Assessment of colour vision as a screening test for sight threatening diabetic retinopathy before loss of vision

G L Ong, L G Ripley, R S B Newsom, A G Casswell

Aim: To assess the effects of sight threatening diabetic retinopathy (STDR) on colour vision and to evaluate automated tritan contrast threshold (TCT) testing for STDR screening before significant visual loss.

Method: Patients were recruited from a hospital based photographic screening clinic. All subjects underwent best corrected Snellen visual acuity (BCVA) and those with 20/30 vision or worse were excluded. Automated TCT was performed with a computer controlled, cathode ray tube based technique. The system produced a series of sinusoidal, standardised equiluminant chromatic gratings along a tritan confusion axis. Grading of diabetic retinopathy was made by one of the team of experienced ophthalmic registrars (SpR) using slit lamp biomicroscopy and a 78D lens; HbA1c, and urine albumin were also tested.

Results: Patients with STDR had significantly worse TCT despite normal BCVA (p<0.0001). TCT yielded a sensitivity of 100% for detecting diabetic maculopathy and 94% for STDR with a specificity of 95%. Logistic regression analyses showed that TCT (p<0.001) and HbA1c (p<0.05) correlated significantly with the presence of STDR but duration of diabetes, urine albumin counts, and BCVA failed to show any significant correlation. No associations between TCT and duration of disease, TCT and HbA1c, and TCT and urine albumin counts were found.

Conclusion: Tritan colour vision deficiency was observed in patients with STDR despite their normal BCVA. These results indicate that automated TCT assessment is an effective and clinically viable technique for detecting STDR, particularly diabetic maculopathy, before visual loss.

Despite effective treatments, diabetic retinopathy is still the most common cause of blindness in industrialised nations among the 20–60 year old age group.1

The detection of presymptomatic sight threatening diabetic retinopathy (STDR) remains difficult. Early treatment of proliferative diabetic retinopathy and diabetic maculopathy improves visual outcome. However, STDR should be detected before visual damage has taken place as only a minority of patients have improvement in vision following laser treatment. With effective screening blind registrations for patients under the age of 70 could be reduced by 10%.1 Screening is clinically viable and cost effective and will be increasingly important as the incidence of diabetes rises over the next 10 years.1

There are considerable variations in screening methods in the developed world including examination by ophthalmologists, ophthalmic opticians,17 diabetes physicians, and general practitioners,17 and by conventional fundus or digital colour fundus photography,18–20 and fluorescein angiography.17 The gold standard for assessing macular thickness has been stereo-macular colour or fluorescein photography with a 30° fundus camera with assessment by a trained grader but neither of these is practical as a screening tool.18 19 Recent development of higher resolution scans such as optical coherence tomography (OCT) and the retinal thickness analyser (RTA) have demonstrated considerable thickening of the diabetic retina before the onset of diabetic retinopathy.15 16 However, the high cost of these instruments have made them prohibitively expensive for screening.

Owing to inadequacies of current screening programmes, many patients with diabetes do not receive treatment before developing severe visual loss.27 Consequently, there is a need for screening that is inexpensive, and yet yields high sensitivity and specificity for the detection of sight threatening diabetic retinopathy before visual loss.

Colour vision testing provides a sensitive, non-invasive method to assess macula damage. Deterioration in colour vision often precedes changes in other clinical measures such as visual acuity and morphological changes. Several studies have shown a correlation between tritan colour vision deficiency and stage of diabetic retinopathy.19

However, no traditional colour vision test has demonstrated adequate efficacy to warrant widespread use for screening. Cathode ray tube (CRT) based colour vision tests have been shown to have significant advantages over traditional tests of colour vision.19

We therefore evaluated the automated tritan contrast threshold (TCT) test to screen for sight threatening diabetic retinopathy.

Methods

A total of 510 consenting diabetic patients were prospectively recruited from a photographic screening clinic over a 2 year period. Medical details of each patient were recorded including: duration of diabetes, age, type of treatment, and history of eye treatment. Diabetic control was assessed with the HbA1c percentage test. In addition, patients' early morning urine samples were also collected for urinary albumin counts.

The patients underwent a TCT test after a short demonstration. The TCT test was carried out on both eyes monocularly. Best corrected visual acuity (BCVA) was measured using the Snellen chart followed by pupillary dilatation with tropicamide 1%. Patients were then examined by one of the team of experienced ophthalmic registrars (SpR) using a slit lamp biomicroscope with a 78 dioptre Volk lens to provide a gold standard. The results of the TCT test and retinopathy status for each patient were obtained in a masked and independent manner.

See end of article for authors' affiliations

Correspondence to: Anthony G Casswell, Sussex Eye Hospital, Eastern Road, Brighton BN25BF, UK; tony.casswell@bsuh.nhs.uk

Accepted for publication 3 October 2002

D
Classification of diabetic retinopathy
The stage of retinopathy was graded using the European grading protocol. In this study, eyes classified as having pre-proliferative retinopathy, proliferative retinopathy, or maculopothy were considered as having sight threatening diabetic retinopathy (STDR). Eyes classified as having no retinopathy or background retinopathy were considered as having non-sight threatening diabetic retinopathy (NSTDR).

TCT procedure
The TCT was measured using a computerised, CRT based system developed by the biomedical engineering division of the University of Sussex. The system consists of a computer unit with a custom made plug-in card stimulus generator, a hand held response module and a 15 inch non-interlaced high resolution (SVGA) colour monitor (Qume: QM857) with a surround to provide a matching background for the monitor screen.

The system generates a series of vertical, sinusoidal, low spatial frequency (0.66 cycle per degree) and standardised equiluminant chromatic gratings on the display monitor, which has a uniform background luminance of 20 cd/m². The chromaticity of the gratings is changed under computer control along a tritan confusion axis. The chromatic contrast of the grating at which the subject can just distinguish the gratings from the background (TCT) is determined using a modified double staircase reversal algorithm as described by Cornsweet. The test procedures have been more fully described elsewhere.

In essence, the lower the TCT score, the more severe is the tritan contrast deficit.

Exclusion criteria
The exclusion criteria for this project included a corrected visual acuity of worse than 6/9, previous history of photocoagulation therapy, history of previous eye disease known to affect colour vision such as glaucoma, and signs of significant media opacification as determined by slit lamp examination through a dilated pupil. Furthermore, in clinical practice, around 1.5% of subjects are unable to compete the test satisfactorily.

Matching controls for lens optical density
Studies have demonstrated that diabetes duration correlates strongly with increases in lens optical density, even among patients with relatively short diabetes duration. Using psychophysical techniques, it has been shown that patients with diabetes suffer an increased rate of lens yellowing similar to that in non-diabetics older than 60 years. Consequently, any tritan deficit seen in diabetics may be wholly or partly due to the preretinal absorption of short wavelength light resulting from lens yellowing.

Moreland has derived two simple equations which relate Lutze and Bresnicks’s linear lens yellowing model with Pokorny et al’s linear model of normal lens ‘ageing.’ Using Moreland’s “lens equated” formulas, we derived the lens equated age of each diabetic and the appropriate lens equated control data were then selected from our laboratory’s control database (n = 310; mean age 48.0 (SD 19.1) years; range 16–87).

Standardisation of diabetic TCT data
Standardised scores were based on data obtained from a randomly selected eye of each of 310 control subjects. The data for these normal subjects were stratified by decade and the means and standard deviations for each decade age group were determined. The TCT from each diabetic eye was compared to the appropriate normal decade-age group matched according to lens yellowing.

Statistical analysis
One eye was randomly selected from each patient for analysis. The Snellen visual acuities were expressed as logarithmic values (logMAR) as proposed by Holladay and Prager. We used analysis of variance (ANOVA) and non-parametric tests (Mann-Whitney U and Kruskal-Wallis H) to analyse the differences between diabetics and “lens equated” controls and

Table 1

<table>
<thead>
<tr>
<th></th>
<th>STDR</th>
<th>NSTDR</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>17</td>
<td>493</td>
<td>&gt;0.5</td>
</tr>
<tr>
<td>Age (years)</td>
<td>60.4 (11.3)</td>
<td>60.9 (13.9)</td>
<td>0.43</td>
</tr>
<tr>
<td>Diabetes duration (years)</td>
<td>11.8 (6.9)</td>
<td>10.4 (8.6)</td>
<td>0.19</td>
</tr>
<tr>
<td>Urinary albumin counts (mg/l)</td>
<td>28.2 (28.7)</td>
<td>26 (47.6)</td>
<td>0.02*</td>
</tr>
<tr>
<td>Glycohaemoglobin (%)</td>
<td>9.8 (1.6)</td>
<td>8.1 (2.2)</td>
<td>0.0001**</td>
</tr>
<tr>
<td>Functional tests</td>
<td>BCVA (log)</td>
<td>0.1 (0.11)</td>
<td>0.06 (0.09)</td>
</tr>
<tr>
<td></td>
<td>TCT</td>
<td>24.7 (7.2)</td>
<td>42.3 (6.5)</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.001; STDR = sight threatening diabetic retinopathy; NSTDR = non-sight threatening diabetic retinopathy; BCVA = best corrected visual acuity; TCT = tritan contrast threshold.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>No retinopathy</th>
<th>Background retinopathy</th>
<th>Preproliferative retinopathy</th>
<th>Proliferative retinopathy</th>
<th>Maculopathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>383</td>
<td>110</td>
<td>3</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>TCT</td>
<td>42.5 (6.3)</td>
<td>41.7 (7.1)</td>
<td>29.6 (8.5)</td>
<td>21.7 (3.3)</td>
<td>24.0 (7.2)</td>
</tr>
<tr>
<td>BCVA (log)</td>
<td>0.06 (0.09)</td>
<td>0.07 (0.09)</td>
<td>0.06 (0.10)</td>
<td>0.09 (0.13)</td>
<td>0.11 (0.11)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>61.4 (13.9)</td>
<td>59.0 (13.9)</td>
<td>61.7 (13.1)</td>
<td>45 (18.4)</td>
<td>62.9 (8.8)</td>
</tr>
</tbody>
</table>

BCVA = best corrected visual acuity; TCT = tritan contrast threshold.

www.bjophthalmol.com
between patients with STDR and patients with NSTDR. Regression and correlation analyses were also performed on patient data, particularly for significance of correlation between the presence of STDR and TCT, age, duration of diabetes, glycosylated haemoglobin (HbA₁c), and urine albumin counts. Values of p < 0.05 were considered to be statistically significant.

Weighted kappa functions derived from the weighted kappa coefficient of association, κ, were used to assess the discriminability and to determine the optimal decision criterion of the TCT test to detect STDR. κ, which is weighted according to the costs associated with the potential outcomes, is defined as follows:

\[
κ_r = \frac{(n_p n_n - n_{pn} n_{np})}{r(n_p + n_n)(n_{fp} + n_{fn}) + r'(n_{fp} + n_{fn}) n_{tp} + r_n n_{fn}}
\]  

where \(r (0 \leq r \leq 1)\) is the weighting factor, \(r_0 = 1 - r\), \(n_p\) is the number with a true positive outcome, \(n_n\) is the number with a true negative outcome, \(n_{fp}\) is the number with a false positive outcome, and \(n_{fn}\) is the number with a false negative outcome. The application of weighted kappa functions to evaluate the discriminability and to determine the optimal decision criterion of vision measures has been described in detail elsewhere.³⁵

Using the optimal decision criterion as the pass/fail criterion for the TCT test, the screening result (pass or fail) was then correlated with the reference standard—that is, the criterion for the TCT test, the screening result (pass or fail). Using the Mann-Whitney U test we have also found that patients with STDR had significantly worse TCT and HbA₁c compared to those of the patients with NSTDR (TCT: \(p<0.0001\); HbA₁c: \(p = 0.02\), see Table 1). Though patients with STDR also had poorer BCVA, longer duration of diabetes and higher urine albumin counts compared to those of the NSTDR group, no significant difference was established (log BCVA: \(p = 0.13\), duration of diabetes: \(p = 0.43\), urine albumin counts: \(p = 0.19\); Mann-Whitney U). Logistic regression analyses showed that TCT (\(p<0.001\)) and HbA₁c (\(p = 0.018\)) correlated significantly with the presence of STDR but duration of diabetes, urine albumin counts and log BCVA failed to show any significant correlation.

Using the Mann-Whitney U test we have also found that patients with STDR have significantly abnormal TCTs when compared to “lens equated” controls (STDR: \(p<0.0001\)). No significant TCT differences were found between diabetic eyes with NSTDR and lens equated controls (NSTDR: \(p >0.1\)).

### RESULTS

Of the 510 patients seen, 107 were categorised as having type I insulin dependent diabetes and 403 as having type II non-insulin dependent diabetes. The mean age of the patients was 60.8 years (range 16–94 years) with a mean disease duration of 10.4 years (range 1.5–50 years). Table 1 summarises the details of all the patients, with and without STDR. The biomicroscopic assessments showed that, out of the 510 eyes examined, 383 eyes (75%) and 110 eyes (21.6%) were found to have no retinopathy and background retinopathy, respectively. The remaining 17 eyes (3.4%) were categorised as having STDR (three with pre-proliferative retinopathy, two with proliferative retinopathy, and 12 with maculopathy). The mean and standard deviation of the TCT according to the grade of retinopathy are summarised in Table 2. A one way analysis of variance (ANOVA) established that statistically significant differences existed between some of the diabetic retinopathy categories. The Scheffe multiple comparisons test showed that the differences are predominantly between the STDR data and NSTDR data. No statistically significant difference in TCT was found for diabetic eyes between the subgroups in the NSTDR category (\(p = 0.23\)) or in the STDR category (\(p = 0.42\)).

A Pearson correlation analysis of the relation between TCT, age, duration of diabetes, HbA₁c, and urine albumin counts was also carried out. Significant correlation was found between age and TCT (\(p <0.0001\)), as well as age and duration of diabetes (\(p <0.0001\)) and age and HbA₁c (\(p <0.001\)). None of the other variables showed significant correlation with TCT (HbA₁c: \(p >0.4\); urinary albumin counts: \(p >0.1\); duration of diabetes; \(p >0.8\)).

The Colour vision as a screening test for diabetic retinopathy 749

![](image)

**Figure 1** Weighted kappa functions for \(κ_r\) (solid triangle), \(κ_{0.9}\) (diamond), \(κ_{0.5}\) (solid circle) and \(κ_0\) (square) for a range of z scores, standardised TCT measure.
shown in Figure 1. The peak of each curve indicates the optimum value of z score to be chosen as the pass-fail criterion and the height of the peak indicates the discriminability.

In selecting the value of \( r \), other authors have opted to use \( r = 0.5 \), thus assuming the cost of a decision was equal for those with and without disease. However, in a diabetic screening setting, the costs of false negative decisions are significantly more expensive. Consequently, the value of \( r \) should be closer to 1. In our experience, an appropriate value of \( r \) for a diabetic screening setting is \( r = 0.9 \). In this study, using \( r = 0.9 \), we have found that the optimal decision criterion was \( z = B1.75 \) which gives \( \kappa = 0.81 \) (see Fig 1). With this optimal pass-fail criterion, the TCT screening results achieved a good sensitivity and specificity of 94% (95% CI: 73 to 100) and 95% (95% CI: 92 to 96) respectively, as shown in Table 3. TCT test results also correlated highly significantly with the presence of STDR (p < 0.0001, \( \chi^2 \) test).

