Authors investigated gene expression patterns of peripheral blood samples from mouse breast cancer models in attempt to identify potential diagnostic markers (or gene expression signature). The information provided could have very well relevant to early detection of breast cancer if the data are properly analyzed. They found that expression patterns of 4,276 gene features are sufficient to identify tumor-bearing mice in their initial discovery group and the association of the signature with breast cancer is later validated in independent cohort. Later, the signature was also validated in patients with breast cancer. This is an important clinical area and the study is generally well performed though there are a number of comments to make specifically with regards to the likely applicability of this signature in clinical practice.

Major concerns
1. According to description in methods (p17), authors identified 4276 genes first before dividing them into two groups. If this is true, prediction model would not be valid since information in validation set is already included in prediction model. Please clarify.
2. Have authors tried to use mouse data as training set to predict human samples?

**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