Reviewer's report

Title: PIK3R1 underexpression is an independent prognostic marker in breast cancer

Version: 1 Date: 5 July 2013

Reviewer: Josh Lauring

Reviewer's report:

This is a technically well done study testing the general hypothesis that alterations in the PI3-kinase pathway may have prognostic value in breast cancer. There have been many such studies examining the prognostic impact of PIK3CA mutations, PTEN expression, AKT phosphorylation and other biomarkers of pathway activity. Conflicting results have been demonstrated regarding the prognostic value of PIK3CA mutations, but the largest studies appear to show a favorable prognostic impact, even when adjusting for the fact that the mutations occur most frequently in the more favorable ER+ subset. These investigators have previously published an analysis of the same 450+ breast cancer patients analyzed for PIK3CA exon 9 and 20 hotspot mutation status and prognosis, finding that mutations were a favorable factor on univariate analysis, but not on multivariate analysis.

In this manuscript they re-analyze PIK3CA mutations, adding in rarer mutations in exons 1 and 2, as well as mutations in AKT1 and PIK3R1, which are found at about 3% frequency each (in contrast to the 30-40% PIK3CA mutation frequency. They also analyze RT-PCR based mRNA expression level for a number of PI3-kinase pathway related genes. The study focuses on the degree of association between various mutations and expression levels and then examines the prognostic utility of the analyzed mutations or individual gene expression levels for metastasis free survival in the whole dataset and in individual breast cancer subtypes.

None of the findings regarding PIK3CA, AKT1, or PIK3R1 mutations is especially novel. There are now quite a number of large scale studies showing the relative frequency of these mutations, their mutual exclusivity in breast cancer, and their association with particular breast cancer subtypes (ER+ versus triple negative, for example). These authors have already published their analysis of PIK3CA hotspot mutations in this same cohort. The addition of 7 tumors with exon 1 or 2 mutations (the majority of which already have additional hotspot mutations) adds nothing to their previous findings. It could have been predicted that all of the other mutations analyzed were sufficiently rare that they would not be able to find any statistical association that would add to our current state of knowledge.

The gene expression-based analysis is a bit of a fishing exercise performing multiple comparisons to try to identify various significant associations. However, the focus on PIK3R1 expression adds something to the existing literature since
PIK3R1 expression has not previously been examined for its association with prognosis.

The suggestion that PIK3R1 expression could become a clinically useful independent prognostic marker in breast cancer overreaches a bit, although the investigators do note in their discussion that further studies are needed. Many investigators have identified individual genes or gene signatures which are prognostic in a given dataset of patients. To be a candidate for clinical utility, a marker needs to be replicated in an independent data set and studied prospectively in a population who are untreated or subject to uniform treatment. Furthermore, there is no evidence that PIK3R1 is as good or superior to other widely used clinical parameters with prognostic or predictive value. To be useful as a clinical test from a practical standpoint, IHC would be a better marker than RT-PCR. Although limited data showing correlation of RT-PCR expression and IHC of PIK3R1, the study’s conclusions cannot really be extended to protein IHC. However, these issues are for future research and do not necessarily detract from the relevance of these initial reported findings. If the finding of the prognostic relevance of PIK3R1 expression holds up, it would be of some interest to the breast cancer research community and worthy of publication.

Major compulsory revisions:

1. A major issue with this clinical dataset is the potential heterogeneity in terms of treatment. The patients in this series were treated over a 30 year period, during which there have been major changes in screening, surgery, radiation, and adjuvant systemic treatment recommendations for breast cancer. The prognostic role of PIK3R1 expression and PIK3CA mutation is assessed in terms of metastasis-free survival of the cohort. However, patients were not necessarily uniformly treated, and no data are presented showing the frequency of hormonal therapy or chemotherapy in the patients with normal expression or underexpression of PIK3R1, for example. Although 344 tumors were ER+, only 268 of them were treated with hormonal therapy, which would be considered sub-standard care by present-day standards. What if there is unequal representation of untreated tumors in the low PIK3R1 subset? The same issue applies to chemotherapy. 183 women received chemotherapy, although 339 were ER-negative. Given that only 26% of the tumors were node-negative, the rate of chemotherapy usage also seems low by today’s standards. Obviously one cannot retrospectively change practice patterns, and one has to work with the data one has, but the authors should explore the possible confounding of their results by differences in treatment and receipt of treatment in the different groups they are analyzing (by PIK3R1 expression level, PIK3CA mutation, etc). A prognostic biomarker cannot be analyzed in the absence of details regarding treatments received.

2. Another possible statistical issue that may apply to this analysis is the lack of correction for multiple comparisons when calculating the significance level of various associations. Many variables (expression level and mutation status of many genes) are analyzed for their associations with survival in the whole cohort and again in individual tumor subtypes. I believe a correction for such multiple
hypothesis testing is needed, but I will defer to a statistical expert.

3. There is no discussion of how the expression cut-off of N<0.5 or N>2.0 was chosen. While it might make biological sense to define as important an expression level relative to normal tissue of <0.5 or >2.0 times normal, it does not necessarily make sense for identification of the performance of a prognostic marker. In this study 61% of breast cancers “underexpress” PIK3R1. This is a high proportion of cancers for a putative adverse prognostic marker. While the choice of a cut-off is to some extent arbitrary, one wonders how such a prognostic marker would be useful clinically if it applies to 61% of patients. The utility of a prognostic marker depends on its ability to discriminate between better and worse outcomes. Many prognostic biomarker studies examine expression of the marker in the set of tumors and then separate the tumors into groups based on their relative expression level (highest to lowest quartile, etc.). Such an approach might better define what level of PIK3R1 expression confers the greatest risk of relapse than the arbitrary choice of N<0.5, which places 61% of tumors in the high-risk category. Any future validation of PIK3R1 expression as a prognostic marker would require going forward with the optimal cut-off expression value, so it bears improving upon in this pilot study. The authors could evaluate the performance of different cut-offs, or potentially examine the performance of their marker as a continuous variable.

4. p.12.-13. The experiments correlating IHC and RT-PCR results for PTEN and PIK3R1 are methodologically flawed. It seems correct that each individual marker must be compared (RT-PCR versus IHC) rather than lumping both proteins together in the analysis. This combined analysis covers over the fact that 2/13 tumors analyzed were discordant for PTEN. I also think it would be better if the authors could show multiple tumors with IHC for PIK3R1 to give the readers a better sense of how PIK3R1 staining intensity was scored and how well the staining correlated with the mRNA level. The criteria for 1+, 2+, and 3+ staining are not explicitly described in the methods. It would also be helpful to see the relative level of PIK3R1 protein expression in the tumor and adjacent normal tissue from the same patient, rather than a comparison of a tumor from one individual with normal tissue from another. The evidence for PIK3R1 underexpression as a relevant driver event in breast tumorigenesis would be stronger if most tumors showed lower expression than adjacent normal breast ducts.

5. Much of the results section is descriptive of associations that are not very meaningful without further analysis. For example, page 12 paragraph 3: “PIK3R1 underexpression was also associated with AKT3 and WEE1 underexpression.” Was there any hypothesis being tested here? Is this interesting in any way? Does this tell us anything about the biology of PIK3R1 underexpressing tumors? If it is worth discussing in the results at such length, there should be some discussion of these particular results in the Discussion section, and there is not, in most cases. Otherwise, the tables can show this less interesting data sufficiently well.
6. p. 13 “Survival analysis” paragraph. It is strange that mutation of either PIK3CA, PIK3R1, or AKT1 is associated with a worse prognosis, yet PIK3CA mutation considered alone is associated with a better prognosis. Given that there were 151 PIK3CA mutations and a combined 26 PIK3R1 or AKT1 mutation, this contradiction seems strange, particularly since AKT1 mutations are almost exclusively found in better prognosis ER+ tumors.

Minor essential revisions:

1. Page 3 Abstract. The conclusion of the abstract is misstated. Alterations in PIK3CA but not PIK3R1 were shown to be prognostic. Lower expression of PIK3R1 was shown to be prognostic in this study, but low expression is not an “alteration.”

2. Page 7, Mutation screening paragraph. In one line it says PIK3R1 exons 11-15 were screened and in the next line it says exons 10-14.

3. p.10-paragraph 2, the synonymous nucleotide change in PIK3R1 is discussed as a potential polymorphism, yet it is counted as one of 11 “mutations.” This “mutation” should be verified as a somatic change in the tumor by sequencing normal tissue from the same patient. This change is most likely a sequencing artifact or a germline polymorphism. If it is a somatic change, it is still of doubtful significance, as it does not change the amino acid sequence of the protein. I think it would be best to report the frequency of non-synonymous changes.

4. p.10 paragraph 2. The sentence beginning “PIK3R1 mutations were found in only 1 of the 151…” should state that PIK3R1 mutations are mutually exclusive with PIK3CA mutations, rather than AKT1 mutations.

5. p.12 The statement beginning “On the other hand. The authors speculate about what altered gene expression of pathway components “might” mean for pathway activity, but they should explicitly note that their data does not in fact provide any actual evidence about the relative level of PI3-kinase pathway activity in these tumors.

6. Table 1 and Table 2 duplicate a lot of data.

7. In p.13 “survival analysis” paragraph and Additional Figure 1, it is not clear what comparison the p value refers to, since there are four lines on the plot.

8. On page 17, In the sentence “PIK3R1 underexpressing tumors were also prone to…” the word “cumulate” should be “accumulate.” In the following sentence, the claim that PIK3R1 underexpression is associated with increased signaling activation is unsupported and hypothetical. The authors should state that it “could be associated with increased signaling activation.”

9. On page 18, paragraph 2, the authors should suggest that PIK3R1 should or could be explored as a predictive marker for trastuzumab resistance, etc., since there are currently no data to suggest that it can be used as such.
10. On page 18, paragraph 3. The sentence beginning “The present study…” Again, the authors provide no evidence supporting an effect of PIK3R1 underexpression on PI3-kinase pathway activation and should simply state that they found an effect on survival.

Discretionary revisions:

1. p. 11 last paragraph. “Marked overexpression” of several genes is discussed, but the numbers reported only show the percentage of tumors with N>2.0 overexpression. The range of values indicates that some tumors have very high expression, which might be termed “marked.” It would be more interesting to see what proportion of tumors have such “outlier” very high expression. Such tumors are more likely to have chromosomal amplification of the genes in question.

2. The data is all presented in tabular format. For ease of comprehension of the major findings, it might be good to also present the Table 2 data in graphical format, as box-and-whisker plots.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.