

Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland

G S Scotland, P McNamee, S Philip, A D Fleming, K A Goatman, G J Prescott, S Fonseca, P F Sharp, J A Olson

Br J Ophthalmol 2007;**91**:1518–1523. doi: 10.1136/bjo.2007.120972

See end of article for authors' affiliations

Correspondence to:
G S Scotland, Research Fellow, Health Economics Research Unit, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD; g.scotland@abdn.ac.uk

Accepted 18 June 2007
Published Online First 21 June 2007

Aims: National screening programmes for diabetic retinopathy using digital photography and multi-level manual grading systems are currently being implemented in the UK. Here, we assess the cost-effectiveness of replacing first level manual grading in the National Screening Programme in Scotland with an automated system developed to assess image quality and detect the presence of any retinopathy.

Methods: A decision tree model was developed and populated using sensitivity/specificity and cost data based on a study of 6722 patients in the Grampian region. Costs to the NHS, and the number of appropriate screening outcomes and true referable cases detected in 1 year were assessed.

Results: For the diabetic population of Scotland (approximately 160 000), with prevalence of referable retinopathy at 4% (6400 true cases), the automated strategy would be expected to identify 5560 cases (86.9%) and the manual strategy 5610 cases (87.7%). However, the automated system led to savings in grading and quality assurance costs to the NHS of £201 600 per year. The additional cost per additional referable case detected (manual vs automated) totalled £4088 and the additional cost per additional appropriate screening outcome (manual vs automated) was £1990.

Conclusions: Given that automated grading is less costly and of similar effectiveness, it is likely to be considered a cost-effective alternative to manual grading.

Systematic screening for diabetic retinopathy has been identified as a cost-effective use of health service resources.^{1–4} The Health Technology Board for Scotland recommended a national screening programme using digital photography and a multi-level manual grading system (fig 1), which is currently being implemented in Scotland.⁵ Similar programmes are also underway in England, Wales and Northern Ireland.

With 161 946 individuals recorded on diabetes registers in Scotland,⁶ manual grading is a resource-intensive activity. Current policy is implemented by capturing digital images at local screening centres, which are then sent electronically to one of nine regional grading centres.⁷

However, a system of automated grading could provide cost savings to the NHS. Our research group recently developed and evaluated an automated grading system that can assess digital retinal images for quality⁸ and the presence of retinopathy.⁹ This system could thereby potentially replace manual level 1 grading. The purpose of this paper is to assess the cost-effectiveness of replacing this manual disease/no disease grading with the automated system, in the context of the three-level grading system used in Scotland (fig 1). A decision tree model was developed to compare NHS grading costs and screening outcomes over a one-year period for these two alternative strategies.

METHODS

The model was developed and analysed using TreeAge Pro 2005 (Treeage Software Inc, Williamstown, MA, USA). Disease prevalence and screening efficacy parameters (table 1) were all obtained from a study assessing the efficacy of automated and manual grading on a test set of 14 406 images from 6722 patients.⁹ The cohort was almost entirely Caucasian, 55% of the patients were male, the median age was 63 (IQR 19) years,

88.5% were over the age of 45, and 52.8% over 65. These demographics are similar to those in the rest of Scotland.

Patients undergoing screening fall within one of four disease categories based on their underlying retinopathy status according to the Scottish Diabetic Retinopathy Grading Scheme:¹⁰ no retinopathy, mild background retinopathy, observable retinopathy (observable background retinopathy and/or observable maculopathy), or referable retinopathy (referable maculopathy, referable background retinopathy, and/or proliferative retinopathy). Details of the grading scheme and assignment of grades are provided in the paper reporting the efficacy findings on which this analysis is based.⁹ Patients within each disease category either experience successful image capture or image capture failure leading to ungradeable images (technical failures). The image sets of all patients are first graded by either manual level 1 graders or the new automated system (fig 1). Patients are then either recalled 1 year later, or their images are referred to a level 2 grader based on efficacy data for the alternative level 1 grading approaches (table 1). Following referral to level 2, patients can either be recalled 1 year later, recalled 6 months later, referred for a slit-lamp examination, or have their images referred to a level 3 grader. The probabilities of each of these four outcomes are reported in table 1, for each underlying category of retinopathy. The detection rates presented reflect the actual decisions made at each level (recall or refer), rather than the retinopathy grades assigned. For the base case analysis, level 3 grading and slit-lamp grading were assumed to be 100% sensitive and specific.

