Reviewer’s report

Title: The ChromaTest, a Digital Color Contrast Sensitivity Analyzer, for Diabetic Maculopathy

Version: 2 Date: 15 February 2008

Reviewer: Keith A Goatman

Reviewer’s report:

Major Compulsory Revisions

1. The inclusion of the control data is a useful addition. However, why call it unpublished data? Why not just include it in this paper (together with the necessary information about subjects including number, age, exclusion criteria etc.). The plots in figures 1 - 3 are a useful addition too.

2. I am still concerned about the Wilcoxon test comparing PCCT in NPDR and CSME and TCCT in NPDR and CSME. Using the data from the tables in the original manuscript (which appears to be correct as the summary statistics I calculate from it agree with those in the paper) BOTH PCCT and TCCT appear to be significantly different between the two groups:

   PCCT (p = 0.0091)
   TCCT (p < 0.001)

   i.e. both p-values are less than 0.05.

   I get exactly the same result using MATLAB statistics toolbox (version 6.1) and the online stats calculator at http://www.fon.hum.uva.nl/Service/Statistics/Wilcoxon_Test.html

   I could not find a Wilcoxon test in standard Excel (XP version) as suggested in the paper to try.

3. I am also still concerned about the thresholds used to calculate the sensitivity and specificity. Although the revised manuscript now mentions the multiple thresholds for different age ranges, it still does not say HOW the thresholds were determined for each age range or what values were used. Clearly dividing a relatively small dataset into several smaller parts and then using them for both training and testing is a very weak performance test.

   The justification for multiple thresholds is the age dependancy of PCCT and TCCT. Performing a regression analysis on the data from the original manuscript there is a statistically significant positive slope for both TCCT and PCCT with increasing age. It's difficult to tell from fig. 1 if this is also the case for controls also. If the variable thresholds were based on estimates of the slope, for instance, this may provide stronger justification for the multiple thresholds (which
could then be applied as a linear, rather than piecewise, threshold).

Minor Essential Revisions

Stray full-stop after "left." on page 4.

Discretionary Revisions

I prefer "seven field" to "7 field".

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I still have no competing interests