

Sensitivity and specificity of a new scoring system for diabetic macular oedema detection using a confocal laser imaging system

L Tong, A Ang, S A Vernon, H J Zambarakji, A Bhan, V Sung, S Page

Abstract

Aim—To assess the use of the Heidelberg retina tomograph (HRT) in screening for sight threatening diabetic macular oedema in a hospital diabetic clinic, using a new subjective analysis system (SCORE).

Methods—200 eyes of 100 consecutive diabetic patients attending a diabetologist's clinic were studied, all eyes had an acuity of 6/9 or better. All patients underwent clinical examination by an ophthalmologist. Using the HRT, one good scan was obtained for each eye centred on the fovea. A System for Classification and Ordering of Retinal Edema (SCORE) was developed using subjective assessment of the colour map and the reflectivity image. The interobserver agreement of using this method to detect macular oedema was assessed by two observers (ophthalmic trainees) who were familiarised with SCORE by studying standard pictures of eyes not in the study. All scans were graded from 0–6 and test positive cases were defined as having a SCORE value of 0–2. The sensitivity of SCORE was assessed by pooling the data with an additional 88 scans of 88 eyes in order to reduce the confidence interval of the index.

Results—12 eyes in eight out of the 100 patients had macular oedema clinically. Three scans in three patients could not be analysed because of poor scan quality. In the additional group of scans 76 out of 88 eyes had macular oedema clinically. The scoring system had a specificity of 99% (95% CI 96–100) and sensitivity of 67% (95% CI 57–76). The predictive value of a negative test was 87% (95% CI 82–99), and that of a positive test was 95% (95% CI 86–99). The mean difference of the SCORE value between two observers was -0.2 (95% CI -0.5 to $+0.07$).

Conclusions—These data suggest that SCORE is potentially useful for detecting diabetic macular oedema in hospital diabetic patients.

(*Br J Ophthalmol* 2001;85:34–39)

underpins the importance of screening for maculopathy in diabetics.

Recently, we have described a new method for the detection and quantification of macular oedema by volumetric analysis using the Heidelberg retinal tomograph (HRT).³ This study was performed in patients referred to the ophthalmology department with a suspicion of maculopathy.

This study evaluates the role of the HRT in screening the maculas of unselected patients attending a hospital based clinic for control of their diabetes. In an attempt to improve on our original method of analysis, we have developed the System of Classification and Ordering of Retinal Edema (SCORE) which uses the subjective evaluation of the HRT colour map and the reflectivity images for the detection of macular oedema.

Methods

PATIENTS

The entry criteria for the study were:

- (1) Patients attending a hospital diabetic clinic for management of their diabetes.
- (2) Best corrected visual acuity in each eye of 6/9 or better.
- (3) There must be no impairment of ability to perform the HRT scanning—that is, no serious musculoskeletal deformity or inability to understand simple instructions.
- (4) No previous ophthalmic treatment such as laser treatment or treatment for glaucoma.

Patients with previous squint surgery who fulfilled the other criteria were included.

All suitable patients were recruited when they were reviewed at a hospital diabetologist's clinic.

All patients gave informed consent and the study was approved by the ethics committee of Queen's Medical Centre, Nottingham.

All patients had cyclopentolate 1% eye drops in each eye before examination. Slit lamp biomicroscopy using a 78D lens was performed by an experienced ophthalmologist (LMT). A Goldmann fundus contact lens was used for borderline cases with possible macular oedema. The presence of clinically significant macular oedema (CSMO) according to ETDRS⁴ criteria was noted, and if so the eye was deemed to be "positive for oedema".

HEIDELBERG RETINAL TOMOGRAPH

Details of the acquisition of HRT images have been described in a previous paper.⁵ HRT software 2.01 was used in image acquisition, image alignment, as well as data analysis. An operator

Departments of
Ophthalmology and
Endocrinology,
Queen's Medical
Centre, University
Hospital, Nottingham
NG7 2UH, UK

L Tong
A Ang
S A Vernon
H J Zambarakji
A Bhan
V Sung
S Page

Correspondence to:
Mr S A Vernon
stephen_vernon@hotmail.com

Accepted for publication
2 August 2000

Diabetic maculopathy is the major cause of visual loss in diabetics.¹ In the early stages most patients have good vision and a recent large audit indicated that 86.5% of those cases requiring treatment have no symptoms.² This

with at least 1 month's experience performed HRT with the machine (LMT).