Costs per patient for grading at each of the three levels were estimated from a survey of the diabetic retinopathy screening programme in Grampian. Table 2 outlines how these unit costs were calculated. In addition, the extra cost of quality assuring the manual grading strategy was estimated at £32 340 per annum (see Appendix for full details). Table 3 presents all the

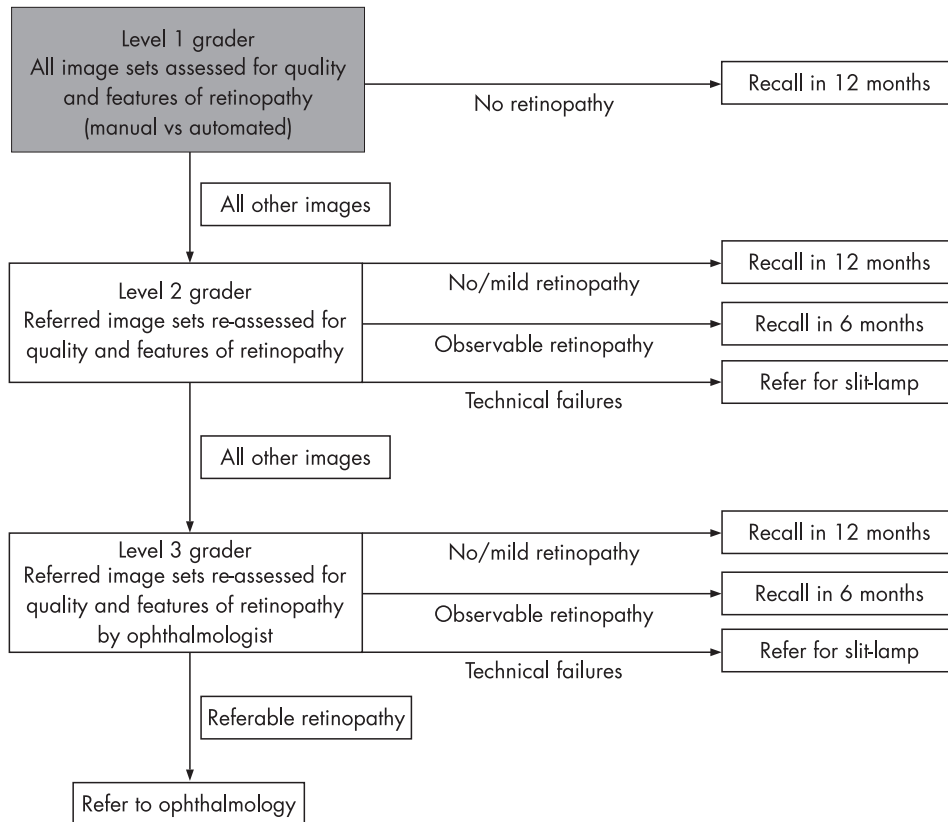


Figure 1 Grading pathway recommended by the Health Technology Board for Scotland, 2002 (the shaded box indicates the choice between manual and automated grading). Level 1 graders first of all assess image quality and identify whether or not images show any signs of retinopathy. Patients whose images show no signs of retinopathy are recalled for a further screening appointment a year later, whereas image sets showing signs of retinopathy, or poor quality image sets, are referred to Level 2 graders. The level 2 graders review these image sets and pass on those with suspected referable retinopathy for final grading by a level 3 grader (ophthalmologist). Level 2 graders can also recall those with no retinopathy or mild retinopathy at one-year, recall those with observable retinopathy at 6 months, and refer patients with poor quality image sets for a slit-lamp examination.

unit costs that were incorporated in the model and the plausible ranges assessed through sensitivity analysis.

A cost per patient for the automated system, was calculated under the assumption that software would run from an existing central server covering all of Scotland. Implementation costs were estimated as the cost of further software development (12 months of a grade 6 computer analyst's time (£31 542)), and the cost of integration with the current system (£60 000) plus ongoing support and maintenance at £12 000 per annum (estimation from Siemens Medical Solutions). The non-recurrent costs were annuitised over a period of 10 years (the assumed useful lifespan of the software) using a discount rate of 3.5%. The total annual equivalent implementation cost (£23 007) was divided by the number of patients screened annually in Scotland (160 000) to give a cost per patient of £0.14. Given the uncertainty surrounding this cost estimate, the impact of halving and doubling the implementation cost was also assessed. All unit costs were estimated in sterling for the year 2005/2006. All other costs relating to the screening programme (eg, IT equipment, office space and overheads) were assumed not to vary between the two grading strategies.

Analysis

Cost-effectiveness was calculated by estimating the grading cost, and number of appropriate outcomes and referable cases detected, for Scotland's diabetic population, assuming 100% coverage and uptake of screening. Appropriate outcomes were defined as final decisions (recalls and referrals) appropriate to actual grade of retinopathy present. The outcomes reported reflect the overall sensitivity and specificity of the three-level grading system (fig 1). The cost per appropriate outcome and cost per referable case detected were calculated for each strategy, as were the corresponding incremental cost-effectiveness ratios.

Sensitivity analysis

To characterise uncertainty in base case estimates, alternative cost and effectiveness estimates were computed by varying all individual parameters, one at a time, within their 95% confidence intervals or plausible ranges presented in tables 1 and 3. Following this, the impact of changing several key assumptions was assessed. First, changes were made to annual throughput. Secondly, as the automated system would result in more patient referrals to level 2 graders, and thus possibly more patients being recalled at 6 months, the impact of including the cost of extra screening visits was assessed. Finally, the impact of reducing the sensitivity and specificity of level 3 graders was calculated.

In addition, probabilistic sensitivity analysis using Monte Carlo simulation was employed.¹² Values were simultaneously selected for each parameter from an assigned distribution and the results recorded. The process was repeated one thousand times to give a distribution of cost and effect differences between the two strategies. Beta distributions were assigned to all the probability parameters as is recommended where the data informing parameters are binomial.¹² Gamma distributions were chosen to represent variation in cost variables, as health care costs most frequently follow this type of distribution (ie, skew to the right and constrained to be non-negative). Finally, due to the greater uncertainty surrounding the cost of automated grading, we halved and doubled this cost and assigned a uniform distribution (table 3).

RESULTS

Base case

Table 4 shows the results obtained using the base case point estimates. For a diabetic population of 160 000, with prevalence of referable pathology detectable by digital imaging at 4% (6400 true cases), the partially automated strategy would be expected to identify 5560 cases (86.9%) and the manual strategy 5610

Table 1 Key efficacy variables used in the model

Variables	Point estimate (95% CI)
Prevalence variables	
Prevalence (normal cases)	0.681 (0.670–0.692)
Prevalence (mild retinopathy)	0.266 (0.255–0.276)
Prevalence (observable retinopathy/maculopathy)	0.013 (0.011–0.016)
Prevalence (referable retinopathy/maculopathy)	0.040 (0.035–0.045)
Technical failure rate	0.082 (0.076–0.089)
Efficacy of level 1 manual grading	
Detection rate for technical failures	0.969 (0.951–0.981)
Proportion of normal cases appropriately recalled	0.920 (0.911–0.927)
Detection rate for mild retinopathy	0.819 (0.800–0.837)
Detection rate for observable retinopathy/maculopathy	1.000 (0.956–1.000)
Detection rate for referable retinopathy/maculopathy	0.992 (0.971–0.998)
Efficacy of level 1 automated grading	
Detection rate for technical failures	0.998 (0.990–1.000)
Proportion of normal cases appropriately recalled	0.674 (0.660–0.688)
Detection rate for mild retinopathy	0.859 (0.841–0.875)
Detection rate for observable retinopathy/maculopathy	0.976 (0.917–0.993)
Detection rate for referable retinopathy/maculopathy	0.980 (0.953–0.991)
Efficacy of level 2 manual grading	
Technical failures	
Detection rate for technical failures*	0.819 (0.786–0.852)
Proportion of technical failures incorrectly referred to level 3	0.103 (0.077–0.128)
Proportion of technical failures incorrectly recalled at 12 months	0.076 (0.054–0.099)
Proportion of technical failures incorrectly recalled at 6 months	0.002 (0.000–0.010)
No retinopathy	
Proportion of normal cases appropriately recalled*	0.938 (0.912–0.964)
Proportion of normal cases incorrectly referred to slit-lamp	0.056 (0.032–0.081)
Proportion of normal cases incorrectly referred to level 3	0.006 (0.002–0.021)
Proportion of normal cases incorrectly recalled at 6 months	0.000 (0.000–0.011)
Mild retinopathy	
Detection rate for mild cases*	0.955 (0.944–0.966)
Proportion of mild cases incorrectly referred to slit-lamp	0.007 (0.004–0.014)
Proportion of mild cases incorrectly referred to level 3	0.025 (0.016–0.033)
Proportion of mild cases incorrectly recalled at 6 months	0.013 (0.007–0.019)
Observable retinopathy	
Detection rate for observable cases*	0.607 (0.503–0.712)
Proportion of observable cases incorrectly referred to slit-lamp	0.000 (0.000–0.044)
Proportion of observable cases incorrectly referred to level 3	0.143 (0.068–0.218)
Proportion of observable cases incorrectly recalled at 12 months	0.250 (0.157–0.343)
Referable retinopathy	
Detection rate for referable cases*	0.885 (0.845–0.925)
Proportion of referable cases incorrectly referred to slit-lamp	0.000 (0.000–0.016)
Proportion of referable cases incorrectly recalled at 6 months	0.045 (0.019–0.071)
Proportion of referable cases incorrectly recalled at 12 months	0.070 (0.038–0.102)

*Denotes appropriate level 2 grading outcomes for each category of retinopathy (beta distributions were applied to all efficacy parameters in probabilistic sensitivity analysis).

cases (87.7%). However, automated grading would result in cost savings to the NHS of £201 600 per year. The additional cost per additional referable case detected (manual vs automated) was calculated as £4088. The number of referable cases detected by the alternative strategies is a function of the combined efficacy of the level 1 and level 2 grading systems in performing their respective tasks (table 1). A greater number of cases of referable retinopathy are missed at the level 2 stage due to the fact that level 2 graders perform the more difficult task of assigning grades of retinopathy.

In terms of all screening outcomes, the partially automated strategy would be expected to deliver 158 170 appropriate outcomes compared to 158 271 expected with the current manual system. The additional cost per additional appropriate screening outcome (manual vs automated) was calculated at £1990.

Sensitivity analysis

The results were most sensitive to changes in the proportion of referable cases detected by manual level 1 and automated

Table 2 Model grading cost variables, assumptions and values

Grading level	Salary scale	Total annual cost (salary, on costs and training)	Weeks worked/year	Hours/week	Hourly cost	Grading rate (patients per h)	Grading cost per patient
Level 1	D-grade nurse	£23 248.40	42	37.5	£14.76	10.00	£1.48
Level 2	E-grade nurse	£25 577.65	42	37.5	£16.24	11.50	£1.41
Slit-lamp	F-grade nurse	£29 868.31	42	37.5	£18.96	4.29	£4.42
Level 3	Hospital consultant	£91 276.35	41	48.2	£46.19	15.00	£3.08

Note: time worked per year/week taken from the Unit Costs of Health and Social Care.¹¹

Table 3 Cost parameters and ranges used in the model

Variable	Cost per patient (range) (£)	Distribution for sensitivity analysis
Level 1 manual grading (D-grade nurse)	1.48 (1.14–2.11)	Gamma (α 10.95, β 0.135)
Additional QA costs (manual grading)	0.20 (0.15–0.25)	Gamma (α 16, β 0.0125)
Level 1 automated grading	0.14 (0.07–0.29)	Uniform (min 0.07, max 0.29)
Level 2 grading (E-grade nurse)	1.41 (1.12–1.98)	Gamma (α 9.95, β 0.142)
Level 3 grading (consultant ophthalmologist)	3.08 (2.57–3.85)	Gamma (α 47.43, β 0.065)
Slit-lamp grading (F-grade nurse)	4.42 (3.16–5.69)	Gamma (α 130.24, β 0.034)
6 month re-screen (manual system)	13.95*	
6 month re-screen (automated system)	12.70*	
Ophthalmology referral	65.00†	

*Based on fixed screening cost plus grading costs per patient in Grampian.
†Average cost for an ophthalmology outpatient visit in Scotland (http://www.isdscotland.org/isd/files/Costs_RO44X_2005.xls).

grading. When these parameters were varied within their 95% confidence intervals, the strategy that identified the most referable cases changed. The automated system became more effective when the detection rate of manual level 1 graders for referable retinopathy fell below 98.2%. By contrast, it remained less costly when all cost parameters were varied individually within the ranges presented in table 3. All else remaining constant, the cost per patient for automated level 1 grading would have to reach £1.40 for the automated strategy to become more costly than manual grading. This would require the equivalent annual implementation cost to be greater than £224 000.

Including the cost of re-screening individuals recalled inappropriately at 6 months had no effect on the cost-effectiveness of strategies. For changes in uptake, if only 80% of people with diabetes attended screening (128 000), the cost per patient for automated grading would increase to £0.18 and the additional quality assurance cost associated with the manual system would decrease to £0.19 per patient. The cost saving associated with the automated strategy would be £154 880, and the number of referable cases detected would be 4448 compared to 4488 for manual grading. With 50% uptake, the automated strategy would reduce grading costs by £86 400, and reduce the number of referable cases identified by 25 (from 2805 to 2780).

When the specificity of level 3 graders was set at 80% as opposed to 100%, the manual and automated strategies resulted in 227 and 262 inappropriate referrals to ophthalmology respectively. Thus the cost savings were slightly lower with this scenario (£198 400 compared to £201 600). When the specificity of slit-lamp graders was also set at 80%, the cost savings associated with automated grading decreased further to £176 000. By contrast, when the sensitivity of level 3 and slit-lamp grading was set at 80%, the difference in effectiveness between the two strategies, in terms of the overall number of referable cases identified, was reduced (automated grading identifying 40 fewer cases than manual grading for a cohort of 160 000).

Figure 2 summarises the results of the probabilistic sensitivity analysis by plotting the 1000 estimates of the difference in cost and effect of the automated vs the manual strategy, where effect is measured as the probability of detecting a true referable case per patient screened. Of the points, 18% lie in

the quadrant that clearly favours automated grading (ie, where automated is less costly and more effective). The remaining points lie in the quadrant where automated grading is less costly but also slightly less effective.

DISCUSSION

The results reported here suggest that automated grading could reduce the workload and associated costs of grading patients' images for diabetic retinopathy. The partially automated strategy was calculated to reduce grading and quality assurance costs by £201 600 for a cohort of 160 000 people with diabetes. Despite having lower specificity, the cost savings associated with the automation of level 1 grading more than offset the cost of increased referrals to level 2 graders. However, automated grading also reduced the number of referable cases identified by 50 (5560 cases compared to 5610 cases). Cost-effectiveness was most sensitive to changes in the level 1 detection rates for referable retinopathy/maculopathy. If the proportion of referable cases detected by manual level 1 graders were to fall below

