Dr Sam Philip, MBBS,MD,MRCP is a Specialist Registrar at the Grampian Retinal Screening Programme, NHS Grampian in Scotland. His areas of research and interest are in diabetic retinopathy, retinal imaging and medical informatics.

John Olson is a Consultant Ophthalmic Physician at Aberdeen Royal Infirmary. Special interests in inflammatory disorders, neurological disorders and vascular disorders affecting vision. He is also the Lead Clinician for Scotland's National Managed Clinical Network for Uveitis and NHS Grampian's Eye Health Network.

Correspondence:
Dr Sam Philip,
Grampian Diabetes Centre,
Woolmanhill,
Aberdeen,
AB25 1LD,
UK.
Email: sam.philip@nhs.net

Improving the Efficiency of Diabetic Retinopathy Screening using Automated Retinal Image Analysis

Introduction

Screening for retinopathy is an important part of the annual care of people with diabetes and has been shown to be cost-effective.[1] All four nations in the UK have set up primary care based screening programmes using digital fundus photography as the preferred technique for screening.[2] The number of patients with diabetes is expected to increase dramatically over the coming years and this will place considerable strain on regional retinal screening programmes. Moreover, the national screening committee has strongly recommended that screening programmes be quality assured. Delivering an annual quality assured, screening programme for a constantly increasing population of people with diabetes will require innovative and more efficient methods of screening.[3]

It was in the early 1980s that clinical interest in applying image analysis techniques to retinal photographs started.[4] Various groups developed algorithms to automate the detection and measurement of various normal and abnormal features of interest in the retina.[5] Initial studies were on fluorescein angiograms and with digitised fundus photographs. However the successful introduction of digital fundus photography gave a new impetus to research in this area. The 'Ophthalmic Imaging Group' at the University of Aberdeen has been involved in this field over the last two decades.[6] Initially work involved developing algorithms for microanuerysm (MA) detection in fluorescein angiograms and digitised retinal images. The efficacy of software developed was assessed in a pilot study involving 586 patients and the automated software had a comparable sensitivity (79%) to manual graders (80%).[7] Seventy percent of the patients screened have no retinopathy.[8] Screening normal images is tedious and can increase the probability of missing abnormal images. Hence if automated grading separated patients whose images had no retinopathy, the efficiency of the grading process could be greatly increased.

The Scottish Diabetic Retinopathy Screening Collaborative has been set up according to the recommendations of the Health Technology Board for Scotland.[9] Single 45° disc / macula photographs are taken using a non-mydriatic digital fundus camera. If an adequate quality image is not obtained then the photograph is repeated after mydriasis. Images are then graded using a three level grading process. The level-one grader identifies those patients with any features of retinopathy of whose images are of poor quality and passes them on to the level-two grader. The level-two grader identifies those with features of referable retinopathy and passes them on to the level-three grader who is a consultant ophthalmologist. The level-three grader adjudicates whether a patient needs to be referred to the hospital eye service. The bulk of the manual grading workload is at the level-one stage and we felt that this was the key classification task that should be addressed by automated grading. As microaneurysms are the first visible signs of retinopathy, software that detected microaneurysms could perform the key classification task of identifying abnormal

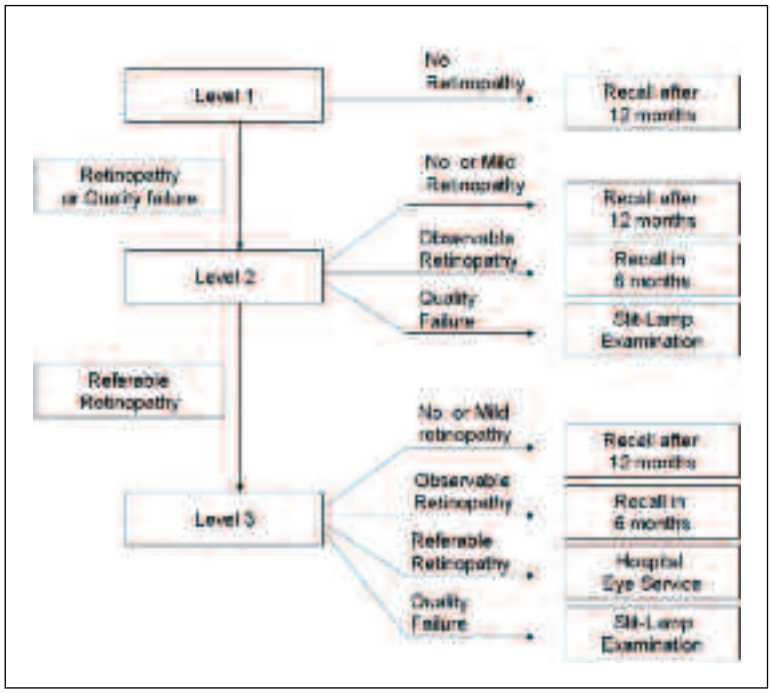

Figure 1: Three level grading structure.

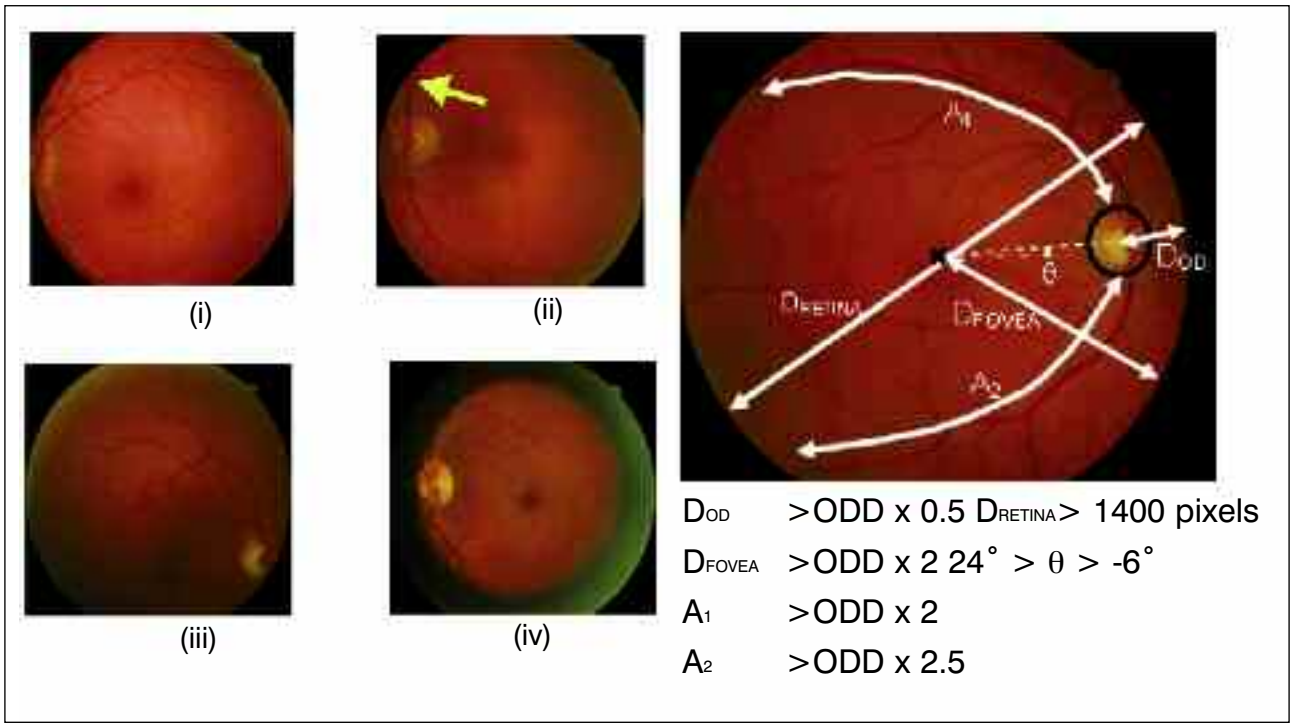

Figure 2a: Examples of an inadequate field definition :
(i) optic disc incomplete
(ii) superior temporal arcade incomplete
(iii) macula incomplete
(iv) restricted field of view.

Figure 2b: Image constraints for field definition: Condition for adequate field definition based on distance of optic disc and fovea from edge (DOD and DFOVEA), visible length of temporal arcades (A1 and A2), diameter of visible retina (DRETINA), fovea to optic disc angle.

images. There was however, a major technical hurdle before automated grading could be implemented. The level-one grader also checked every image to see if it has sufficient quality for grading. So it was essential that this step should also be automated. The efficacy of the combined software compared against manual graders using a large prospective cohort of patients.

Image quality assessment

There are two aspects to retinal image quality namely field definition and image clarity. Field definition is an assessment of whether the appropriate part of the retina was photographed and image clarity checks whether details in the retina are clearly visible. Figures 2a and 2b show examples of inadequate field definition.

The automated software identified the presence and position of the key landmarks in a 45° disc / macula photograph namely the optic disc, the fovea and the vascular arcades. This information was used to assess whether the position of the landmarks were within the limits expected in a standard photograph.[10] The software then identified vessels of all sizes in the retinal image. Image clarity was assessed on how clearly visible the fine vessels in the macula were.[10] An image was classified as having adequate quality if it passed both the field definition and clarity assessments. If the patient did not have at least one image of adequate quality for the right and left eyes, then the screening episode was classified as a technical failure by automated assessment.

Retinopathy detection

In an earlier study we had demonstrated that image processing techniques drawn from the field of mathematical morphology can be used to distinguish the dot like microaneurysms from the retinal vasculature and other dark objects. However, there are many such similar looking structures within the retina and each of these candidates were also evaluated on features like area, eccentricity and mean intensity. Further improvements to the classification process were made during the study by explicitly delineating the vessels and the texture of the retina surrounding each candidate (Fleming 2006). If at least one MA is detected in any of the images of adequate quality then the patient was classified as having retinopathy by automated assessment.

Comparison with manual level-one grading

The performance of the automated software was compared with that of manual level-one grader on retinal images obtained from 6722 consecutive patients attending the Grampian diabetes retinal screening programme.[11] All images were also reference graded by a clinical research fellow who was quality assured by the lead clinician of the retinal screening programme. The reference grader classified 8.2% of the patients as having ungradable images, 62.5% of patients as having no retinopathy and 24.4% as having mild retinopathy (requiring a subsequent screening visit in one year), 1.2% as having observable retinopathy (requiring a subsequent screening visit in six months), and 3.7% as having referable retinopathy (requiring referral to the ophthalmology department).

The sensitivity and specificity for manual and automated systems relative to the reference standard are shown in Table 1. The reference grader classified 330 patients as having observable or referable retinopathy recommending that they should be observed at a shorter screening interval (six months) or referred to the eye clinic respectively. The automated missed seven patients while manual level-one graders missed two. None of the patients missed had new vessels. The difference between the fully manual and combined automated / manual systems for detection of referable or observable retinopathy was not statistically significant. The overall agreement between the reference standard and the manual graders was measured using the kappa statistic, = 0.79 and = 0.53 for manual and auto-

Table 1: Sensitivity and specificity of manual and automated grading.

		Manual grading	Automated grading
Retinopathy detection in good quality images	Sensitivity	83.8%	85.5%
	Specificity	96.3%	79.4%
Referral to next stage grading	Sensitivity	86.5%	90.5%
	Specificity	95.3%	67.4%

mated grading respectively. All cases of referable retinopathy missed by the automated grading system had referable maculopathy and none were found to have sight-threatening macular oedema upon referral to ophthalmology.

Workload reduction

The automated system, despite its lower specificity, would have had a significant impact on grading workload. Of the 6722 patients in the study, all were graded by level-one graders and 2545 were graded by both level-one and level-two graders. If the level-one graders were replaced by our automated software the level-two graders would have graded 3652 patients. This equates to a workload reduction of 60%.

Cost-effectiveness

We assessed the cost-effectiveness of implementing automated grading within the Scottish National Screening Programme. Using a decision tree model base, the prevalence, costs and detection rates from the Grampian Retinal Screening Programme were extrapolated to a national screening programme for an estimated Scottish diabetes population of 160,000.[12] The cost of the automated system included costs for integration with the existing national screening software, salary for a grade 6 computer analyst and annual maintenance charges; all annuitised over 10 years. The resulting costs of grading were £432,000 for manual grading and £230,400 for automated grading representing an annual saving of £201,600.

Conclusions

We have developed and validated the performance of a fully automated level-one diabetic retinopathy grading system in a large prospective study. This is the first system to combine formal image quality assessment and retinopathy detection. The system has been shown to be safe, effective and can provide significant reduction in the manual grading workload. By increasing the efficiency of manual grading it will enable screening programmes to increase their coverage and provide screening at recommended intervals. Automated grading will also reduce the burden of quality assurance by decreasing the number of manual graders required. ***EN***

Take Home Message

- Screening for retinopathy is an important part of the annual care of people with diabetes and has been shown to be cost-effective.
- Screening normal images is tedious and can increase the probability of missing abnormal images.
- The performance of the automated software was compared with that of manual level-one graders on retinal images obtained from 6722 consecutive patients attending the Grampian diabetes retinal screening programme.
- For an estimated Scottish diabetes population of 160,000, automated grading resulted in an annual saving of £201,600.

Acknowledgements

Members of the Ophthalmic Imaging Group at the University of Aberdeen.

Dr John Olson, Clinical Director, Grampian Diabetes Retinal Screening, Foresterhill, Aberdeen.

Prof. Peter Sharp, Medical Physics, University of Aberdeen.

Dr Keith Goatman, Radiology, University of Aberdeen.

Dr Alan Fleming, Radiology, University of Aberdeen.

Mr Graham Scotland, Health Economics Research Unit, University of Aberdeen.

Ms Sofia Fonseca, Department of Public Health, University of Aberdeen.

Mr Gordon Prescott, Department of Public Health, University of Aberdeen.

Lorraine Urquhart and members of her team from the Grampian Retinal Screening Programme.

This project was funded by the Chief Scientist Office, Scottish Executive Health Department (grant number CZH/4/76).

References

1. James M, Turner DA, Broadbent, DM, Vora J, Harding S. Cost effectiveness analysis of screening for sight threatening diabetic eye disease. *BMJ* 2000;**320(7250)**:1627-31.
2. Harding S, Greenwood R, Aldington S, Gibson J, Owens D, Taylor R, Kohner E, Scanlon P, Leese G. Diabetic Retinopathy Grading and Disease Management Working Party. Grading and disease management in national screening for diabetic retinopathy in England and Wales. *Diabet Med* 2003;**12**:965-71.
3. Leese GP, Ellis JD. Quality assurance for diabetic retinal screening. *Diabet Med* 2006;**23(6):**629-34.
4. Baudoin C, Maneschi F, Quentel G, Soubrane G, Hayes T, Jones G, Coscas G, Kohner EM. 1983 Quantitative evaluation of fluorescein angiograms: microaneurysm counts *Diabetes* 1983;**32 (Suppl 2)**:8-13.
5. Teng T, Lefley M, Claremont D. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Med Biol Eng Comput* 2002;**40(1)**:2-13.
6. Cree MJ, Olson JA, McHardy KC, Sharp PF, Forrester JV. A fully automated comparative microaneurysm digital detection system. *Eye* 1997;**11**:622-8.
7. Olson JA, Strachan FM, Hipwell JH, Goatman KA, McHardy KC, Forrester JV, et al. A comparative evaluation of digital imaging, retinal photography and optometrist examination in screening for diabetic retinopathy. *Diabet Med* 2003;**20**:528-34.
8. Philip S, Cowie LM, Olson JA, The impact of the Health Technology Board for Scotland's grading model on referrals to ophthalmology, *British Journal of Ophthalmology* 2005;**89**:891-6.
9. Facey K, Cummins E, Macpherson K, Morris A, Reay L, Slattery J. Organisation of Services for diabetic retinopathy screening. Health Technology Assessment Report 1. Glasgow: Health Technology Board for Scotland; 2002.
10. Fleming AD, Philip S, Goatman KA, Olson JA, Sharp PF. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Invest Ophthalmol Vis Sci* 2006;**47**:1120-5.
11. Philip S, Fleming AD, Goatman KA, Fonseca S, Mcnamee P, Scotland GS, Prescott GJ, Sharp PF, JA Olson. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme, *Br J Ophthalmol* 2007;**91(11)**:1512-7.
12. Scotland GS, McNamee P, Philip S, Fleming AD, Goatman KA, Prescott GJ, Fonseca S, Sharp PF, Olson JA. Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland. *Br J Ophthalmol* 2007;**91(11)**:1518-23.