Author's response to reviews

Title: Translating a gene expression signature for Multiple Myeloma prognosis into a robust high-throughput assay for clinical use.

Authors:

Ryan K van Laar (rvanlaar@signalgenetics.com)
Rachel K Flinchum (rflinchum@signalgenetics.com)
Nathan S Brown (nbrown@signalgenetics.com)
Joseph Ramsey (jramsey@signalgenetics.com)
Sam Riccitelli (sriccitelli@signalgenetics.com)
Christoph Heuck (CJHeuck@uams.edu)
Bart Barlogie (BarlogieBart@uams.edu)
John D Shaughnessy (jShaughnessy@signalgenetics.com)

Version: 4 Date: 26 April 2014

Author's response to reviews: see over
Author's response to editorial requests:

Information about the ethics committee who approved the study is included.
Acknowledgements section has been added.

Referee 1: Reviewer's report

Version: 2 Date: 7 April 2014 Reviewer: Zhengdeng Lei

I. Major Compulsory Revisions:

(1) Because this study was based on the 17 peer-reviewed publications described in Table 1, the authors should point out which studies (especially large cohort) clearly validate GEP70.

All publications listed in Table 1 used the GEP70 algorithm to stratify patients as either high or low risk for disease relapse and/or overall survival. There are many more publications which refer to GEP70 in discussion but where the algorithm has not been applied to any GEP data – these were not included in the table. We also did not include studies in which GEP data from previous studies were re analyzed (eg meta-analyses, such as MAQC Consortium[1]).

(2) In Figure 4a and 4b, please show us the plot and R^2 without using log10 transformation of x-axis, as did in 4c and 4d.

Figure 4a and 4b are shown on linear scale below. Log10 scale is used in the manuscript to better visualize the spread of these (very left-skewed) data.
We feel it is more informative to view these data on a log scale as the vast majority of specimens we receive are <10% pre-sort CD138%, but were interested see if a trend was evident when the entire spectrum was analyzed.

(3) Page 7, "For each MyPRS® specimen analyzed, the 70 gene expression values used to compute the patients risk score are combined with a matrix of 70-gene data from 559 patients used to originally train and validate the prognostic algorithm", please show us how to combine the value and how to deal with batch effect.

To generate a personalized gene expression data heat map for each patient analyzed, the following steps are performed by our ResultsPX analysis software:

i. Load the matrix of the MAS5 normalized gene expression data for the 70 genes by 559 patients.
ii. Attach the 70 gene data from the patient being analyzed to the matrix, making it 70 genes by 560 patients.
iii. Order 560 columns of the matrix (1 column = 1 patient) left-to-right by increasing GEP70 score.

We do not perform any ‘batch effect’ modification during this procedure. The 70 gene by 559 patient dataset was that used to originally develop the GEP70 algorithm. These data were produced by the UAMS Myeloma Institute for Research and Therapy laboratory and were used in the validation studies performed to ensure assay produces statistically equivalent results when performed in our clinical laboratory. As such we have verified that there is no batch effect in data generated between our two laboratories, with regard to the 70 gene profile in use.

II. Minor Essential Revisions:
(1) Page 2: "Over a 12 month period, the 70-gene prognosis score (range 0-100) has a standard deviation of 2.72 and a variance of 0.03", please highlight that this is based on cell line (positive control), and it is "coefficient of variance", not "variance".

This section now reads: “Over a 12 month period, the 70-gene prognosis score (range 0-100) of our multiple myeloma cell-line control sample had a standard deviation of 2.72 and a coefficient of variance of 0.03. “

III. Minor Essential Revisions:

(1) Page 3: "In this paper we describe the use of a high-throughput process, combining cell isolation, flow cytometry and gene expression profiling to provide physicians with personalized prognostic assessments multiple myeloma", missing "of" after "assessments".

The word ‘of’ has been inserted before “assessments”.

(2) Page 6: "Figure 4a and 3b" => "Figure 4a and 4b".

Corrected.

(3) Page 6: missing a period in the last sentence.

Corrected.

References: