clinics and performance. (D-F) The association between the time since initial professional qualification and performance (lower panels).

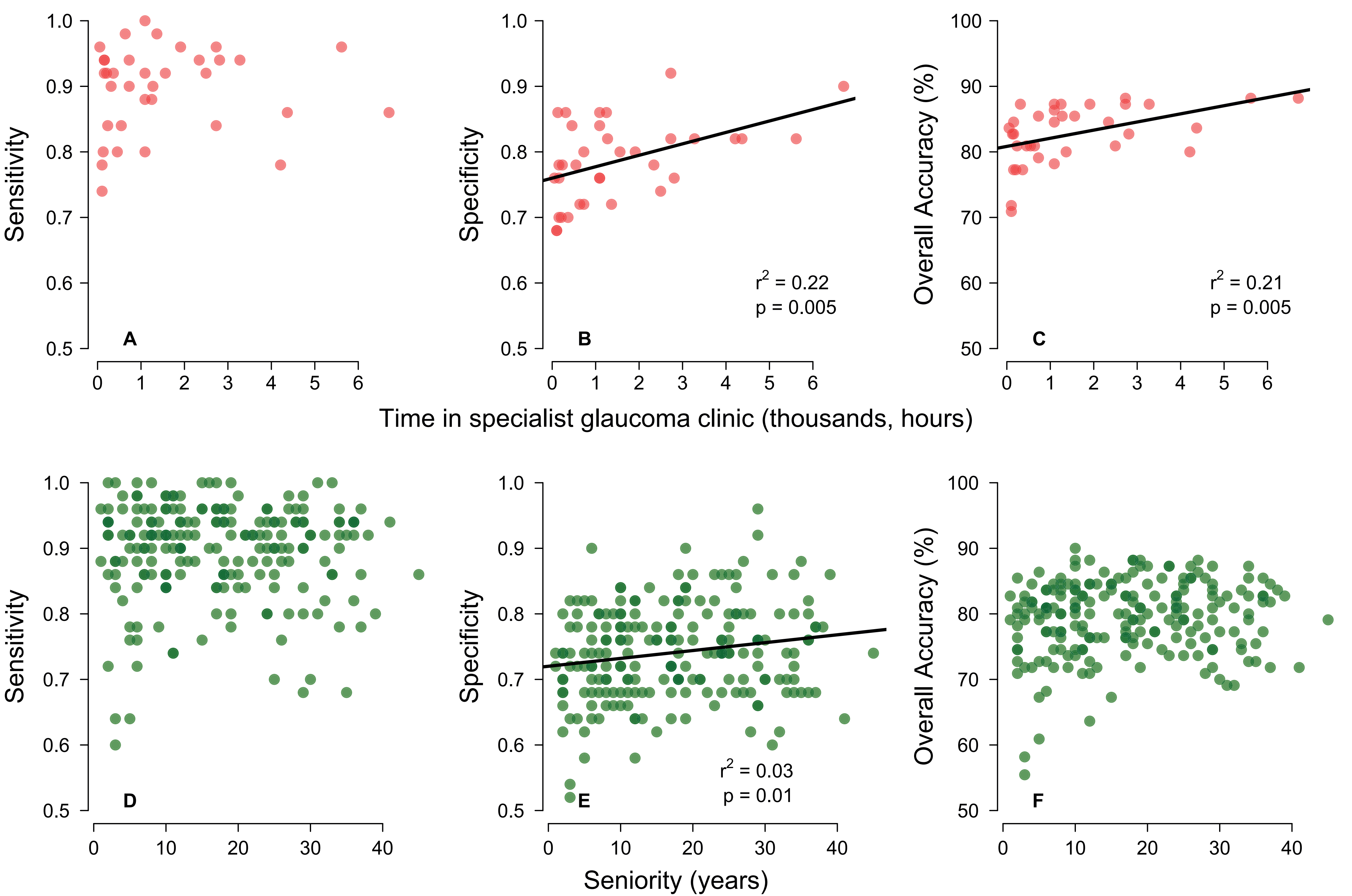
<u>Figure 5</u>: (A) Confidence levels of optometrists, in their assessment of optic discs. Numbers in red indicate the reports of all participants in stage 1. Numbers in light grey indicate those optometrists undertaking stage 2. Numbers in dark grey indicate those with higher qualifications. (B - D) The distribution of performance levels of optometrists reporting each confidence level.


A
B

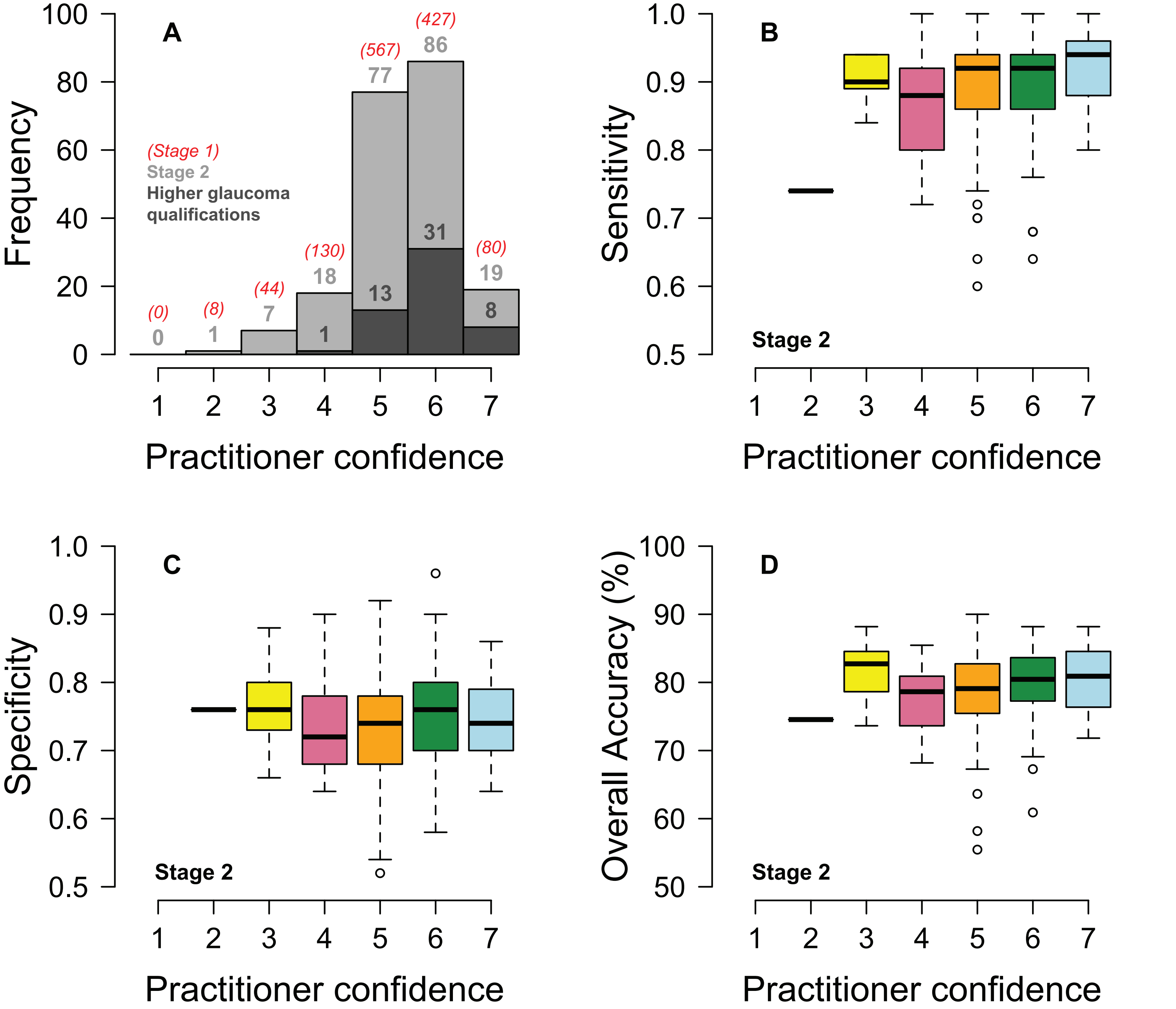


1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity
A
Current Study
Reus et al (2010)
B


Sensitivity
Frequency
p = 0.48
δ
Specificity
p < 0.001
Accuracy



Sensitivity
Specificity
Overall Accuracy (%)
r^2 = 0.22
p = 0.005
r^2 = 0.21
p = 0.005
A
B
C
Time in specialist glaucoma clinic (thousands, hours)
D
E
F
r^2 = 0.03
p = 0.01
Seniority (years)



A
B
C
D
Frequency
Sensitivity
Specificity
Overall Accuracy (%)
Practitioner confidence
(Stage 1)
Stage 2
Higher glaucoma qualifications
(0) 0
(8) 1
(44) 7
(130) 18 1
(567) 77 13
(427) 86 31
(80) 19 8
Stage 2