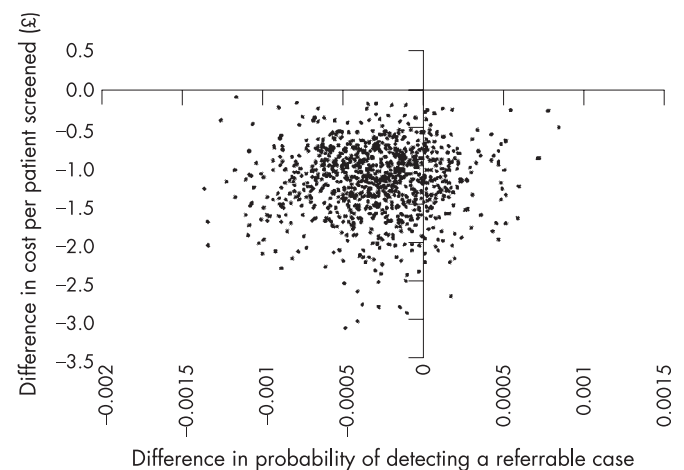


Figure 2 Incremental cost-effectiveness scatter plot of automated vs manual grading from 1000 Monte Carlo simulations.

Table 4 Base case results for a cohort of 160 000 patients

Strategy	Total cost of grading	Referable cases detected	Appropriate outcomes	Additional cost per additional referable case detected	Additional cost per additional appropriate outcome
Automated	£230 400	5560	158 170		
Manual	£432 000	5610	158 271	£4088*	£1990†

*Differs slightly from £201 600/50 due to rounding.

†Differs slightly from £201 600/101 due to rounding.

98.2%, the automated strategy would not only be cheaper, but also more effective.

The saving in grading cost might be a conservative estimate as it assumes that D-grade nurses, some of whom also act as screening photographers, undertake all manual level 1 grading at the rate of 10 patients/h. While it might be possible to conduct manual level 1 grading at a higher rate than this, it is likely that this would require more experienced and more expensive staff. There could be unforeseen costs associated with implementation of the automated system, but sensitivity analysis indicates that automation of level 1 grading would still result in savings of £179 200 if implementation costs were doubled. Another potential benefit of automated grading is that it could potentially reduce the time it takes for patients to receive their results, due to the fact it could run 24 h a day.

A weakness of this analysis is that it only considers costs and consequences over a one-year time horizon. This period was chosen partly because of uncertainty surrounding the implications of missing referable cases as defined by the Scottish Diabetic Retinopathy Grading Scheme.¹⁰ The base case analysis predicts that for a screening cohort of 160 000, automated grading would reduce by 50 the number of referable cases identified (referable maculopathy, referable background retinopathy, or proliferative retinopathy). This prediction is due to a small non-significant difference in the referable retinopathy detection rates reported for automated and manual level 1 grading.⁹ In a direct comparison on 14 406 images from 6722 patients, automated grading missed three more cases of referable disease than manual grading. However, all of these were cases of referable maculopathy, as opposed to referable or proliferative retinopathy, and none were found to have sight threatening macular oedema upon referral to ophthalmology. Data from the Grampian screening programme suggests that only 12% of maculopathy cases referred to ophthalmology actually have indications of macular oedema when examined by slit-lamp biomicroscopy (John A Olsen, unpublished data). Estimates from the literature suggest an annual hazard rate for loss of central acuity of 0.017 for untreated vs treated macular oedema.¹³ This equates to a 0.2% (0.12×0.017) chance of missed referable maculopathy cases leading to sight loss within the space of a year (the maximum screening interval). Therefore, the small increase in false negatives with automated grading is unlikely to lead to additional cases of sight loss and additional health care costs.

Relation with other studies

Despite increased interest in the automation of retinopathy grading,^{14–16} the cost-effectiveness of an automated system compared with manual grading has only been assessed in one previous study,¹⁷ which found that automation was less cost-effective than manual grading due to having lower sensitivity. Our estimate of cost-effectiveness is more favourable due to substantial improvements in the sensitivity of automated detection. Automated image analysis systems have also been developed for use in screening programmes for various cancers.^{18–19} In a recent review, Willis and colleagues found insufficient evidence to draw conclusions on the cost-effectiveness of using automated image analysis in screening programmes for cervical cancer.¹⁸ This was partly due to uncertainty surrounding the cost of implementation – some estimates indicating that the required technology might be prohibitively expensive. Prohibitive implementation cost is unlikely to be a barrier to automated retinopathy grading in Scotland, as the infrastructure required (IT links, national database, universal screening software) is already in place. However, the cost-effectiveness of automated grading will be less favourable in areas where such infrastructure is lacking.

