Reviewer's report

Title: A Kallikrein 15 (KLK15) single nucleotide polymorphism located close to a novel exon is associated with poor ovarian cancer survival

Version: 2 Date: 15 November 2010

Reviewer: kathryn terry

Reviewer's report:

This is a study of polymorphisms in the KLK15 gene identified through a variety of in silico tools to identify likely functional polymorphisms. In this effort the authors identified a novel exon likely involved in alternative splicing. 9 SNPs were selected to tag 22 likely functional SNPs and genotyped in 319 ovarian cancer cases from a population based case control study and Royal Brisbane Hospital in Australia. The most promising SNP (rs266851) in the Australian data was genotyped in 1815 cases from multiple UK studies as well as 413 cases from the TCGA pilot project. The SNP was not significantly associated with ovarian cancer survival in these validation data sets.

Strengths of this study includes the authors exhaustive attention to in silico modeling to identify the most promising SNPs in the KLK15 gene, control for age, stage, histologic subtype, and grade in the analysis, and genotyping in 3 independent data sets. However, there are several limitations of the study that need to be addressed.

Major Compulsory Revisions

1. Although the initial findings in the Australian data set are interesting for rs266851 the genotyping in the UK and TCGA data sets show no significant associations. Table 1 shows hazard ratio estimates from the UK GWAS (HR=1.07, 95CI% 0.94-1.24) and TCGA (HR=1.24, 95% CI=0.90-1.61) consistent with a null association and have confidence intervals that span 1. Granted these estimates go in the same direction as the original finding but could not be considered a “validation” of the original finding. With that said, the authors may want to point out that there could be differences in these populations that may have lead to different results including differences in histologic grade/subtype distributions, longer time to case ascertainment leading to a survival bias (that is, cases may be too sick to participate or have died before enrollment if the time between diagnosis and enrollment in the UK and TCGA data sets is longer than the Australian data set then a survival advantage for this snp could be missed entirely in the UK and TCGA data sets). However, with the data presented here that is impossible to assess. Given these differences are not explored, and the UK and TCGA data sets are much larger than the original (and therefore better powered to detect a true association), the current data leads to a conclusion of no association. Consequently, the main conclusion of the paper and the title of the manuscript need to be revised.
2. More detailed information about case ascertainment, including time between diagnosis and enrollment, should be added to the methods. Also, how were the 207 cases from the population based study selected? Presumably the original study includes far more than 207 cases. Similarly, how were the hospital cases identified... were these consecutive incident cases? If not, how were they selected?

Minor Essential Revisions

3. The authors should include a discussion of the limitations of the study including selection of the cases, lack of validation, and possibility of survival bias.

4. Residual disease is an important predictor of survival. If this data is available it should be added to the statistical analysis, if not this needs to be noted in the limitations.

Discretionary Revisions

5. The statistical analysis description indicates that survival analyses were censored at Sept 1, 2004. Presumably, this date coincides with the end of follow up. It would be useful to update this analysis with more recent survival data if possible.

6. In the section describing the sequencing of cancer cell ines and aggressive cancer patients, were the authors looking at germline or somatic mutations in the aggressive cases?

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

**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.