One good quality scan of each eye was utilised in all analyses. A 2 mm diameter circle was drawn using the circle draw facility of the HRT and this was centred on the fovea. This was performed by visualising the pattern of the retinal vasculature on the reflectivity image. The positioning of the reference plane was performed as described previously.⁵

We devised a novel system for detecting macular oedema using the HRT, called SCORE. SCORE represents the sum of two grades. The first grade is classified from the extent of retinal elevation using the HRT colour map facility, and the second grade from the HRT reflectivity image. A combined score of 0–2 was designated to “abnormal” or requiring “ophthalmic referral” and 3–6 was “normal” or “not requiring referral”.

Two observers, trainees in ophthalmology (AB and VS), were trained in SCORE by familiarising themselves with a set of standard images (see Figs 2 and 3) and some instructions on the use of these images. These observers had no knowledge pertaining to the clinical findings in these patients. The ability of these observers to use these standard images to grade patients' images was assessed by a set of “test images” of patients with and without macular oedema. This was to ensure that the observers could use SCORE in a reliable and consistent way. In order to minimise observer fatigue, the grading of the study patients was performed over a few sessions lasting not longer than 1 hour each.

Colour map scoring

Under the stereo measurement feature of the HRT software, it was possible to superimpose a colour map over the 2 mm diameter circle of the topographic image. As the topographic map was coded by the position rather than magnitude of the maximum reflectivity along the z-axis (axis perpendicular to the two dimensional HRT images), we speculated that it would detect macular oedema more predictably than the reflectivity image. Visual evaluation of the topographic image without further software analysis would be difficult, as there is little contrast between points on this image. The colour map illustrates the topography of the macula within the 2 mm diameter circle using colour codes. The colour map was graded 0–3 where 0 corresponds to the greatest area of retinal elevation and 3 to the least (Fig 1).

Reflectivity image scoring

The reflectivity images were also graded by the two observers independently from 0 to 3 where 0 corresponds to the most abnormal image and 3 to a definitely normal image. A grade of 0 was given in cases where discrete white spots or grains were observed within the temporal arcades; or in cases where blot haemorrhages could be seen in all four quadrants of the macula. A grade of 3 was given for a “normal reflectivity image” which could be quite variable in appearance (Fig 3). Normal appear-

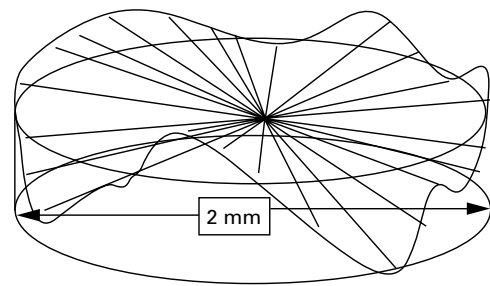


Figure 1 Diagram showing the definition of coloured regions within the 2 mm diameter circle. The top circle is the mean height of the contour line shown as the curved line. The curved surface is shown as a series of lines joining a central point of the circle at the level of the mean height of the contour line to an infinite number of points along the contour line at the circle perimeter.

ance may include that of red “dots” over the fovea without involving all quadrants of the macula, or small “dot haemorrhages” without “blot haemorrhages in the macula”, or a “bow tie” type of appearance over the centre of the macula (Fig 3).

DATA ANALYSIS

The number of ungradable images was recorded. The reasons for any failure to assign SCORE were noted.

For the first observer, the specificity and predictive value calculations, together with 95% confidence intervals using SCORE to detect eyes “positive for oedema” were performed in each eye as well as in each patient.