Implications for policy

National implementation of automated grading would reduce the grading and quality assurance costs of the National Diabetic Retinopathy Screening Programme in Scotland. The question of whether automated grading is more cost-effective than manual grading depends on whether the probable effects of a slight reduction in the probability of identifying true referable cases, likely to be around 0.03%, are outweighed by the likely cost savings of around £200 000 per year. NHS Quality Improvement Scotland (NHSQIS) recommend that clear grading errors - failure to notice unequivocal signs of referable retinopathy or failure to notice that an image is of insufficient quality for grading - should not exceed 1 in every 200 patients screened.²⁰ For the 6722 patients studied in Grampian,⁹ the clear error rate for automated grading (1 in 1120) was actually lower than that for manual level 1 grading (1 in 258) due to its higher sensitivity for technical failures. Thus, as automated grading operates well within the clinical standards recommended by NHSQIS, it is likely to be considered a cost-effective alternative to manual grading in Scotland.

ACKNOWLEDGEMENTS

We thank all the members of the Grampian Retinal Screening Programme who supported this project, especially Lorraine Urquhart and Julie Hughes who provided information on resource use and unit costs for the cost analysis. We also thank Paul Burdett (Siemens Medical Solutions), Angela Ellingford (Tayside Retinopathy Screening Programme), Andy Davis (NHS National Services Division), and Ian Esslemont and Gary Mortimer (NHS Grampian) for the information they provided for the cost analysis. JAO, PFS, KAG, GJP and PM designed the study and obtained funding. SP, JAO and KAG coordinated the study on a day-to-day basis. SP acted as the gold standard against which the effectiveness of the alternative grading approaches was assessed. ADF developed the image processing computer algorithms underlying the automated system. SP, ADF, SF, and GJP conducted the analysis of the effectiveness data. GSS, SP, JAO, KAG collected the resource use and cost data for the economic analysis. GSS and PM conducted the cost-effectiveness modelling and took the lead on writing the paper. All authors commented, contributed to, and reviewed the drafts and final version of the paper.

Authors' affiliations

G S Scotland, P McNamee, Health Economics Research Unit, University of Aberdeen, Foresterhill, Aberdeen
S Philip, J A Olson, Grampian Retinal Screening Programme, David Anderson Building, Foresterhill, Aberdeen
A D Fleming, K A Goatman, Department of Biomedical Physics, University of Aberdeen, Foresterhill, Aberdeen
G J Prescott, S Fonseca, P F Sharp, Department of Public Health, University of Aberdeen, Foresterhill, Aberdeen

Funding: This study was funded by the Chief Scientist Office of the Scottish Executive Health Department (SEHD). The views expressed here are those of the authors and not necessarily those of the SEHD.

Competing interests: Implementation in Scotland is being considered. If this occurs it is likely that there will be some remuneration for the University of Aberdeen, NHS Grampian and the Scottish Executive.

Guarantor: GS.

REFERENCES

- 1 Maberly D, Walker H, Koushik A, et al. Screening for diabetic retinopathy in James Bay, Ontario: a cost-effectiveness analysis. *JAMC* 2003;**168**:160–4.
- 2 Davies R, Roderick P, Canning C, et al. The evaluation of screening policies for diabetic retinopathy using simulation. *Diab Med* 2002;**19**:762–70.
- 3 James M, Turner DA, Broadbent DM, et al. Cost-effectiveness analysis of screening for sight threatening diabetic eye disease. *BMJ* 2000;**320**:1627–31.
- 4 Javitt J, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Annals of Internal Medicine* 1996;**124**:164–9.
- 5 Facey K, Cummins E, Macpherson K, et al. *Organisation of services for diabetic retinopathy screening*, Glasgow, Health Technology Board for Scotland; 2002, Health Technology Assessment Report 1.

