Author's response to reviews

Title: Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status

Authors:

Xi Zhao Ph.D. (xi.cameron@me.com)
Einar A. Rødland Ph.D. (einarro@ifi.uio.no)
Therese Sørlie Ph.D. (Therese.Sorlie@rr-research.no)
Hans Kristian M. Vollan M.D. (hans.kristian.vollan@me.com)
Hege G. Russnes M.D., Ph.D. (hegrus@rr-research.no)
Vessela N. Kristensen Ph.D. (v.n.kristensen@medisin.uio.no)
Ole Christian Lingjærde Ph.D. (ole@ifi.uio.no)
Anne-Lise Børresen-Dale Ph.D. (a.l.borresen-dale@medisin.uio.no)

Version: 5 Date: 7 January 2014

Author's response to reviews: see over
Palo Alto, January 5, 2014

Dear editors of the BMC Cancer:

We hereby submit the revised manuscript “Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status”. In this letter, please find our point-by-point response to referees' comments and explanations of how the manuscript has been changed.

A Sweave document is also provided together with the revised version of the manuscript. This allows other investigators to reproduce the analyses performed in our study.

We strongly believe that our study fits well with the scope of BMC Cancer and is well suited to the interest of your audiences.

Thanks for your consideration.

Sincerely,

Xi Zhao, Ph.D
Center for Cancer Systems Biology
Stanford University
School of Medicine
Response to comments on “Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status” (BMC Cancer: MS 8553457659452288)

Below, we list in details our responses to each of the reviewers’ comments (cited by the quotes) and the changes we have made to the manuscript.

Reviewer: Giampaolo Bianchini

“Major Compulsory Revisions
1) As also pointed out by both the two previous reviewers, “the observation that the performance of gene signatures depends on hormone receptor status is not new” and “is well-known for a long time”. A review on gene expression profiling in breast cancer (Reis-Filho JS, Pusztai L, Lancet 378:1812-1823) stated as follow: “Many gene expression prognostic signatures have been described during the past decade. First, despite differences in the genes that compose each of the signatures, they largely identify the same group of patients as having poor prognosis disease. The unifying characteristic is the high expression of proliferation-related genes. Meta-analyses of large cohorts of breast cancers subjected to gene expression profiling revealed that only the proliferation-related component of prognostic signatures can forecast the outcome of patients with breast cancer. Second, proliferation and the degree of expression of proliferation-related genes are one of the strongest prognostic factors in ER-positive cancers. Third, since the levels of expression of proliferation-related genes are usually high in ER-negative cancers, first generation prognostic signatures almost invariably classify ER-negative cancers as of poor prognosis….. Sixth, the accuracy of the outcome predictions of most prognostic signatures seems to be time-dependent, with more accurate predictions at 5 years than at 10 years after diagnosis.” The different biological processes associated with prognosis and chemotherapy response in ER-positive and ER-negative breast cancers have been deeply and purposely investigated (Iwamoto T J Natl Cancer Inst 103:264-72, 2011). Accordingly with this concept, nowadays, biomarker studies in breast cancer should investigate and develop biomarkers in a subgroup specific way (i.e. Rody A, Breast Cancer Res 11:R15, 2009; Coutant C, Clin Cancer Res 17:2591-601, 2011; Bianchini G, Cancer Res 70:8852-62, 2010; Ignatiadis M, Journal of Clinical Oncology 30:1996-2004, 2012; Santarpia L, Oncologist, 2013; and many others).

Therefore, all the assessments of the biomarkers in the unselected overall population should be removed from the manuscript because these analysis can be considered inappropriate, not informative and potentially misleading. All univariate and multivariate analysis as well as time dependent evaluations should be performed and presented only for the two ER group separately.”

All the analysis are performed and presented for the two ER groups separately.

2) The authors assessed the time dependent effect of the signatures separately in the untreated group of patients (Figure S3). For consistency, the same analysis for the adjuvant treated group should be presented. In fact, it is important to know if the amount of the time dependent effect is influenced by administered treatment or if it is retained in patient treated with adjuvant therapy.”
Results for the treated group have now been included.

