Reviewer's report

Title: Validation and simplification of microarray-based grading in breast cancer

Version: 1 Date: 29 March 2011

Reviewer: Mauro Delorenzi

Reviewer's report:

The authors modify a proposed method for using gene expression microarrays to determine the grade of a breast cancer.

The aim of the present work is to avoid using histological grade to calibrate a gene expression index (GGI), essentially by devising a parameter-free classification into low (G1) or high (G3) grade.

Their score is the difference between the average rank of the expression level for a set of genes that are high in G3 and low in G1 minus the average rank of a set of genes with the opposite trend.

They use only genes determined previously and a training set to eliminate some of the G3-high genes so that the split when the difference is 0 (average G3 higher than average G1 genes to call a high grade) is a cutoff that performs well.

The authors show that the newly defined grade (GGR):

a) is in good agreement on independent data with histological grade 1 and 3

b) splits histological grade 2 so that its G1 and G3 components have different survival in the independent datasets, significantly where the sample size is large to have good power.

The authors regard the result as a more readily acceptable and/or more correct demonstration of the efficacy of the principle of using gene expression for grading than those presented by other authors previously, as no histological grade information is used and the method classifies a single expression profile without needing or depending on a batch.
Discussion of weak points:

Minor Essential Revisions

We think that the loss of performance in the NKI data visible in grade calling and in a possibly smaller separation of G1 and G3 in the survival curve Fig. 1E, points to the sensitivity of this system on the particular probes that represent a gene or other technical platform-dependent factors.

Better platform-independence was a reason why for calibration of GGI histological grade was used in Sotiriou et al. and seems a disadvantage for the proposed method; a calibration that - at least in theory - has to be done only once after which the GGI can also be used on single arrays and quite simply. (also normalization by RMA can used a fixed quantile distribution).

The results presented in Fig. 1 would be more informative if the authors would compare them to those that are obtained, if the GGI is used, as one would like to see that survival prediction is of comparable or higher quality.

While "insensitivity to composition" of a cohort (p.4) can be an advantage for testing the ability of the GGR to reproduce histological grade 1 and 3, for clinical applications of grading one would like a method that is thoroughly prospectively tested on the platform used, calibrated and compared to histological grade. Then, it seems less clear, if the proposed method still represents an improvement.

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 am a named inventor of the GGI-based method to grade tumors based on gene expression.