Title: Development and preliminary evaluation of the VPS ReplaySuite: a virtual double-headed microscope for pathology

Authors:
- Dan DJ Johnston (daniel.johnston2@mail.dcu.ie)
- Sean SP Costello (sean.costello@slidepath.com)
- Peter PA Dervan (pdervan@mater.ie)
- Daniel DG O'Shea (Donal.OShea@dcu.ie)

Version: 2 Date: 7 March 2005

Author's response to reviews: see over
To the editor,

Please find attached the redrafted manuscript of “The VPS ReplaySuite: A virtual double-headed microscope pathology training tool” [MS: 2376928415662474]. We have addressed all the points made by peer reviewer Dr. Rebecca Crowley, as illustrated below.

Abstract:
Page 2-3 - Results section. It seems very unfair to only report opinions of those 3 participants who replayed more than 10 examinations. Either include the results for the entire set of participants or remove this entirely.

Now reads “77.7% (7/9) of all participants, and all 3 participants who replayed more than 10 examinations stated the ReplaySuite to be of some or great benefit in pathology training and quality assurance.”

Page 3 - Conclusions section. Change the last sentence. For a study of this kind, the sample set is simply too small to know whether it bodes well or not. Remove this claim. Simply state that the sample set was small and potentially biased. Indicate the need or intention for future work to validate and generalize these findings

Removed reference to future bodes well for the technology. Now reads “While the evaluation set was limited and potentially subject to bias, the response of participants was strongly favourable. Further work is planned to determine whether use of the ReplaySuite can result in improved diagnostic ability.”

Background:
Page 4 - Novices often make errors when searching the slide. Use the following reference which contains the original data:


Reference changed

Methods:
Page 8 - Diagnostic Concordance Graphs. Indicate how these are constructed by the system? You seem to have different categories (B2, B5, etc), which seem to come from The Core Reporting Guidelines. Do they only exist for breast? Is this scalable to other domains by simple changes in your system?

Sentence added: “The ReplaySuite generates graphs by setting the dimensions of each bar according to the percentage of participants it represents, and can represent alternate scoring regimens by making simple change to the system.”


Now reads “Preliminary evaluation of the ReplaySuite“

Page 9 Study architecture. Retitle this Study procedure
Retitled to study procedure

Page 9 - In the last line of the design paragraph - Participants who had previously participated. Do the authors mean that these individuals were not required to re-diagnose the cases, but did have an opportunity to view the examinations of others? If so, please clarify.

Paragraph now reads “Participants who had previously participated in the VPS validation study (by examining cases) were not required to re-examine cases, and could review any examinations.”

Results:
Page 12 If User 5 was biased and you know this from the start- I would just remove that person from the beginning and say that you had 9 participants, rather than 10. Sentences referring to user 5 have been deleted, and the participant group is referred to as only having 9 members. Table 1 & 2 have been modified to reflect this change.

Discussion:
Page 15 It seems very questionable to single out the change in diagnoses by individuals, and try to interpret the potential for generalization. I would remove this paragraph entirely, or talk more generally about the kinds of re-assessments that are possible.

Sentence removed “The subsequent group concordant re-diagnosis after replaying an examination (screen resolution of 1024x768 pixels) illustrates the benefit of observing the diagnostic trace used during the examination.”

Sentences added “It cannot be suggested, based on these individual examples, that group concordant re-diagnosis subsequent to ReplaySuite use will be the rule rather than the exception. However, these examples are worth noting, as it highlights that using the system can result in diagnostic re-evaluation.”

Page 18 The paragraph describing the relationship between use of the system and perceived benefit does not contribute much to the paper. I would suggest removing this paragraph as well as Figures 8 and 9.

This paragraph remains in the discussion, but in a reduced format, as we feel it is important to highlight the fact that participants who used the system the most thought it to be of the greatest potential benefit in pathology training and QA. Figures 8 and 9 remain as they are important for displaying the distribution of answers to 3 important questions.

Section 1 removed: “Figure 9 illustrates that, the more participants used the ReplaySuite to replay examinations, the greater potential benefit they perceived the ReplaySuite as having in pathology training. This trend is again replicated in Figure 10;”

Replaced with: “Figures 9 & 10 illustrate that the more participants used the system, the greater potential benefit they perceived it having in pathology training and quality assurance.”
Section 2 removed: “User 60, who did not replay any examinations and considered the ReplaySuite of little benefit to both training and QA work, may illustrate this point, stating “I think am not a reliable test as I did have a very limited exercise[sic].”

Figures: Figures, especially screen shots, are of poor quality. Almost impossible to see any detail. Capture figures at higher resolution using SnagIt or CaptureEase or similar software, if this hasn’t already been done. All figures and screenshots have been re-captured, using SnagIt.