“3) The METABRIC dataset has been used to confirm the time-varying effect of biomarkers. However, the outcome information available in the METABRIC dataset (breast cancer specific survival) is significantly different from distant event free survival available in the first cohort. In particular, the time from distant metastasis and death could be strongly influenced by treatment administered and molecular subgroup. This call for caution in interpreting the METABRIC data and should be properly discussed.”

The evaluation on METABRIC has been added for treated and untreated groups. Similar indications were observed in both the treated and untreated group.

“4) The observation that histologic grade influenced the time-dependent effect of biomarkers is interesting but it is not well supported. First, the analysis of time dependency should be performed in ER-positive and ER-negative separately as described in point 1. Second, because in this series the grade itself seemed to have a strong time dependent effect (Figure S4 B), the time varying effect of biomarker in the ER+ group should be assessed separately in grade 3 and grade 1-2 groups to assess if the time dependency is only masked by grade in the proportional hazard assessment.”

The analysis of time dependency is now performed in ER-positive and ER-negative separately. The effect remains in the ER-positive group (Table3 & Figure S2). However, we have further investigated the relation between the time-dependency and grade (or tumor size), and agree that the inclusion of grade (or size) in the models may have masked the time-dependent effect of the signature. We have adjusted the manuscript accordingly, and have also included tables with more detailed analyses. As tumor size is another parameter describing tumor progression, the indication that “In the context of long-term survival, incorporating characteristics for tumor aging may help to alleviate the time-dependency” from previous analysis still stand. In the ER-negative group, because the prognostic power for included clinical parameters and gene signatures are dismal to start with, we were not be able to uncover the nature of the time-dependency for this tumor group.

“Minor Essential Revisions
1) “In the context of long-term survival, incorporating characteristics for tumor aging may help to alleviate the time-dependency….” “Aging” should be replaced with “grading””

The analysis indicates that the time dependency of the gene signature can be captured by parameters describing the degree of tumor advancing such as histological grade (in previous analysis with both ER+ and ER- groups) and tumor size (for ER+ group). The generic term “aging” was intended to reflect this, but has been replaced with “advancement” which should be able to reflect either grade or tumor size. In addition, after revisiting the link between the time-dependency and grade (or size), we now refer to the inclusion of these clinical...
parameters as “might help improve the quality of the prognosis”, rather than claim that they can remove the time-dependency.

“2) “Additionally, immune-related gene modules have been implicated to be prognostic in high-risk ER-positive breast cancers [53] and ER-negative breast cancers [54, 55].” (Discussion pg 20) Reference 53 (Rody A, 2009) is incorrectly referenced. It swohed that an immune-module (T-cell) is associated with good prognosis in HER2-positive and TN groups.”

Rody et al. reported a positive prognostic value of lymphocyte-specific kinase (LCK) metagene for ER-negative tumors as well as for ER-positive tumors with HER2 overexpression (Figure 4d and Table 3 in Rody et al 2009). Hazard ratio associated with LCK metagene among the ER-positive tumors with HER2 overexpression is 4.17, comparing LCK low vs high, indicating high expression of the LCK metagene predicted for better disease-free survival in this high-risk ER-positive breast cancer group. This led us believe that immune-related gene modules have been implicated to be prognostic in high-risk ER-positive breast cancers.

“3) “Different normalization procedure from the original study [6] may explain why the 76-gene signature did not predict any good prognosis in our data” (pg 18) This sentence should be explained (the 76-gene signature is actually prognostic in both ER-positive and ER-negative).”

The 76-gene signature did not produce any good prognosis using the original cutoffs. Only by our proposed population-based cutoffs, this gene signature is able to be prognostic in both ER-positive and ER-negative. We have tried to make this formulation more clear in the manuscript.

“The title of table S5 is incorrect “Time- & ER-dependent effect assessment of individual gene signatures in predicting Distant Metastasis Free Survival (DMFS) on the METABRIC set.” The outcome used in the METABRIC is the breast cancer specific death”

The title has been corrected for Table S5 (now as Table S7).  

“4) The title of table S5 is incorrect “Time- & ER-dependent effect assessment of individual gene signatures in predicting Distant Metastasis Free Survival (DMFS) on the METABRIC set.” The outcome used in the METABRIC is the breast cancer specific death”

We included a discussion about previous relevant work, specifically for Desmedt et al. work. This work is the most relevant work compared to ours in term of biology for risk stratifications underlying various gene signatures. We now reference Albain KS, Lancet Oncol 11:55-65, 2010 for the time trend observation for 21-RS signature (Oncotype DX). Nielsen TO, Clin Cancer Res 16:5222-32, 2010 is a detailed work comparing PAM50 expression based signature vs IHC classifications in term of prognostic power. It has different focuses than our study.
Reviewer: Charles C Theillet

“Remarks and concerns
In the comparative analysis of the different signatures, the absence of the recent Guedj et al. (2012 Oncogene 1;31(9):1196) signature is regrettable. This signature has shown to segregate accurately ER+ from ER- and proliferative from non-proliferative subtypes. In this sense it could perform just as well as the GGI.”

We thank reviewer for pointing us to the interesting study. The Guedj et al. signature does not provide a prognostic score. And for our study, since this signature does not produce the same subtypes, we can’t use the same ROR model by incorporating four intrinsic subtypes (LumA, LumB, HER2-enriched and Basal-like) to predict risk scores on the studied datasets.

“The dependance of prognostic power to time could be undermined by multiple confounding factors. The test set is a compound dataset formed of 6 independent Affymetrix U133A datasets, with different aims, different treatment regimen and non-unified clinical follow up procedures. Furthermore, it is now a known fact that while Basal-like tumors are at elevated risk of early recurrence, the number of metastatic events drops dramatically after 5 years. In ER+ breast cancer the situation is more complex, since two groups can be identified, one at risk of rapid recurrence, the second one showing late (between 6 and 10 years) or very late recurrence (>10 years). These late recurrences may, thus correspond to different biological settings, explaining why the signatures that are mainly fitted for aggressive breast cancer (high vs. low proliferation, high vs low hypoxia) perform poorly for late or very late recurrence. Both these points should be discussed.”

First point, we have evaluated that the cohort/batch effect for pulling the data together. And the cohort effect is not statistically significant in term of survival prediction. Reference to Figure 1 and Box 5 of the Supplement had been added. We also tested and confirmed our hypothesis on METABRIC data, this eases the concerns about the validity of our study due to confounding factors arising from different studies.

Upon reinvestigating the link between the time-dependency of the signatures and grade (as well as tumor size) in greater detail, we found that some of our initial suggestions may have been somewhat hastily made, and have removed speculations that correction for tumor advancement (e.g. grade) could make the signature predictive of late relapse.

We have been a bit reluctant to go into discussions of why the prognostic effect falls with time. We have tried to speculate on it privately, but have found little in terms of empirical evidence to help us assess the various alternatives. A proper discussion, we feel, would warrant more space that there is room for.

“Minor points
Section "possible effects of cohort differences" the text states that cohort differences do not contribute significantly to the tested models, but no data are shown to support this claim. Please correct.”
Reference to Figure 1 and Box 5 of the Supplement has been added.

“The HER2 stratification came to no conclusion, which could be expected considering the work by Johan Staaf and colleagues in JCO (2010) 28, 1813, showing that HER2+ cancers do not form a homogeneous breast tumor subset and that major differences exist between HER2+/ER+ and HER2+/ER-. Hence, a double stratification could have been more informative.”

Double stratification has power issue. There was no event observed in the 5-10 time interval for the HER2+/ER- group; and no event in the >10 year time interval for the HER2+/ER+ group. Hence, the associated HR of individual signatures cannot be estimated for these two groups within the two time intervals respectively. For the HER2-/ER+ and HER2-/ER- groups where at least two events are presented for each time intervals, we observe a decreasing time trend for the signature’s prognostic power, and HER2-/ER+ is generally better than HER2-/ER- in term of prognostic power. This can be explained by the ER stratification, which presented as the main analysis in the study.

Table. HER2-ER dependent effect assessment of individual gene signatures in predicting Distant Metastasis Free Survival (DMFS). Main effect associated with a signature for DMFS prediction in a certain follow-up time interval was estimated by a Cox model within each HER2 and ER stratification. The Hazard Ratio (HR) along with its 95% confidence interval and the p value from the Wald test are shown. Numbers of patients at risk (n_{risk}) were computed at time point 0, 5 and 10 year, respectively.

<table>
<thead>
<tr>
<th>Gene signature</th>
<th>Time</th>
<th>n_{risk}</th>
<th>n_{event}</th>
<th>HR [95% CI]</th>
<th>p</th>
<th>n_{risk}</th>
<th>n_{event}</th>
<th>HR [95% CI]</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic-RORs</td>
<td>0-5yr</td>
<td>49</td>
<td>12</td>
<td>1.34 [0.72-2.51]</td>
<td>0.3513</td>
<td>65</td>
<td>22</td>
<td>1.17 [0.76-1.79]</td>
<td>0.469</td>
</tr>
<tr>
<td></td>
<td>5-10yr</td>
<td>37</td>
<td>0</td>
<td>-</td>
<td>0.1714</td>
<td>41</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>&gt;10yr</td>
<td>37</td>
<td>3</td>
<td>0.54 [0.23-1.30]</td>
<td>0.0716</td>
<td>41</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PAM50-RORs</td>
<td>0-5yr</td>
<td>49</td>
<td>12</td>
<td>1.12 [0.64-1.98]</td>
<td>0.6921</td>
<td>65</td>
<td>22</td>
<td>1.76 [1.02-3.04]</td>
<td>0.0433</td>
</tr>
<tr>
<td></td>
<td>5-10yr</td>
<td>37</td>
<td>0</td>
<td>-</td>
<td>0.0716</td>
<td>41</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>&gt;10yr</td>
<td>37</td>
<td>3</td>
<td>0.45 [0.19-1.07]</td>
<td>0.0378</td>
<td>41</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>70-gene</td>
<td>0-5yr</td>
<td>49</td>
<td>12</td>
<td>0.77 [0.45-1.31]</td>
<td>0.3333</td>
<td>65</td>
<td>22</td>
<td>1.15 [0.76-1.74]</td>
<td>0.5182</td>
</tr>
<tr>
<td></td>
<td>5-10yr</td>
<td>37</td>
<td>0</td>
<td>-</td>
<td>0.0378</td>
<td>41</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>&gt;10yr</td>
<td>37</td>
<td>3</td>
<td>0.21 [0.05-0.92]</td>
<td>0.0118</td>
<td>65</td>
<td>22</td>
<td>1.49 [0.99-2.25]</td>
<td>0.0572</td>
</tr>
<tr>
<td>76-gene</td>
<td>0-5yr</td>
<td>49</td>
<td>12</td>
<td>2.08 [1.18-3.67]</td>
<td>0.0118</td>
<td>65</td>
<td>22</td>
<td>1.49 [0.99-2.25]</td>
<td>0.0572</td>
</tr>
<tr>
<td></td>
<td>5-10yr</td>
<td>37</td>
<td>0</td>
<td>-</td>
<td>0.0118</td>
<td>65</td>
<td>22</td>
<td>1.49 [0.99-2.25]</td>
<td>0.0572</td>
</tr>
</tbody>
</table>

* p ≤ 0.05  •  p > 0.05
We also redid the analysis on HER2 stratification. We observe the decreasing time trend in the HER2+ group. However, due to limited events in the 5-10 year followup interval, we cannot draw conclusions about the time trend associated with Her2+ group and the differences in the prognostic power between the two HER2 groups.
"In the discussion "we did not find any notable effect of cohort differences..." followed by a call on supplementary figure 1 and Box5. Supplementary figure 1 has nothing to do with dataset differences and I could not find Box5. Please correct."

Please refer to the Supplement, Figure 1(different from Figure S1) and Box5, which are not listed in the main text.
Reviewer: Anita Grigoriadis

“Zhao and coworkers describe their meta-analysis of prognostic signatures for breast cancer. In this study the authors validate the already well known finding of a strong influence of hormone receptor status on the performance of the gene signatures. However, given that an extensive and new analysis of follow-up times is proposed, and since the authors have addressed all the previously raised points, this work merits to be published. Maybe to stress their survival analysis, we would suggest to change the conclusion accordingly.”

We have modified the Conclusions to emphasize the main points of the manuscript.