
A FULLY AUTOMATED COMPARATIVE MICROANEURYSM DIGITAL DETECTION SYSTEM

MICHAEL J. CREE¹, JOHN A. OLSON², KENNETH C. McHARDY³, PETER F. SHARP¹ and JOHN V. FORRESTER⁴

Aberdeen

SUMMARY

A fully automated digital image processing system, which provides an objective and repeatable way to quantify microaneurysms in digitised fluorescein angiograms, has been developed. The automated computer processing includes registration of same-eye retinal images for serial studies, cutting of regions-of-interest centred on the fovea, the detection of microaneurysms and the comparison of serial images for microaneurysm turnover. The microaneurysm detector was trained against a database of 68 images of patients with diabetes containing 394 true microaneurysms, as identified by an ophthalmologist. The microaneurysm detector achieved 82% sensitivity with 2.0 false-positives per image. An independent test set, comprising 20 images containing 297 true microaneurysms, was used to compare the microaneurysm detector with clinicians. The microaneurysm detector achieved a sensitivity of 82% for 5.7 false-positives per image, whereas the clinician receiver-operator-characteristic (ROC) curve gives 3.2 false-positives per image at a sensitivity of 82%. It is concluded that the computer system can reliably detect microaneurysms. The advantages of the computer system include objectivity, repeatability, speed and full automation.

Microaneurysm counting is a useful technique for assessing the development and progression of diabetic retinopathy,^{1,2} but has proved to be a time-consuming and tedious operation that is subject to observer error. Automating counting procedures on computer, so that little or no operator intervention is

required, provides the potential for efficient and easily repeatable analysis.

Early attempts at automated microaneurysm detection^{3,4} have addressed the detection process for isolated images and relied on simplistic morphological and thresholding techniques. Unfortunately, the threshold can only be adjusted to achieve good sensitivity at the expense of poor specificity. Furthermore the monitoring of diabetic retinopathy requires a serial study of each patient-eye, and this issue has not been addressed.

Spencer *et al.*⁵ significantly improved the specificity of microaneurysm detection by using over-sensitive morphological/thresholding techniques to identify all possible candidate (microaneurysm) locations. A region-growing algorithm is then used at each candidate location to find the underlying candidate morphology. Knowledge of the candidate morphology is necessary for subsequent classification of the candidates. Eight features (such as area and total intensity) are measured on each candidate. Analysis of the features enables microaneurysms to be distinguished from other microvascular abnormalities and background texture. The method of Spencer *et al.*, while a significant improvement on previous attempts, is still suboptimal in terms of execution speed and its ability to distinguish microaneurysms from other pathologies. In addition, the method requires human intervention to identify the specific region of the fundus for counting, and to register images for serial studies.

Ideally, image registration and region identification should be automated as well as microaneurysm counting. The need for image registration arises from the nature of monitoring the progression of retinopathy, which requires a serial study of each patient-eye, with the images being taken, typically, months apart. To ensure consistent microaneurysm counting between images, the same region-of-interest (ROI) must be used for each image. Normal procedures for capturing fluorescein images cannot ensure that

From: ¹Department of Bio-medical Physics and Bio-engineering, University of Aberdeen, Foresterhill, Aberdeen; ²Eye Clinic, Aberdeen Royal Infirmary, Foresterhill, Aberdeen; ³Diabetic Clinic, Aberdeen Royal Infirmary, Woolmanhill, Aberdeen; ⁴Department of Ophthalmology, University of Aberdeen, Foresterhill, Aberdeen, UK.

Correspondence to: Dr M. J. Cree, Department of Bio-medical Physics and Bio-engineering, University of Aberdeen, Foresterhill, Aberdeen AB9 2ZD, UK. Tel: +44 (0)1224 681818 ext 53812. Fax: +44 (0)1224 685645.

exactly the same region of the eye is captured at each patient visit; thus some method is required to ensure a consistent ROI for each image. One method is to register the 'later' images to the baseline image. This process involves rotating, scaling and translating the later image until its permanent features are matched to those in the baseline image. Then a ROI defined on the baseline image is also valid for any later images that have been registered, and microaneurysm counting that has been performed within the ROI will be consistent from image to image.

The microaneurysm detection system described here is based upon the method of Spencer *et al.*⁵ but involves a complete redesign of the region-growing and classification algorithm used in microaneurysm detection, and addresses automated registration and identification of the macula. This paper reports the improvements in the image processing strategy and gives evidence that the microaneurysm detection system is adequately robust for clinical use.

MATERIALS AND METHODS

Fluorescein angiograms of 35° field-of-view were captured with a Topcon fundus camera onto 35 mm film. Normally one frame at the early venous stage was chosen for analysis; however, a later frame was used if no earlier frames of satisfactory quality were available. In this way, any complications due to leakage were minimised. The chosen frame was digitised by back-illumination of the film-strip and the image recorded with a Kodak MegaPlus CCD (charged-coupled device) camera at 1024 × 1024 pixel resolution. Each pixel (picture-element) in the digitised image represents the intensity of light received from the corresponding point of the film with a number in the range 0–255. The resultant pixel resolution achieved is approximately 8 μm. In a negative the fovea appears bright white, and the brightness of each image, before digitisation, was adjusted so that a few pixels in the foveal region were just saturated. In this manner, the full dynamic range of the digitiser was exploited. A 'flood' image was also captured with no negative in the film-strip holder. The captured fluorescein image was divided by the flood image and then rescaled to the original intensity contrast range, to correct for any non-uniform illumination present in the digitisation system.

All computer processing was carried out on a SUN-Sparc IPX workstation, whose speed is comparable to a mid-range 80486 IBM-compatible PC. All processing was carried out on the images in the negative state. However, for presentation purposes, the fluorescein angiograms are shown here as positives.

For a serial same-eye study, a number of images of the same patient-eye were captured. A cross-

correlation algorithm⁶ was used to automatically register the later images to the baseline image. This process can be time-consuming – it takes approximately 30 minutes per image on a SUN IPX – but with current state-of-the-art workstations can be completed within 5 minutes per image pair. If there were laser burns due to photocoagulation performed during the period being studied, or if the images were of poor photographic quality and contained significant regions of poor contrast, then the registration was unlikely to succeed. Out of 142 registrations, 20 initially failed. Of the 20 failures, 12 were finally registered by adjusting parameters of the registration program, or by recapturing a different frame if the image was of poor contrast. The final 8 registrations were abandoned; however, manual registration techniques could have been used to successfully register those image pairs.

The fovea was found by correlating the baseline image with a model that approximates the gross shading of a fluorescein image over the macular region. The peak of the correlation function gave the location of the fovea. For this study a 512 × 512 pixel ROI was defined to be centred on the fovea. This is large enough to cover the macula, which is of primary interest due to its functional importance. The same ROI was applied to all later images of the same-eye study; the registration process ensured that it was valid for all the images. All microaneurysm counting occurred within the ROI only. The automated ROI calculation process is almost instantaneous and is robust; we have observed it to fail on only two images (out of 95), which had shading very similar to that of the fovea near the edge of the image.

As a check on accuracy two clinicians were asked to identify the fovea in eight digitised fluorescein angiograms on two separate occasions. The automated foveal detector was also run on the same eight images. The Euclidean distance between two points was used as a measure of error. The two clinicians had a mean error of 13.8 pixels (± 8.5, SD) and 10.1 pixels (± 4.7) between their two observations. There was a mean error of 10.2 pixels (± 2.9) between the two clinicians. The automated fovea detector had a mean error of 19.2 pixels (± 12.2) from the average of the four clinician points of each image. The worst case was 41.2 pixels. Note that 19.2 pixels corresponds to approximately 0.15 mm or about 1/10 optic disc diameters.

An advantage of automated ROI placement is that should the study need to be repeated then exactly the same ROI will be found. For manual techniques, the regions-of-interest have to be carefully documented to enable selection of precisely the same ROI at a later date.

