Reviewer's report

Title: PAM50 Breast Cancer Subtyping by RT-qPCR and Concordance with Standard Clinical Molecular Markers

Version: 1 Date: 2 May 2012

Reviewer: Thomas Hughes

Reviewer's report:

In this article the authors examine the use of a newly developed assay that allows determination of expression levels of the PAM50 gene set in breast cancers in order to classify the tumours. They characterise and optimise the performance of their assay on a training set of tumours and then analyse its ability to classify tumours, as compared to standard pathology techniques, using a large tumour cohort derived from a clinical trial. Overall, I find the study to be potentially worthwhile, although I have some doubts about whether there are any novel scientific findings beyond the validation of the new assay itself. However, I find considerable problems with the article including a lack of detail on key aspects of what was done and why, lack of focus on the specific findings that the authors wish to highlight, and some over-statements or even exaggerations. Specific points are listed below.

Major points.

1) Abstract: A key finding seems to be “RT-qPCR scoring… provided more prognostic information than histological-molecular scoring”. I think this claim relies solely on Figure 6. I find this figure to be rather difficult to interpret since hormone receptor status is not broken down into ER and PR separately and because of the confounding factor that the cohort includes patients who received a variety of therapies (the two different arms of the chemotherapy trial; also with or without hormonal therapies – a fact not mentioned in this manuscript anywhere). The hormone therapy is particularly problematic since this was presumably given on the basis of ER expression as detected by IHC – and therefore is specific for some groups shown on the right, but is probably variable in all groups on the left. Could the authors justify this conclusion more thoroughly in the relevant results section?

2) Throughout the manuscript there is a difficulty in articulating how the accuracy of one method of determining cancer subtypes is being assessed. It is important to note that when assigning a tumour to a subtype there is not necessarily a “right answer”. All that can be done is one method can be compared to another method – but it is not really possible to say which is right. For example, in the abstract, “IHC/CISH biomarkers were not able to accurately identify all intrinsic subtypes”. It is true that the results from IHC differ from those of qPCR – but it is not fair to say that the qPCR must be correct and IHC has failed. This difficulty is embedded within the writing – especially within the “Subtype,
immunohistochemistry and RT-qPCR…” section of the results.

3) Introduction, 2nd para. It is simply not true that the PAM50 gene set has become the gold standard for subtyping (at least certainly not in the UK, and I doubt in many other countries). Nor is it “the standard method of classification” as claimed in the Discussion.

4) Methods, top page 9. How were IHC scores dichotomised into positive and negative?

5) I don’t feel the NKI analyses, which are barely mentioned in the results text, contribute enough to justify their inclusion.

6) Results, “Training set” and Figure 1. Could the authors more fully explain why non-neoplastic tissue tissue should be included in a training set for analysis of tumours? Could the authors please highlight the 16 non-neoplastic samples on Figure 1 – are these within the right hand group of 17 samples, therefore presumably one tumour has grouped with the non-neoplastics? Or are they elsewhere? In the text the green highlighted samples are described as “normal-like” (a tumour classification) whereas I think they are probably actually non-neoplastics (as suggested by the figure legend). The phrase “The accuracy of the RT-qPCR training set in correctly assigning tumor subtype was 93% as determined by comparing to our previously reported research training set that used microarray data” is an example of my point 2 – surely the authors simply mean concordance, not accuracy or correctly. I’m not completely sure of the relevance of the statement anyway as the data are not shown.

7) Results “Interference from normal…”. Which non-neoplastic sample was used (ie which on Figure 1) and did it matter which with respect to the change in classification seen? Which tumour subtype RNA were used? “The switch from Luminal B to Luminal A required less than 50% contribution from the normal breast tissue signature”, this is directly contradicted in the Discussion. A very surprising result is seen in Additional file 6 – where the esr1 call changes from intermediate to low and back as additional non-neoplastic RNA is added to the HER2-E sample; please would the authors comment.

8) Results “Validation of single and…”. How was it decided that medians were appropriate cut offs for some markers and quartiles for others? To what is high/intermediate/low in red/yellow/blue for proliferation and positive/negative in red/blue for “luminal” referring in additional file 7? Could the authors explain why these plots “validate” the cut offs (as in the title of this section), rather than merely demonstrating their effect?

9) Results “Subtype, immunohistochemistry…” and Figure 3. Could the authors explain more fully and clearly what the ROC curves demonstrate. They say that “Intermediate and high gene scores were combined to dichotomise the continuous gene expression data” – giving the impression that they have used the lower cut offs in Table 2. But I believe they have used the ROC curves to test different cut offs for the qPCR data basing these sensitivity_specificity analyses.
on the positives defined by IHC/ISH. These analyses would certainly be appropriate – but I’ve had to guess that this is what was done. Most worryingly, the graphs seem to demonstrate that the best cut offs have not been selected – particularly in graph C, where the cut off has inflated the false positive rate for no gain in true positive rate.

The rest of this section is very difficult to read for at least two reasons. A - it is consistently unclear exactly what the authors mean by each description of a patient group – for example, “Luminal A tumours more frequently…”. Do the authors mean luminal A as defined by their PAM50 classification – or by IHC staining? This needs to be very clear since different groupings are defined by the two technologies and since the findings here are counter-intuitive with some tumours found to be negative for the classic markers usually associated with their group. B – there is simply too much detail and too few take home messages. What are the main important points the authors wish to make?

10) Discussion, page 17. “We used a novel approach for selecting… cut offs”. I am not clear that using either median or quartile cut offs is novel or was “shown to be reproducible in an independent data set”. Could the authors explain this more fully?

11) Discussion, page 18. I believe the discussion of transcription factors should be omitted since these data are not a focus of the manuscript.

12) Conclusions. The surrogate subtyping is barely mentioned in the results text yet is highlighted as a key conclusion – please could the authors either reconsider this conclusion or more fully explain the data. The final sentence is without justification and should be removed.

Minor points
1) The abstract is filled with jargon and abbreviations that are likely to be uninterpretable to many readers. For example: what is CLIA, what does “central scoring” mean, what is IHC/CISH?
2) Methods, page 10. “due to a lower concentration of a single gene within a sample”. Please change to gene product, or transcript or something similar.
3) Please increase the size of the dendrogram relative to the heat map in Figure 1.

Level of interest: An article whose findings are important to those with closely related research interests

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 interest