The post-test likelihood of oedema (predictive value of positive test) was calculated using the Bayes theorem:

$$\text{Post-test probability} = \frac{(\text{pretest probability}) \times (\text{sensitivity})}{((\text{pretest probability} \times \text{sensitivity}) + (1 - \text{pretest probability})(1 - \text{specificity}))}$$

The pretest probability would be the disease prevalence in the screening population.

Because of the low prevalence of cases “positive for oedema” expected in such a population, we doubted whether it would be possible to obtain a sensitivity and positive predictive value with narrow 95% confidence limits. The sensitivity and positive predictive value of SCORE was therefore assessed on the scans above pooled with an additional 88 scans of 88 eyes of diabetic patients with a greater prevalence of clinical macular oedema. These patients were referred to the ophthalmic department for suspected retinopathy. The corrected visual acuity in this group of eyes was similar to that of the first cohort—that is, 6/9 or better. The clinical examination, HRT imaging, and analysis and SCORE values were obtained in an identical fashion to that described above, although the HRT imaging was performed by another operator experienced in the technique (HJZ). It is important to note that the graders performed many grading sessions over different days, without any knowledge regarding the proportion of cases with clinical oedema in each session. This proportion could vary greatly between sessions as the scans were randomly distributed to different sessions.

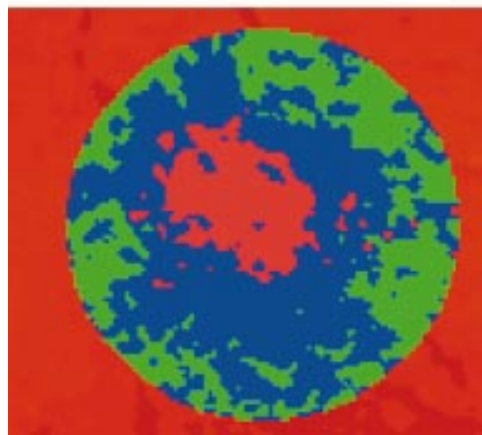
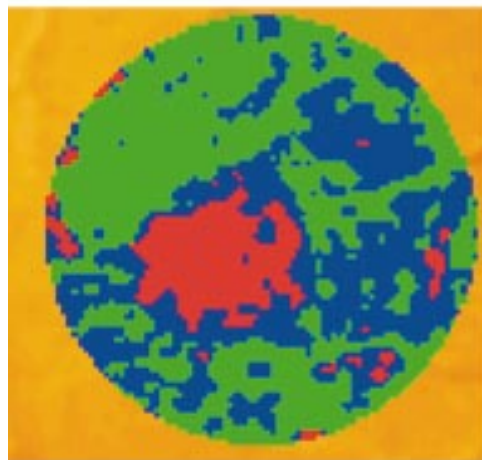
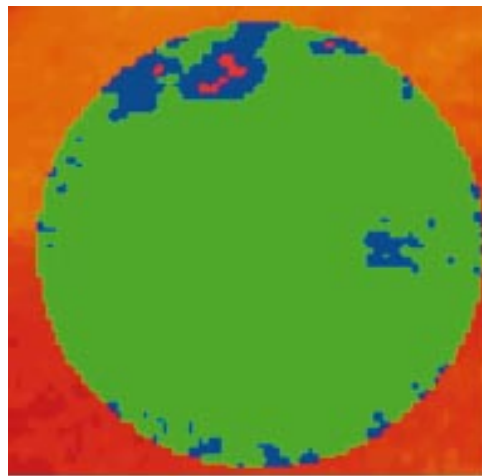


Figure 2 Standard colour maps. From top to bottom, Grade 0, 2, 3. Any maps with extent of green area intermediate between 0 and 2 would be graded as 1. Green equates with an area of retina above the curved surface (defined in Fig 1), blue relates to the area of retina whose height lies below the curved surface but above the plane defined by the lowest point on the contour line, and red defines the area of retina lying below the plane of the lowest point of the contour line.

In order to assess the interobserver agreement of the technique, the raw SCORE value of the second observer was compared with that of the first. For this purpose only the SCORE from all the right eyes of the 100 patients screened in the diabetic clinic was used. A plot

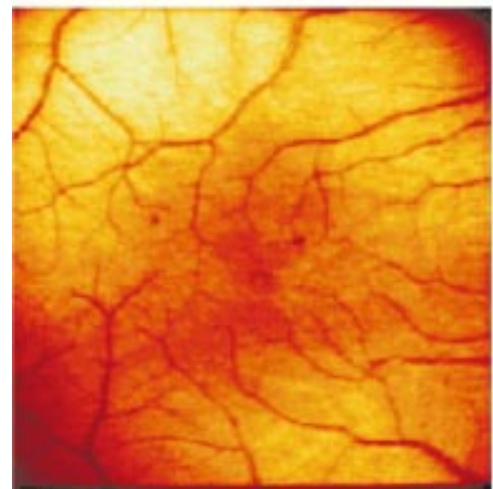
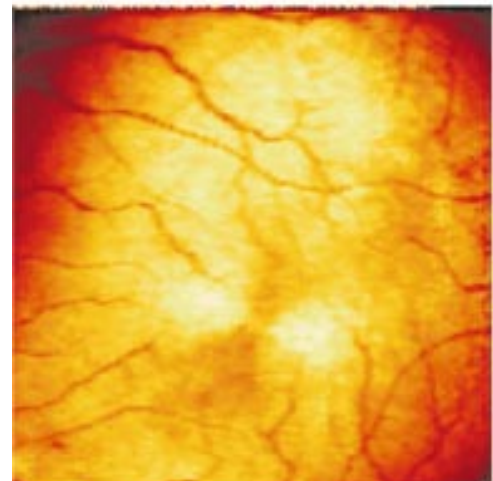
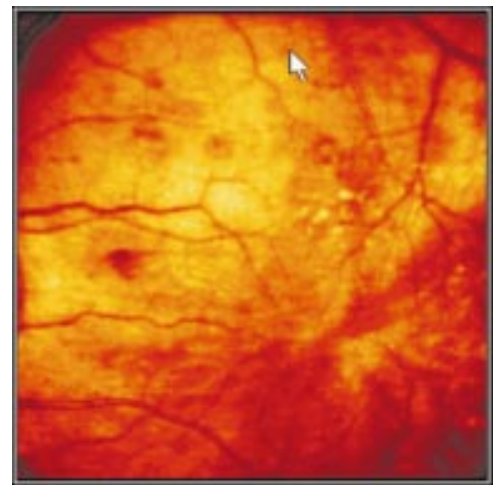


Figure 3 Examples of standard reflectivity images. From top to bottom, grade 0 (showing discrete yellow spots and blot retinal haemorrhages), grade 3 (example 1), grade 3 (example 2). Grade 1 would be designated if a grader is more than 50% but less than 90% certain that the image should be graded as grade 0. Grade 2 would be designated if a grader is more than 50% certain but less than 90% certain that an image should be graded as grade 3.

of the SCORE difference between each pair of readings was made against the mean of each pair of SCORE readings to ensure that there is no systematic relation between the magnitude of the differences when the mean increases. The mean difference and 95% confidence

Table 1 Specificity, sensitivity, positive and negative predictive values for observer 1 in percentages (95 CI)

	Specificity	Sensitivity	Positive predictive values	Negative predictive values
Patients (97)	100 (96–100)	75 (32–95)	100 (60–100)	97 (91–99)
Eyes (197)	100 (100–100)	58 (29–82)	100 (78–100)	97 (94–98)
Pooled eyes (285)	99 (96–100)	67 (57–76)	95 (86–99)	87 (82–91)

A

By patients

		Clinical finding		total	3 ungradable
		0	1		
HRT finding	0	89	2	91	
	1	0	6		
total		89	8	97	

B

By eyes

		Clinical finding		total	3 ungradable
		0	1		
HRT finding	0	185	5	190	
	1	0	7		
total		185	12	197	

C

By pooled eyes

		Clinical finding		total	3 ungradable
		0	1		
HRT finding	0	194	29	223	
	1	3	59		
total		197	88	285	

Figure 4 Two × two tables illustrating the results of screening by observer 1: 0 refers to a negative result; 1 refers to a positive finding.

interval of the differences were calculated. The repeatability coefficient, defined as $2 \times$ (standard deviation of the differences), was also calculated as was the kappa statistic.

Results

We studied 200 eyes (100 patients). There were 51 males and 49 females. The patients' mean age was 58.51 years (SD 13.06 years, range 27–81 years).

Twelve eyes in eight patients were “positive for oedema” clinically. Three scans in three patients could not be analysed because of poor scan quality, none of whom had clinical macular oedema.

The sensitivity, specificity, and predictive values of the scoring system are shown in Table 1. As expected, the sensitivity and positive predictive values of the test were difficult to interpret because of the wide confidence interval. For this reason, Table 1 also shows results following pooling of the additional 88 scans. Figure 4 shows the 2×2 tables for observer 1, where the indices were calculated. Within the subgroup of 88 scans, the prevalence of eyes “positive for oedema” on clinical examination was 86% (76/88). Analysis of the “pooled scans” was only performed on eyes rather than on patients as not all patients in the additional group had both eyes included. (Some of these patients, unlike those from the “screening

population”, had one eye which failed the visual acuity criterion.)

Using the Bayes theorem, the post-test probability of a “positive test patient” to have macular oedema in at least one eye would be 100%.

In the interobserver agreement section of the study, the SCORE difference of each SCORE pair from the two observers was not related to the magnitude of the mean SCORE. The correlation coefficient ($r = -0.1$) was close to 0 and hence the differences were not correlated with the mean of each pair. The mean difference was found to be -0.2 (95% CI -0.5 to $+0.07$). The repeatability coefficient was 3.0 and the kappa statistic was 0.45 (indicates good agreement).

Discussion

Our data suggest that, in a hospital setting, the use of the HRT with SCORE may be potentially useful to screen for asymptomatic diabetic macular oedema.

The main strengths of using SCORE in screening are its high specificity and relative ease of performance. In our previously published report³ the volume above reference plane (VARP) was used to identify CSMO. Using the suggested cut off points of the VARP index at 1.8, the sensitivity to detect CSMO (effectively eyes positive for oedema) per eye was 58% (95% CI: 38%–78%).³ The specificity, however, was 75% (66–84%), and the predictive values of a positive test and a negative test were 39% (23–59%) and 87% (79–94%), indicating a high rate of false positive results. In the same study, when reflectivity images and pseudo three dimensional maps were evaluated subjectively in 41 patients for CSMO, the sensitivity was 100% and predictive value of the negative test was 91% (71–100%); however, both the specificity and the positive predictive value were only 37% (20–56%). The corresponding figures using SCORE in this study represent a significant improvement.

Altering the cut off value to “a positive screen equates with a SCORE of 3 or less” increases the sensitivity to 87.5% but reduces the specificity to 84.3% when pooled eyes were examined. This indicates that increasing the sensitivity cannot be achieved without a significant loss of specificity and that our original point on the receiver operating curve (SCORE of 2 or less equals a positive screen) is the most appropriate.

The differences between the two SCORE observers were not significantly different from zero at the level of $p=0.05$, as illustrated by the confidence interval. The measured coefficient of repeatability may reflect the ability of SCORE to detect finer levels of differences between images.⁶

In theory the colour maps used in this study lie within a circle of 2 mm diameter and may miss cases of significant oedema. In practice, SCORE is a composite grade also taking into account the features of the reflectivity images outside the above circle.

Essential to our system of oedema detection are good quality scan images. The scan quality

of images with the HRT may vary with experience as at present there are no formal training courses for HRT operators. Furthermore, experience with using SCORE may also influence the grading capability of assessors. We have found it quite demanding for an observer to assess SCORE on a large number of images in one sitting. Although we have not formally tested the following, one method to reduce observer errors would be to introduce "check images" which were identical to the standard images. Failure to grade these "check images" correctly would strongly suggest assessor fatigue. We do not recommend assessing too many images in one sitting. Assessment of less than 60 scans in one session would be reasonable.

This study involved patients from a diabetic medical clinic, and therefore our results may not be applicable to diabetic patients from the general population, where a lower incidence of true positives might be found. For example, diabetic patients under the care of only general practitioners may have less severe systemic disease and a lower incidence of retinopathy and maculopathy. In addition, patients in the hospital may be better at keeping still and maintaining fixation and therefore be more likely to provide a well centred scan.

The detection of CSMO is adopted by many clinicians as the indication to perform laser treatment.⁴ A strength of this study was the use of a clinically accepted gold standard for the identification of treatable disease—that is, stereo biomicroscopy of the macula. Current screening for maculopathy by non-ophthalmologists relies largely on the presence of macular exudates and haemorrhages rather than macular oedema⁷⁻⁹ but treatment for maculopathy might not be necessary in the presence of macular exudates.¹⁰ Moreover, about 20% of maculopathy presents as retinal thickening as the sole manifestation without exudates.² Identification of oedema rather than the signs often associated with oedema could therefore be considered advantageous when screening for treatable maculopathy from diabetes.

In addition, treatment in cases of early disease (that is, patients with good vision) is more effective because the treated patients tended to retain good vision years later.¹¹ Hence it is important to pick up macular oedema early so that it can be assessed and treated.

Currently, stereoscopic methods of examination such as slit lamp biomicroscopy or stereo pair photographs are not widely used in screening systems.¹² In previously reported findings on screening systems:

(1) The precise method of ophthalmoscopy was not stated in an optometric study.¹³ It was possible that some optometrists used direct ophthalmoscopy whereas other might have used stereoscopic slit lamp systems or a combination of these, depending on available facilities or their own competence. The sensitivity was 77% for detecting "moderate to severe maculopathy".

- (2) Many reports on ophthalmoscopy¹⁴⁻¹⁶ did not compare their maculopathy findings with a gold standard examination—that is, contact lens examination with slit lamp biomicroscopy.
- (3) The results of some studies on ophthalmoscopy would not apply to screening in a larger population because of the extensive clinical experience of the screener (registrars in ophthalmology,¹² or doctor with special interest in diabetes and previous training in eye clinics¹⁷).
- (4) No specific figures for ability to detect maculopathy were reported in some studies.^{18 19}
- (5) In one study, the sensitivity of photography with mydriasis was 61% for detecting sight threatening maculopathy, and the specificity was 99%.¹² A significant proportion of cases were ungradable in some photographic studies.^{12 15 17} In the Liverpool study¹² ungradable cases were the result of media opacity in 9% and for other reasons in 2%.

The cost effectiveness of our technique also needs to be evaluated before considering introducing the HRT in a screening programme. As the SCORE system is currently relatively labour intensive, it would be helpful if future HRT software packages could incorporate SCORE in a semiautomated fashion to enable the HRT operator to refer eyes failing the test directly to an ophthalmologist.

The high specificity supports the possible role of the HRT in screening a primary care population, where one would expect the prevalence of CSMO to be lower than that found in our study. Is the level of sensitivity acceptable for screening? There are two important features of this study which may be relevant to this discussion. Firstly, the eight "positive" patients in the group from the diabetic clinic had been screened and found to be negative by the diabetic physicians (by direct ophthalmoscopy). Secondly, our definition of a true positive relied upon a contact lens examination of the macula which would be expected to be more sensitive than the methods used in other screening studies for the detection of early retinal thickening (this could also go some way towards explaining the previous observation). Thus we consider that the sensitivity found in our study should not be considered too low to countenance continuing our investigation of the efficacy of SCORE. Indeed, it is possible that some of the cases missed by SCORE would not progress significantly between screening intervals and might be detected at the next screen. Only a randomised controlled trial of different screening modalities over time against an acceptable gold standard will identify if this is the case.

Proprietary interests: none.
Financial support: none.

1 Government Statistical Service, DHSS. *Causes of blindness and partial sight among adults in 1976/7 and 1980/81 in England*. London: HMSO, 1988.