Automated Microaneurysm Detection

The technical details of computed microaneurysm detection are discussed by Spencer *et al.*⁵ In short, the fluorescein image is processed to remove all gross shading due to the macular region and also due to limitations in image capture. Shade-correction, as the process is called, enables global image processing operators to run more successfully. It is achieved by grossly median-filtering the image to give an approximation to the background fluorescence, which is then subtracted off the original image to give the shade-corrected image.⁵ The vessel structure is removed from the image with a top-hat transform using a long linear mask. The top-hat transform achieves the desired effect by first creating an image that contains all the structure that matches up with various orientations of the mask (in this case, vessels which have long linear sections tend to match well) and then subtracting that image off the original image. It is necessary to remove the vasculature from the image since the next stage, matched-filtering, is susceptible to inadvertent inclusion of vessels. A matched-filter is used to locate all possible microaneurysms in the image. It does this by comparing a model of a microaneurysm (a two-dimensional Gaussian function with $\sigma = 2$ pixels) with each position in the image and generating a number that describes how well that part of the image matches the model. By only accepting those numbers that are greater than a certain threshold value, candidate microaneurysm positions within the image are found. A low threshold value is chosen making it oversensitive so that (ideally) all true microaneurysms present in the image are detected, along with a number of false-positives which bear similarity to microaneurysms. Up to this point the processing for microaneurysm detection is identical to that as described by Spencer *et al.*⁵ and more detail can be found there.

To increase specificity a region-growing algorithm is used on the shade-corrected image, at each candidate's position, to find the morphology of the candidate. Various features of the candidate's morphology are used to classify it as either a microaneurysm or a spurious object. Spencer *et al.*⁵ extract eight features on each candidate; four are morphologically based (for example, how round the candidate is) and four are measurements of the candidate's intensity of fluorescence. For the present study, the features were augmented by a further four intensity-based measurements and the matched-filter response. The additional four intensity measurements were the same as those given by Spencer *et al.* but were multiplied by a scale factor calculated from the image's general contrast. This helped to give better discrimination between microaneurysms and other microvascular abnormalities.

While *a priori* knowledge of microaneurysms can be useful (they are classically said to be round and hyperfluorescent) it was found necessary to use empirically gained knowledge to derive a useful classifier. The classifier is that part of the microaneurysm detector that uses the measured features of each candidate to classify it as either a microaneurysm or a spurious object. The training and testing of the classifier is discussed below.

Having empirically derived a rule-based classifier, the classifier was implemented in the microaneurysm detector software. For a serial same-eye study, the microaneurysm detector was run on the ROI of each image separately, and recorded the locations of the microaneurysms and their total numbers.

Training and Testing the Classifier

To train the classifier a training set of 68 fluorescein images (of 26 patients, and no more than three images of the same patient obtained at different visits) was used. A modified version of the microaneurysm detector was run on each image up to and including feature extraction of each candidate. The same program, without giving any information of what the candidates may be, recorded an ophthalmologist's opinion of what objects were microaneurysms. This was done by displaying the image on a computer screen and allowing the ophthalmologist, who was experienced in viewing fluorescein images on computer screen, to click with a mouse pointer on every microaneurysm in the image. The programme then saved the feature data, including a label for each candidate object of whether the ophthalmologist thought it was a microaneurysm or not. For the training set there are 1627 candidates, of which 368 are true microaneurysms. The ophthalmologist also identified another 26 microaneurysms, which the match-filter/region-growing algorithms never found. Therefore, there is a total of 394 true microaneurysms in the training set.

A medical physicist analysed the feature database to find relationships between features that may help to distinguish microaneurysms from other spurious objects. The primary tool used was to plot scatter graphs of microaneurysms and spurious objects with one feature on the y -axis and the other on the x -axis. From these graphs a number of rules were established that were useful for distinguishing microaneurysms from other objects. These rules were implemented in the microaneurysm detection software, so that it could run entirely unaided. Two computerised non-parametric classification methods (linear discriminant analysis and learning vector quantiser – a type of neural network) were also tried on the feature database.⁷ They gave solutions almost instantaneously, but the classifiers obtained

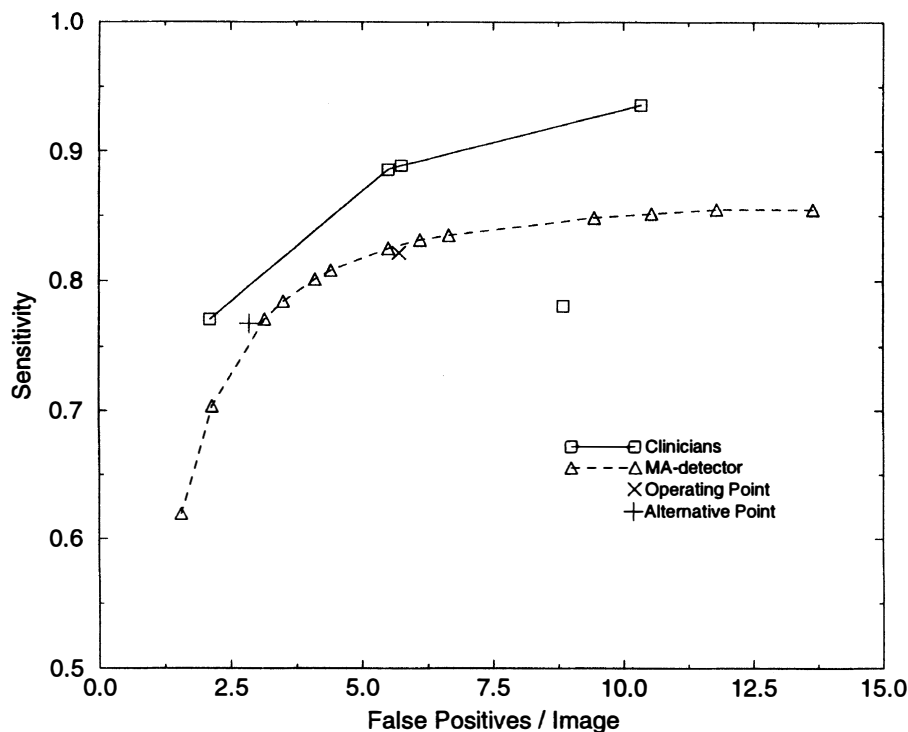


Fig. 1. Free-response receiver-operator-characteristic (ROC) curve for clinicians and the microaneurysm (MA) detector. One clinician has a significantly different opinion from the others and was ignored for the purposes of drawing an ROC curve. The operating point is the recommended point for the microaneurysm detector reported in Results. The alternative operating point is described in the Discussion.

were not as good as was achieved with the manual method described above.

A total of eight intensity features were used in the classifier. Partial component analysis of the correlation matrix of the eight intensity features showed that 96% of the information could be described by four orthogonal linear-combinations of the eight intensity features. However, it was found easier when manually deriving the classifier to work with the original eight features and they were retained for that reason.

To test the microaneurysm detector, another 20 images (each of a different patient who was not used for the training set; the mean frame time was 25.5 s with a range 17–48 s) were used to generate a database of candidate feature data in a manner analogous to the formation of the training set. In this case, however, a medical physicist as well as an ophthalmologist, both experienced with viewing fluorescein images on computer screen, independently identified microaneurysms in the test images. Where the two agreed, this was taken to be authoritative. Where they disagreed over the identification of an object, they reviewed the image involved again, and came to a decision jointly as to whether the candidate was a microaneurysm or not. The joint decision was therefore used as the 'gold standard' for testing purposes. There are 297 true microaneurysms in the gold standard. The micro-

aneurysm detector was compared with the gold standard and figures of sensitivity and false-positives per image were calculated in the normal way. To obtain a full receiver-operator-characteristic (ROC) curve⁸ the classifier rules were varied alternately towards microaneurysms and spurious objects. In this manner, the sensitivity of the microaneurysm detector can be varied against the false-positive rate. Because some of the rules can be varied continuously, a number of discrete points had to be chosen to make the problem manageable. Thus the ROC curve may not be optimal, but is believed to be nearly optimal.

As a comparison, five clinicians were asked to identify microaneurysms within the 20 test images. This was performed from the computer screen rather than from negatives or photographic prints, so the clinicians were working with the same data as the computer. Direct digital fluorescein imaging is becoming clinically acceptable and our system, although digitising from negatives, achieves a similar resolution as commercially available high-resolution digital fundus cameras.

RESULTS

Against the training set, the automated microaneurysm detector achieved 82% sensitivity with 2.0 false-positives per image. Against the test set, it achieved 82% sensitivity with 5.7 false-positives per

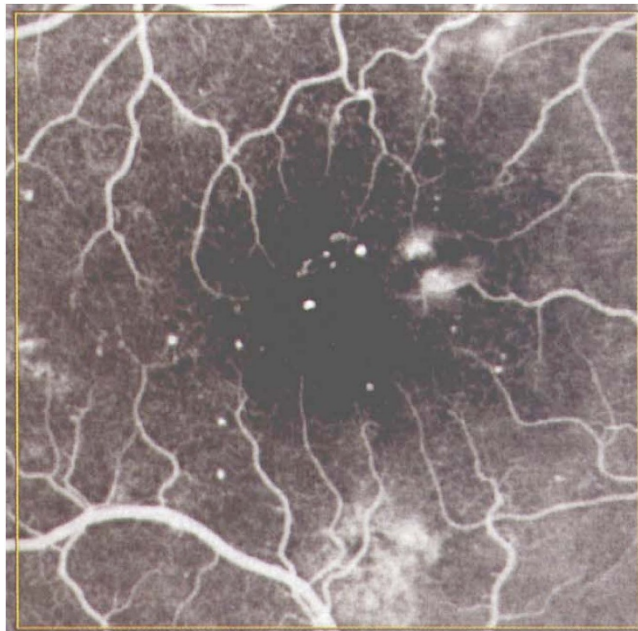


Fig. 2. The baseline fluorescein angiographic image of a diabetic patient. The frame time is 22 s.



Fig. 3. The fluorescein taken 12 months after the baseline image of the same diabetic patient (Fig. 2). The frame time is 15 s.

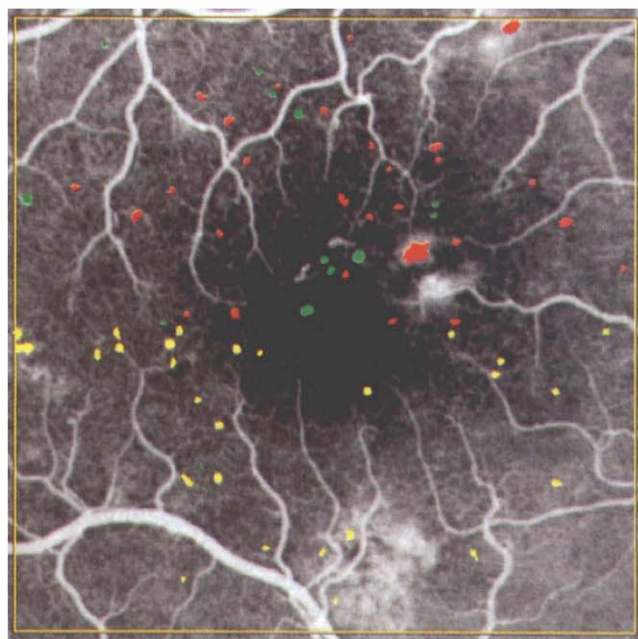


Fig. 4. The microaneurysm classifier in operation (when configured to the standard operating point). All candidates for the baseline fluorescein angiographic image (see Fig. 2) are shown. Those in red have been classified as spurious objects, those in green as microaneurysms, and those in yellow are candidates that have not been classified as yet.

image. An alternative operating point (see Discussion below) achieved 77% sensitivity with 2.9 false-positives per image against the test set. The free-response ROC curves derived from the clinician operating points and from the microaneur-

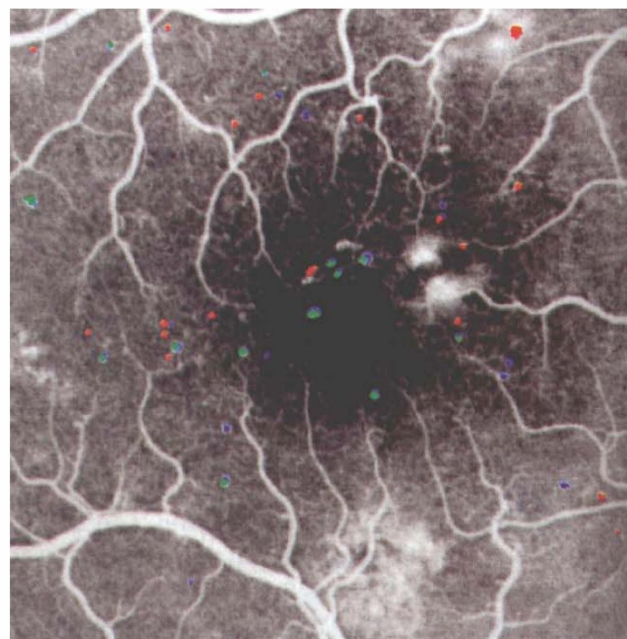


Fig. 5. Illustration of microaneurysm turnover. The fluorescein image shown was taken at baseline (as in Fig. 2). Microaneurysms in blue are present in the baseline image. Microaneurysms shown in green are present in the image taken 12 months later (Fig. 3) and can be matched to a (blue) microaneurysm present in the baseline image. The microaneurysms shown in red are present only in the fluorescein image taken at 12 months.

ysm detector are shown in Fig. 1. An example run of the microaneurysm detector (when configured at the standard operating point) on a fluorescein angiogram of a diabetic patient (Fig. 2) is shown in Fig. 4.

DISCUSSION

The microaneurysm detector and clinicians have been tested against a derived 'gold standard'. The microaneurysm detector performs almost to the same standard as the clinicians and, therefore, is robust enough for everyday clinical use. One clinician had a different opinion of what constitutes a microaneurysm and this further illustrates the difficulty of getting objective results and ensuring that clinicians are trained to the same standard. We ignored that particular clinician for the purposes of generating an ROC curve. However, the existence of such a point suggests that the clinician ROC presented may be a little optimistic. Note that the operating point of the microaneurysm detector (unlike clinicians) can be shifted to be more sensitive (or more specific) at the expense of some specificity (sensitivity) if so desired. A more sensitive operating point may be more suitable for primary screening purposes, for example. The operating point reported here, achieves good sensitivity, with very few false-positives, for images that are of reasonable quality and contain little to mild retinopathy. The same operating point, however, is unsatisfactory for use with severe retinopathy; it reports many false-positives in such cases. Because of this, we have implemented in our microaneurysm detector an alternative operating point that has been designed to handle a wide range of image qualities and types. By types, it is meant, for example, that some fundi are tigroid, others show little or no retinopathy while others contains a variety of microvasculature abnormalities. This alternative operating point is also shown in Fig. 1.

It is worth noting that the images in the test set were chosen to ensure that certain types of images were present. That is, it was ensured that tigroid images, images with little or no retinopathy, and images with various forms of more serious retinopathy were all represented in the test set. The reason for such a choice of images was to ensure that the microaneurysm detector worked for a broad spectrum of images, and to present various 'difficult' cases that would test the operation of the microaneurysm detector in what may be more unusual circumstances. This approach had its rewards; it gave us a better understanding of the operation of the microaneurysm detector and suggested ways it might be improved. However, the test set may not be fairly representative of the general diabetic population, and definitely not so of patients with mild retinopathy only. For these reasons, we believe that the test set probably represents a more difficult set of images than normally encountered in practice.

The computer microaneurysm detection system is automated at all stages of analysis to the extent that human intervention is rarely required. Such pro-

cesses as registering of same-eye images taken at different patient visits and identification of the fovea on a fluorescein image are automated. Only for some images, for example when laser photocoagulation is performed during the study, is human intervention required to aid the registration process. This is because the laser burns present in the later image tend to dominate the general structure of the image and 'fool' the automated registration process. In any case, the failure of the registration process is always easily identifiable and can be corrected with human intervention.

The 512×512 ROI for microaneurysm counting in this study was chosen partly because of hardware constraints encountered early in software development. These constraints have since been surmounted and any size ROI (up to the captured image size) may be used for microaneurysm counting. For the purposes of this study, which was to validate the automated microaneurysm detector, the choice of ROI was relatively unimportant, as long as the ROI was large enough to have a reasonable chance of encompassing some microaneurysms.