Table 4 shows the comparison of TCT and ophthalmological assessment for all patients when the optimum pass-fail criterion of \( z = -1.75 \) was used. The results show that TCT successfully flagged all eyes with diabetic maculopathy giving a sensitivity of 100% despite good visual acuity. However, TCT missed one (pre-proliferative) out of five (three pre-proliferative and two proliferative) eyes with sight threatening retinopathy giving a sensitivity of 80%. A total of 26 out of 493 eyes with NSTDR failed the TCT test, of which 18 eyes were graded with no retinopathy and eight eyes with background retinopathy by ophthalmic examination.

### DISCUSSION

The case for developing a screening programme for diabetic retinopathy has been widely accepted. Javitt et al\(^{16} \) showed that screening and treatment for eye disease in patients with type II diabetes is both clinically and cost effective. Studies in the United Kingdom have demonstrated similar results.\(^ {17} \) Despite this many diabetics are not receiving optimal eye care.

Ideally, all diabetics should have their eyes examined annually by ophthalmologists. However, in many countries, this is impractical. In most countries, and certainly in the United Kingdom, ophthalmologists do not have the time to participate in screening programmes and it is too costly to train and to pay other personnel such as opticians or non-ophthalmic doctors to conduct such programmes. Various screening techniques have been proposed but none has proved to be ideal. Furthermore, studies have shown that the performance of any screening method can be disappointingly low, depending on the experience of the examiners.\(^ {18} \)

Various colour vision assessments, such as FM100, anomaloscope and others, have shown correlation between colour vision deficits and diabetic retinopathy. However, most conventional colour discrimination tests, such as FM 100 hue and Farnsworth-Munsell 100 hue D-15 tests, are inadequate and also not sensitive enough for widespread screening purposes. For example, in 1985, Green et al\(^ {19} \) examined the FM 100 hue test as a screening device for STDR. Although they found the test to have a sensitivity of 73% and a specificity of 66%, they concluded that the test was not sensitive enough for the detection of severe retinopathy. In a similar study, Bresnick et al reported a sensitivity of 65% and a specificity of 59% and concluded that the FM 100 hue test was cumbersome to administer and score in an office setting. In view of this, could a more user friendly and sensitive colour vision test improve the sensitivity and specificity of the screening procedure? Recently, a study by Maár et al reported a sensitivity and specificity of 88.9% and 93.3% respectively in detecting diabetic macular oedema using the Mollon-Reffin “Minimalist” test. However, the study only examined a relatively small number of subjects with macular oedema. In addition, the study did not assess the performance of the test to detect other STDR such as pre-proliferative and proliferative retinopathy.

For a screening programme to be cost effective, the costs of early detection of a target disease need to be balanced with the costs of late detection. These costs determine whether screening for the disease is economically viable. In a diabetic screening setting, the costs of early detection involve the screening examination, patient investigation, treatment and re-examination of false positive patients. These costs need to be balanced against the costs of late detection which include the human and economic consequences of secondary complications of the disease, the loss of potential earnings and the provision of social/housing services that may be incurred.

In this study, using the weighted kappa coefficient of association, \( \kappa \), we evaluated the discriminability, and the optimum pass-fail criterion, of an automated TCT test in screening for STDR. As shown in equation (2), \( \kappa \) is weighted according to the costs associated with the potential outcomes of a screening programme. A choice of \( r = 0.5 \) would assume that the costs of early and late detection of a target disease are equal.

