Reviewer's report

Title: Gene-expression patterns in peripheral blood classify familial breast cancer susceptibility

Version: 2 Date: 27 July 2015

Reviewer: Lisa Hines

Reviewer's report:

This work explores the ability to use gene-expression levels in peripheral blood mononuclear cells as a means of breast cancer risk prediction among individuals who have a family history of breast cancer (FBC). In this study, they utilized two study populations: a training set of 124 women (63 controls and 61 cases) from Utah and a validation set of 73 women (37 controls and 36 cases) from Ontario, Canada. These data were additionally stratified according to BRCA1 and BRCA2 mutation status.

Major:

1. One of the biggest limitations with this study is that it is retrospective – blood samples were obtained after diagnosis and treatment. While these breast cancer patients may have completed their course of treatment, it is plausible that the differences that are being observed are still a consequence of treatment regimen. While the inclusion of women with sporadic breast cancer does provide some control of this confounding variable, tumor pathology and/or treatment regimens for women with a family history could have been more aggressive than sporadic cases. Were these groups comparable with respect to clinical/pathological characteristics and treatment regimen? The manuscript should provide summary tables that reflect risk factors and prognostic factors among the groups being compared for both the test and validation sets. The current Tables 1 and 2 provide no information regarding these datasets, other than sample size and median age. Supplemental Tables 1 and 2 do not provide summary data based according to the groups being compared. There is no data for Ontario population. Was the Ontario population similar with respect to these factors when compared to the Utah population? With respect to Table 2, it appears that these comparison groups were not age-matched? This is another limitation. It appears that, overall, there were differences among these groups. The groups without breast cancer tended to be younger, so we cannot assume that they will not develop breast cancer by the time they reach the average age for the cases.

2. Clarification and corrections are needed in the Results section, as well as Tables and Figures. There are many inconsistencies with respect to Figures and Tables (including Supplementary Tables) that are being referenced in the Results section, and Figure legends are not consistent with what is being illustrated in the Figure. Additional clarification and interpretation should be made with respect to what is being depicted in the Figure, i.e. define genomic model score, do the lines depict 95% confidence intervals, are their differences according to BRCA
status (Figures 2C and 2D), etc. The Results Section could provide more clarity when interpreting data results. For example, it states that similar levels of accuracy were obtained for women with and without BRCA, but Figures 2C and 2D reflect differences that are not discussed. Was any comparison done to see if the gene-expression profile was any better at predicting risk when compared with existing models for estimating risk based solely on personal health history and demographics? This would be an important point to address.

3. The Discussion section should address the limitations of this study. The sample size is actually quite small once subdivided into the different groups, with 23 being the largest subgroup. The Discussion states that the sample size was large enough to obtain statistically meaningful results, yet the data illustrated in Figure 1 appear to reflect very large confidence intervals (assuming these do in fact depict 95% CIs). This small retrospective study provides interesting suggestive data, but prospective studies are imperative to assess potential as a risk prediction tool.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests