Reviewer's report

Title: Predict model: KI67 validation

Version: 2
Date: 10 October 2014

Reviewer: Chi-Cheng Huang

Reviewer's report:

In this manuscript, an updated version of PREDICT (v3) was proposed which incorporated KI67 into the predictive model. The regression weight of KI67 was derived from SEARCH study and the whole PREDICT predictive model was validated in the independent Nottingham cohort. In general, addition of KI67 did enhance the predictive model substantially in terms of 10-year survival.

Major compulsory revisions

1. The regression weight of KI-67 was derived from 2688 SEARCH patients. KI67 is only prognostic in ER+ BC and the authors should reveal the proportion of ER positivity in this discovery cohort. Besides, all hazard ratios and CI of each predictive factors including tumor size, grade, nodal status and HER2 as well as demographic features of discovery/validation samples should be detailed as in the Table 1/2 of Wishart et al: Breast Cancer Research 2010, 12:R1.

2. Missing value imputations were carried out in the validation samples. What’s the proportion of missing data of each prognostic factor within PREDICT?

3. The formula of predictive model, assumed Cox proportional hazard model, should be explained with more statistical details in the method section.

4. Since KI67 was only prognostic in ER+ tumors, PREDICT v3 should be restricted to ER+ subgroup and it’s quite weird to compare the AUC with (v3) and without (v2) KI67 for the whole cohort of 1726 ER-positive/negative patients (Results/discrimination). For the same reason, it’s redundant to compare the results of v2 and v3 for ER-negative patients (Results/calibration).

5. In Discussion, "It is anticipated that this improvement in model performance will contribute to more accurate predictions of the chemotherapy benefit for individual patients. " should be revised since there is no association between the adoption of chemotherapy and PREDICT prediction from this observational study. The "first-generation" chemotherapy in the study cohort inevitably introduced bias for contemporary breast cancer patients.

6. Page 9/line12: "They are based on mRNA expression in up to 70 cell cycle and proliferation genes [15,16,17]". The number of "70" cell cycle and proliferation genes is quite inappropriate because the three signatures, Mammaprint, Oncotype DX and PAM50 did not sum up to 70 genes. Actually the intersect of these signatures was surprisingly low and not all genes within these signatures were categorized into cell-cycle or proliferation-related genes.
7. Page 9/line 21-23: "the clinical validity – the calibration and discrimination of the recurrence predictions of the Oncotype DX RS have not been published." As one of FDA-approved multi-gene assay, the clinical validity is evidenced from the risk groups defined by the recurrence scores and the associated survival discrepancies. For clinical applications, survival discrepancies between defined groups were pursued and the authors' argument needs additional support.

8. The authors stated that "analytical validity of Ki67 IHC" is not of primary important in current study. If the measurement of Ki67 is inconsistent and questionable, how can they declared its clinical validity given the debate on the threshold/cut-off value of Ki67 dichotomy. Maybe a sensitivity analysis of Ki67 threshold could shed light on this ongoing issue. In the study of Karn et al. Breast Cancer Res Treat (2010) 120:567–579

bimodal distributions of ER, PR, and HER2 were identified from microarray experiments but a continuous distribution was revealed for Ki67, further indicating problematic classification of Ki67 status for clinical use.

Minor essential revisions
1. Background lines 12-13: "The classifications based on gene expression can be recapitulated using IHC." This is a strong argument and the authors should provide more details/references that some IHC measures could be the surrogate of the more complex gene expression profiles.

2. The authors also plan to incorporate gene expression profiles in the future version of PREDICT. What's the demand of gene expression profiles if their prognostic/predictive values could be reproduced by the IHC results. It's also interesting to see how to retrieve gene expression profiles from these historical cohorts.

3. Methods: Validation study population lines 24-25. The number of ER+ samples should be either 1274 or 1275. Which one is the correct?


5. Figure 2 should be 1274 ER "positive breast cancers" based on PREDICT v2 and PREDICT v3.

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.