- 6 **Scottish Executive Health Department.** *Scottish diabetes survey 2004.* Edinburgh: SEHD, 2005.
- 7 **National Services Division.** *Diabetic retinopathy screening services in Scotland: outline vision of a national service.* Edinburgh: NSD, 2005.
- 8 **Fleming AD,** Philip S, Goatman KA, *et al.* Automated assessment of diabetic retinal image quality based on clarity and field definition. *Invest Ophthalmol Vis Sci* 2006;**47**:1120–5.
- 9 **Philip S,** Fleming AD, Goatman KA, *et al.* The efficacy of automated “disease/no disease” grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol* 2007;**91**:1512–7.
- 10 **NHS Quality Improvement Scotland.** *Diabetic retinopathy screening services in Scotland: recommendations for implementation.* Edinburgh: NHSQIS, 2003.
- 11 **Netten A,** Curtis L. *Unit costs of health and social care.* Personal Social Services Research Unit. Canterbury: University of Kent, 2004.
- 12 **Briggs AH.** Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;**17**:479–500.
- 13 **Javitt JC,** Aiello LP, Chiang Y, *et al.* Preventive eye care in people with diabetes is cost-saving to the federal government: implications for the federal government. *Diabetes Care* 1994;**17**:909–17.
- 14 **Hipwell JH,** Strachan F, Olson JA, *et al.* Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool. *Diabet Med* 2000;**17**:588–94.
- 15 **Larsen M,** Godt J, Larsen N, *et al.* Automated detection of fundus photographic red lesions in diabetic retinopathy. *Invest Ophthalmol Vis Sci* 2003;**44**:761–6.
- 16 **Niemeijer M,** van Ginneken B, Staal J, *et al.* Automatic detection of red lesions in digital color fundus photographs. *IEEE Trans Med Imaging* 2005;**24**:584–92.
- 17 **Sharp PF,** Olson J, Strachan F, *et al.* The value of digital imaging in diabetic retinopathy. *Health Technol Assess* 2003;**7**. <http://www.hta.ac.uk/fullmono/mon730.pdf> (accessed 6 September 2007).
- 18 **Willis BH,** Barton P, Pearmain P, *et al.* Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK. *Health Technol Assess*, 2005;**9**. <http://www.hta.ac.uk/fullmono/mon913.pdf> (accessed 6 September 2007).
- 19 **Taylor P,** Champness J, Given-Wilson R, *et al.* Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography. *Health Technol Assess*, 2005;**9**. <http://www.hta.ac.uk/fullmono/mon906.pdf> (accessed 6 September 2007).
- 20 **NHS Quality Improvement Scotland.** *Diabetic retinopathy screening – clinical standards.* NHS Quality Improvement Scotland: Edinburgh, 2004.

APPENDIX 1 ESTIMATE OF ADDITIONAL QUALITY ASSURANCE COSTS FOR THE MANUAL GRADING STRATEGY

1. Number of quality assurance (QA) reviews required annually per grader = 500
2. Cost per QA review = £1.54

3. Estimated number of level 1 graders in Scotland = 50
*Estimated quality assurance costs for manual level 1 graders
= $50 \times 500 \times £1.54$
= £38 500
4. Number of QA reviews required for the automated system = 2500 (assumption)
5. Number of patients to be screened per year in Scotland = 160 000
5. Proportion of patients referred to level 2 with manual system = 0.38
6. Proportion of patients referred to level 2 with automated system = 0.54
*Estimated increase in the number of patients referred to level 2 with the automated system compared to manual.
= $(160\ 000 \times 0.54) - (160\ 000 \times 0.38)$
= 25 600
*Estimated number of patients graded annually by one WTE level 2 grader = 16 800 (based on grader working 42 weeks per year, 5 days a week, at rate of 80 patients per day)
*Estimated number of level 2 WTE graders required to grade additional referrals
= $25\ 600 / 16\ 800$
= 1.5
*Estimated actual numbers of graders required to grade additional referrals (assuming graders work half-time
= 1.5×2
= 3
*Estimated costs of quality assuring the automated grading strategy relative to the manual strategy
= £1.54 (2500+ (3×500))
= £6160
*Estimated additional QA cost of the manual strategy compared to the automated strategy
= £38 500 – £6160
= £32 340
= £0.20 per patient screened