Reviewer’s report

Title: Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model

Version: 1 Date: 3 April 2012

Reviewer: Fengxi Su

Reviewer’s report:

Major Compulsory Revisions

--The author might have missed a predictive model for SLN metastasis that was developed by Fabien Reyal(Reyal F, Rouzier R, Depont-Hazelzet B, Bollet MA, Pierga J-Y, et al. (2011) The Molecular Subtype Classification Is a Determinant of Sentinel Node Positivity in Early Breast Carcinoma. PLoS ONE 6(5): e20297. doi:10.1371/journal.pone.0020297). This model should be incorporated in your study since it suggested that molecular subtype is involved in predictive of SLN metastasis. On the other hand, the MSKCC model does not include molecular subtype or ER/PR/HER2 status as predictors. The author should have a discussion about the role of molecular subtype in the prediction model for axillary LN mets.

--In a multicenter study conducted by Rouzier’s team, (J Clin Oncol. 2009 Jun 10;27(17):2800-8. Epub 2009 Apr 6.), RP-ROC and CART were used for model construction. Are there any differences between your "tree" model and their RP-ROC or CART models? If there are some differences, please clarify why you choose your model rather than theirs. Besides, in their study, they performed an unreliability test and calculated Eaver and Emax and obtained a calibration curve. Do you think it is necessary to perform this analysis in your calibration plot?

Minor Essential Revisions

--In table 1, the information about clinical stage (TNM) should be clarify as that would make it clear whether the tumor burden was similar or not among the three dataset. Especially for the LN metastasis, the among of patients with pN0, pN1, pN2 etc should be clarified.

--In appendix C, if the patients has F=5, should the score of F be +0.5 rather than -0.3?

Discretionary Revisions

--The author should have a brief introduction about the rational and purpose of the pruning analysis since most clinical physician may not familiar with this analysis.
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:

Our team are performing similar research as theirs. They did a great job and we really learned a lot from their work.

We have no other competing interests to declare.