2 Bailey CC, Sparrow J, Grey RHB, et al. The National Diabetic Retinopathy Laser Treatment Audit. I Maculopathy. *Eye* 1998;12:69-76.

- 3 Zambarakji H, Butler T, Vernon SA. Assessment of the Heidelberg retina tomograph in the detection of sight threatening diabetic retinopathy. *Eye* 1999;**13**:136–44.
- 4 Early Treatment Diabetic Retinopathy Study Group, Early Treatment Diabetic Retinopathy Study Report No 1. Photocoagulation for Diabetic Macular Edema. *Arch Ophthalmol* 1985;**103**:1796–806.
- 5 Zambarakji H, Evans J, Amoaku W, *et al.* Reproducibility of volumetric measurements of normal maculae with the Heidelberg retina tomograph. *Br J Ophthalmol* 1998;**82**:884–91.
- 6 Bailey IL, Bullimore MA, Raasch TW, *et al.* Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci* 1991;**32**:422–32.
- 7 Retinopathy Working Party (Kohner E, Porta M). A protocol for screening for diabetic retinopathy in Europe. *Diabetic Med* 1991;**8**:263–7.
- 8 Royal College of Ophthalmologists. *Guidelines for diabetic retinopathy*. London: RCO, 1997.
- 9 British Diabetic Association. *Retinal photography screening for diabetic eye disease*. London: British Diabetic Association, 1997.
- 10 Cook H, Newsome R, Long V, *et al.* The natural history of diabetic macular streak exudates: evidence from a screening programme. *Br J Ophthalmol* (in press).
- 11 Early Treatment Diabetic Retinopathy Study Group, Early Treatment Diabetic Retinopathy Study Report No 9. Early photocoagulation for diabetic retinopathy. *Ophthalmology* 1991;**98**:766–85.
- 12 Harding S, Broadbent D, Neoh C, *et al.* Sensitivity and specificity of photography and direct ophthalmoscopy in screening for sight threatening eye disease: the Liverpool diabetic eye study. *BMJ* 1995;**311**:1131–5.
- 13 Hammond C, Shackleton J, Flanagan D, *et al.* Comparison between an ophthalmic optician and an ophthalmologist in screening for diabetic retinopathy. *Eye* 1996;**10**:107–12.
- 14 Burns-Cox C, Hart JCD. Screening of diabetics for retinopathy by ophthalmic opticians. *BMJ* 1985;**290**:1052–4.
- 15 Taylor R, Lovelock L, Tunbridge W, *et al.* Comparison of non-mydriatic retinal photography with ophthalmoscopy in 2159 patients: mobile retinal camera study. *BMJ* 1990;**301**:1243–7.
- 16 Sussman E, Tsurias W, Soper K. Diagnosis of diabetic eye disease. *JAMA* 1982;**247**:3231–4.
- 17 Williams R, Nussay S, Humphrey R, *et al.* Assessment of non-mydriatic fundus photography in detection of diabetic retinopathy. *BMJ* 1986;**293**:1140–2.
- 18 Buxton M, Sculpher M, Ferguson B, *et al.* Screening for treatable diabetic retinopathy: a comparison of different methods. *Diabetic Med* 1991;**8**:371–7.
- 19 O'Hare J, Hopper A, Madhavan C, *et al.* Adding retinal photography to screening for diabetic retinopathy: a prospective study in primary care. *BMJ* 1996;**312**:679–82.



Sensitivity and specificity of a new scoring system for diabetic macular oedema detection using a confocal laser imaging system

L Tong, A Ang, S A Vernon, et al.

Br J Ophthalmol 2001 85: 34-39

doi: 10.1136/bjo.85.1.34

Updated information and services can be found at:

<http://bjo.bmj.com/content/85/1/34.full.html>

These include:

References

This article cites 9 articles, 6 of which can be accessed free at:

<http://bjo.bmj.com/content/85/1/34.full.html#ref-list-1>

Article cited in:

<http://bjo.bmj.com/content/85/1/34.full.html#related-urls>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

[Epidemiology](#) (756 articles)

Notes

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>