However, as mentioned above, the costs of late detection in a diabetic screening setting are much higher than the costs of early detection. In view of this, the aim of a diabetic screening programme should be to detect early all (or most) diabetics with STDR in order to avoid the high cost associated with false negative decisions. Therefore, to determine the optimum pass-fail criterion of the TCT test to detect STDR, the value of \( r \) should be close to 1. From our experience, an appropriate value of \( r \), in a screening setting, is about 0.9. This is supported by the results of this study which show that the optimum pass-fail criterion for screening for STDR based on kappa function \( \kappa_z \), produced better sensitivity and discriminability.
than the optimum criterion based on kappa function \( \kappa \). For kappa function \( \kappa \), the optimum pass/fail criterion was a \( z \) score of \(-2.75 \) which yielded a discriminability, sensitivity, and specificity of 58%, 65% and 98% respectively. On the other hand, for kappa function \( \kappa \), the optimum pass/fail criterion (\( z \) score = \(-1.75 \)) yielded a discriminability of 81%, sensitivity of 94% (95% CI: 69 to 100), and specificity of 95% (95% CI: 92 to 97) indicating that TCT is a robust discriminator for screening for STDR.\(^{30}\)

In this study, using an automated CRT based technique, we have shown the efficacy of using the TCT test in screening for STDR. Our results also show that the best performance of the TCT test in our screening population gives a sensitivity of 94% (95% CI: 73 to 100) and a specificity of 95% (95% CI: 92 to 96). Our results compare favourably with the standards of any screening programme established by the British Diabetic Association (Diabetes UK) of at least 80% sensitivity and 95% specificity.\(^{31}\) But larger numbers will be required to confirm this. Using various colour vision assessment techniques, the relation between different aspects of diabetic retinopathy and colour vision has been studied extensively. To the best of our knowledge, based on a literature search through Medline, this is the largest reported study analysing the use of an automated TCT test in screening for STDR in a diabetic screening programme.

Studies have shown that a proportion of diabetic patients who are suffering considerable visual loss could have been identified earlier in the course of their disease by visual acuity measurements,\(^{32}\) but that does not yield sufficient sensitivity to provide clinical information as to the impact of altered retinal function in early stages of diabetic eye disease.\(^{41}\) In addition, maculopathy can be present for some time before affecting Snellen acuity. In our study, all patients with STDR had corrected Snellen acuity ranging from 6/5 to 6/9. Furthermore, patients with STDR ideally should be identified before acuity has been affected. The success of the TCT test in achieving an overall sensitivity of 94% in detecting diabetic patients with STDR (100% sensitivity in detection of maculopathy and proliferative retinopathy), despite their normal Snellen acuity, makes it a viable technique for early detection of STDR among the diabetic population.

Consistently with other studies, our results showed that patients with STDR experienced tritan colour dysfunction. However, no statistically significant difference was found between patients with NSTDR and lens equated control subjects (\( p > 0.1 \)). The results reported here on the TCT of diabetic patients clearly indicate that diabetics with STDR have significantly worse TCTs (\( p < 0.0001 \)) when compared to those with NSTDR. It is known that microvascular complication increases more frequent with increasing diabetes duration\(^{42}\) and high glycosylated haemoglobin values.\(^{43}\) However, our results, similarly to previous studies, did not show any association between either TCT deficit and duration of diabetes, or the latest glycosylated haemoglobin result.\(^{44}\)

CONCLUSION

Diabetic retinopathy is a sight threatening disease where early and effective treatment has been shown to reduce significantly the incidence of blindness. This study has demonstrated the efficacy of an automated TCT test to detect STDR, especially diabetic maculopathy, before visual loss. In addition, since it measures visual function (rather than features associated with visual loss—for example, retinal morphology), it is possible that this test will also identify early those subjects who will progress to develop more severe retinal disease.

ACKNOWLEDGEMENT

We would like to thank Mr Osama Kamel, FRCOphth, and Mr Ben Moate, FRCOphth, for their contributions to the study.
40 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1997;33:159–74.