The microaneurysm detector could also be implemented on current digital fundus camera control computers. The software is fast enough to analyse an early fluorescein frame, and give a result by the time the late frame is ready to be taken. The advantages are speed of operation and the ability to report back to clinicians preliminary results while patients are still at the clinic. If the image quality is found to be too poor for processing, then images can be re-captured while the patient is still at the clinic.

Conventional manual methods for monitoring the progression of retinopathy can be tedious and are potentially error-prone. For such studies, the automated microaneurysm detector offers clear advantages. All the processing, including registration and macular identification, is fully automated and offers greatly increased speed without a significant loss of sensitivity over what is achievable by clinicians. Furthermore the method is fully repeatable and, unlike clinicians, always gives the same result for the same image. The automated microaneurysm detector makes large pharmaceutical trials particularly manageable.

As a recent innovation the microaneurysm detector has been extended to compare the microaneurysm results, for serial studies, across images. By doing this, it is not limited to reporting microaneurysm counts only, but is capable of reporting newly developed and recently regressed microaneurysms as well. Such information may be as significant as microaneurysm counts themselves.⁹ A comparison of two fluorescein images taken a year apart is shown in Fig. 5; see Figs. 2 and 3 for the two fluorescein angiograms analysed. The result, for the

most part, appears reasonable; however, there are instances where microaneurysms have been detected in the 12-month image and not in the baseline image, despite little change in the underlying feature. Closer inspection reveals that these instances are due to one of two causes. The first is that a patch of the baseline image is slightly out of focus so causing a microaneurysm to be missed. The second reason is that the leakage surrounding a potential microaneurysm has changed over the 12-month period so that the possible microaneurysm missed in the baseline image was found in the 12-month one. It is noted that in any case these examples involve questionable microaneurysm identifications and clinicians themselves may disagree over their interpretation. It is also worth noting that the microaneurysm detector classifier was configured at the 'operating point' (see Fig. 1) for the purposes of producing Fig. 5. Since calculating microaneurysm turnover is essentially differentiation, and it is well known that this differentiation tends to emphasise noise at the expense of signal, it may be advantageous to use a more specific operating point, such as the 'alternative point' shown in Fig. 1, to achieve a form of noise suppression. Such issues need more investigation. We ultimately plan to use the microaneurysm detector to further study microaneurysm turnover in diabetic retinopathy.

The financial assistance of Scotia Pharmaceuticals is gratefully acknowledged.

Key words: Microaneurysm, Diabetes, Retinopathy, Computer, Digitisation.

REFERENCES

1. Kohner EM, Sleightholm M, and KROC Collaborative Study Group. Does microaneurysm count reflect severity of early diabetic retinopathy? *Ophthalmology* 1986;93:586-9.
2. Klein R, Meuer SM, Moss SE, Klein BEK. The relationship of retinal microaneurysm counts to the 4-year progression of diabetic retinopathy. *Arch Ophthalmol* 1989;107:1780-5.
3. Lay B, Baudoin C, Klein JC. Automatic detection of microaneurysms in retinopathy fluoro-angiogram. *Proc SPIE* 1983;432:165-73.
4. Spencer T, Phillips RP, Sharp PF, Forrester JV. Automated detection and quantification of microaneurysms in fluorescein angiograms. *Graefes Arch Clin Exp Ophthalmol* 1992;230:36-41.
5. Spencer T, Olson JA, McHardy KC, Sharp PF, Forrester JV. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. *Comp Biomed Res* 1996;29:284-302.
6. Cideciyan AV, Jacobson SG, Kemp CM, Knighton RW, Nagel JH. Registration of high resolution images of the retina. *Proc SPIE* 1992;1652:310-22.
7. Frame AJ, Cree MJ, Olson JA, Undrill PE, McHardy KC, Sharp PF, Forrester JV. Computer-based classification of retinal microaneurysms. In: *Proceedings of the 2nd International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, Plymouth, UK, 1996:50-6.
8. Metz CE. ROC methodology in radiologic imaging. *Invest Radiol* 1986;21:720-33.
9. Kohner EM, Dollery CT. The rate of formation and disappearance of microaneurysms in diabetic retinopathy. *Eur J Clin Invest* 1970;1:167-71.