Author's response to reviews

Title: Computer-aided assessment of diagnostic images for epidemiological research

Authors:

   Alison G Abraham (agump@jhsph.edu)
   Donald D Duncan (donald.duncan@bme.ogi.edu)
   Stephen J Gange (sgange@jhsph.edu)
   Sheila West (shwest@jhmi.edu)

Version: 4 Date: 29 September 2009

Author's response to reviews: see over
September 25, 2009

BMC Medical Research Methodology

Dear Editor:

We were happy to receive on August 25th the response to our revised manuscript entitled “Computer-aided assessment of diagnostic images for epidemiological research”. While we sought in our revised manuscript to carefully address all the concerns of the reviewers, it was clear from the lingering comments that some additional refinement was necessary. The abstract, in particular, has been greatly refined based upon the insightful comments. We have supplied our responses below to the remaining concerns and recommendations of the reviewers (our responses in red). Some minor additional content was added to the manuscript to increase clarity. We have detailed in the responses the location of changes that were made to the manuscript so the appropriateness of the changes could be assessed.

Thank you once again for providing us with the valuable review of this work and for this opportunity to share our results with the readers of BMC Medical Research Methodology.

Sincerely,

Alison Abraham, PhD (Corresponding Author)
Johns Hopkins Bloomberg School of Public Health
Department of Epidemiology
615 N. Wolfe Street, Room E-7640
Baltimore, MD 21205 USA
Phone: 410-502-9763
Email: agump@jhsph.edu
Comments for Authors:

Reviewer 1:
Discretionary Revisions:
P.4 l.28 "imdilate and imerode functions";
Their functions correspond to dilation and erosion operations in Image Processing Toolbox of MATLAB, do they? A referee will think that it is better to revise a manuscript if authors are possible.

Yes those functions are included with the MATLAB Image processing toolbox. We have specified this fact in the methods (Image processing subsection).

Reviewer 2:
This manuscript has considerably improved with respect to a previously reviewed version.

We appreciate very much the recognition of the work and thought we devoted to the revised manuscript as a result of the careful review. We are grateful for the opportunity to improve the manuscript and credit the insight of the reviewers.

I think the authors are doing themselves a disservice, however, with the uninspiring abstract (discretionary revision). The abstract fails to clearly state the purpose and relevance of this research. The 'background' and 'conclusions' are especially weak since they are too general and don't even seem to necessarily relate to the work presented in the paper and definitely don't generate any excitement for the reader. This manuscript deserves a better abstract inviting readers to read further.

We are grateful to the reviewer for bringing the uninspiring quality of the abstract to our attention. As the reviewer astutely points out, the abstract should convey the essence of the paper and we have tried to revise the wording to more appropriately convey the relevance of the work.

In the methods section the authors explain image segmentation to obtain the pupil area of interest. However, this paper does not appear to use this segmentation since the pupil boundary is known. Inclusion in the Methods section in its current form is hence misleading (minor essential revision).

The reviewer is correct that the segmentation portion of the algorithm was not needed (or consequently put to the test) for this study. However, the algorithm was designed for use with real data, and thus a necessary component is a segmentation step. The paper is meant, in part, to be a didactic for the design of a CAD system, illustrating that readily available software and methods can be used for a complete design. Thus we feel the full design is appropriate to include. To avoid any opportunity for the inclusion to be misleading, we have added a sentence prior to the algorithm methodology description to explicitly state what aspects of the algorithm were used in the validation (Methods section, 2nd paragraph) and a sentence to the validation subsection to restate the same (Validation subsection, 2nd paragraph).
The description of the use of a cutoff value to distinguish between normal and abnormal for use in ROC analysis is unclear (minor essential revision). The authors state that a cutoff value for the severity score of 3 was used to determine the area under the ROC curves. It is unclear in the current form whether this cutoff was applied to the known 'truth' severity score to divide the images into 2 'truth' categories normal/abnormal or (incorrectly) to the scores given by the reviewers/CAD algorithm.

The ROC was performed using the threshold score of 3/16ths to define cortical cataract based on the true cataract severity. Thus the true status, i.e. cataract or no cataract, for each image was known. Scores assigned by the CAD system and the reviewers were dichotomized based on a moving threshold to obtain the ROC curves from which an area was calculated.

Also, why was the threshold value set at 3? The reported areas under the curve are extremely high (around 0.98), which is unusual but would also imply that the reviewers really don't need any help from a CAD algorithm. Please explain.

The value of 3/16ths describes a value of cataract severity at which many experience a noted impact on vision. It has been used as a threshold for defining cataract in publications from the Salisbury Eye Evaluation Study (e.g West et al., JAMA, 1998), which uses the cortical cataract grading scale we drew on for the present analysis. We have noted this in the Results section (1st paragraph). While it is true that a CAD algorithm would seem unnecessary for an analysis using a dichotomous outcome of cataract or no cataract, we believe this speaks to the issue of data coarsening, which can dilute signals that may be of etiologic import. Thus a CAD system that can capture the true continuous nature of the disease process is a boon to researchers seeking to elucidate risk factors from multifactorial mechanisms. Comparing the continuous scale of the CAD system to a standard cortical cataract grading scale, which yields an ordinal categorical variable from 0 to 16, illustrates the resulting bias that is a function of both scale and human limitations. We have added statements to the discussion to emphasize this benefit of the CAD system (Discussion Section, 1st paragraph).