Reviewer's report

Title: Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum

Version: 1 Date: 31 May 2010

Reviewer: Nick Stone

Reviewer's report:

This article is well written and explores a key clinical need in utilising a form of spectroscopy to provide molecular based information on lung disease, via sputum production.

This article is novel and deserves to be published. The mix of controls used is interesting and useful to the study.

The identification of spectral overlap between glycogen and glycoproteins is well discussed.

Major compulsory revisions

I have one key concern on the use of non-gold standard diagnosis: 'clinical diagnosis'. 9/25 patients within the lung cancer were labelled in this way. If this is necessary and histopathological decision making is impossible then this needs to be justified.

Minor Essential Revisions

Raman spectroscopy is a complimentary technique which has also been used in lung cancer diagnostics and sputum analysis (Nick Magee). Reference to a key paper or two would help provide balance to the introduction.

Results line 2/3: Cells from sputum were studied not whole sputum samples, so the differences observed are in the cells and not the sputum.

Use of median spectra is unusual rather than mean spectra. Is this point by point medians across the dataset? If so please explain this.

Some discussion regarding identification of subgroups using pca does not seem to be particularly useful, considering the peaks entered into the PCA have been selected specifically for statistical significance for separation of normals from cancers. Other key data likely to separate other groups and provide additional information has been removed... Explore further or remove this kind of discussion.

Discretionary Revisions

'Raw spectra were baseline corrected' How?
'P value' would normally be 'p value'

Why are the wavenumber axes plotted high to low in the figures.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests