Costs and consequences of automated algorithms versus manual grading for the detection of referable diabetic retinopathy


Record Status
This is a critical abstract of an economic evaluation that meets the criteria for inclusion on NHS EED. Each abstract contains a brief summary of the methods, the results and conclusions followed by a detailed critical assessment on the reliability of the study and the conclusions drawn.

CRD summary
This study assessed the cost-effectiveness of two automated grading system algorithms in screening for diabetic retinopathy, in Scotland. The authors concluded that automated exudate and haemorrhage detection was more cost-effective than the previous automated algorithm, and more cost-effective than manual grading. The methods were satisfactory and the results were well reported. The authors’ conclusions appear to be appropriate, but the source studies should be consulted to assess the validity of the evidence.

Type of economic evaluation
Cost-effectiveness analysis, cost-utility analysis

Study objective
This study assessed the cost-effectiveness of two automated grading system algorithms in screening for diabetic retinopathy in Scotland.

Interventions
Two comparisons were made; two algorithms for level-one grading were compared and automatic level-one grading was compared with manual level-one grading, which was the usual practice. Algorithm A was an image quality assessment and detection of microaneurysms and dot haemorrhages. Algorithm B was the same as algorithm A, but also detected blot haemorrhages, and macular exudates. Algorithm B was used for the comparison between manual and automated grading.

The images were graded as having no retinopathy, mild retinopathy, observable retinopathy or maculopathy, referable retinopathy or maculopathy, or image quality failed.

Location/setting
UK/secondary care.

Methods
Analytical approach:
A developed decision model was used to evaluate and compare the relative cost-effectiveness of the two options for each of the two comparisons. This model was extrapolated over 20 years. The authors did not report the study perspective.

Effectiveness data:
The key epidemiological and efficacy data were from two similar studies (Scotland, et al. 2007, and Philip, et al. 2007, see 'Other Publications of Related Interest' below for bibliographic details). The main clinical effectiveness estimates were the specificity and sensitivity of the grading systems.

Monetary benefit and utility valuations:
The utility decrements for vision loss that were used to derive the quality-adjusted life-years (QALYs) were from published literature.
Measure of benefit:
There were three benefit measures: the number of referable cases detected; the appropriate screening outcomes; and QALYs, which were used for the 20-year extrapolation and were discounted at a rate of 3.5% per annum.

Cost data:
The cost categories were level-one and level-two manual grading, level-one automated grading, ophthalmology consultant level-three grading and referrals, and slit-lamp grading. These costs were from Scotland, et al. 2007 and a survey of five Scottish grading centres. The costs were expressed in 2005 to 2006 UK pounds sterling (£) and were discounted at a rate of 3.5% per annum for the 20-year extrapolation.

Analysis of uncertainty:
A probabilistic sensitivity analysis was used to estimate the distribution of the cost and effect differences between the automated algorithms. A deterministic sensitivity analysis, considering eight scenarios, was conducted for the comparison between manual and automated grading.

Results
Compared with algorithm A, algorithm B had an incremental cost of £7,759; detected an additional 113 referable cases; and had an additional 104 appropriate screening outcomes. This produced an incremental cost-effectiveness ratio of £68 per additional referable case detected or £75 per additional appropriate screening outcome.

Compared with the automated algorithm (B), manual grading had an incremental cost of £212,695; detected an additional 123 referable cases; and had an additional 734 appropriate screening outcomes. This produced an incremental cost-effectiveness ratio of £1,727 per additional referable case detected or £289 per additional appropriate screening outcome.

The probabilistic sensitivity analysis indicated that, in terms of referable cases detected, algorithm B had a 95.1% change of being more effective and more costly and a 4.1% chance of being less costly and more effective. Above a willingness-to-pay threshold of £68 per additional referable case detected, algorithm B had a higher probability of being cost-effective than algorithm A.

Over 20 years, manual grading was found to cost between £25,676 and £267,115 per additional QALY gained, compared with automated grading (algorithm B). The sensitivity analysis found that the incremental cost per case detected for manual compared with automated grading was sensitive to small changes in the sensitivity of manual level-one graders.

Authors’ conclusions
The authors concluded that automated exudate and haemorrhage detection was more cost-effective than the previous automated algorithm, and it had acceptable effectiveness and reduced costs compared with manual grading, making it cost-effective.

CRD commentary
Interventions:
The interventions were described and they included the usual practice. They might be relevant in other settings.

Effectiveness/benefits:
The authors used three measures of effectiveness for the interventions and the reporting of these measures was good. Most of the effectiveness data were from two published studies, which were not sufficiently described to assess the validity and quality of the data; the original publications should be consulted. The benefit measures appear to have been appropriate; they incorporated both disease-specific measures and the more generic QALY, which evaluates both morbidity and mortality, and is comparable with other disease interventions. The QALYs were appropriately discounted, but it was unclear if the other benefit measures were discounted, which increases the uncertainty in these estimates.

Costs:
The authors did not explicitly report the study perspective, but it seems to have been that of the Scottish Health Service. The cost data for automated level-one grading were from another publication by the same authors. The costs for level-two and level-three grading were not reported in detail, making it unclear if all the relevant costs were included. The costs were appropriately adjusted for inflation.

**Analysis and results:**
The analytic approach appears to have been appropriate. The results were presented sufficiently, with more detail in two appendices. The impact of uncertainty appears to have appropriately assessed. The authors discussed the relationship between this and their previous study and the policy implications. They identified and discussed some limitations to their study.

**Concluding remarks:**
The methods were satisfactory and the results were well reported. The authors’ conclusions appear to be appropriate, but the previous studies should be consulted to assess the validity of the evidence.

**Funding**
Funded by the Chief Scientist Office of the Scottish Government Health Directorates (SGHD).

**Bibliographic details**

**PubMedID**
19965826

**DOI**
10.1136/bjo.2008.151126

**Original Paper URL**
http://bjo.bmj.com/content/94/6/712.abstract

**Other publications of related interest**


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Algorithms; Cost-Benefit Analysis; Decision Trees; Diabetic Retinopathy /complications /diagnosis /economics; Diagnosis, Computer-Assisted /economics /methods; Diagnostic Techniques, Ophthalmological; Exudates and Transudates /metabolism; Health Care Costs /statistics & numerical data; Humans; Image Interpretation, Computer-Assisted /methods; Mass Screening /economics /methods; Quality-Adjusted Life Years; Retinal Hemorrhage /etiology; Scotland; Severity of Illness Index

**AccessionNumber**
22010001260