

Reviewer Comments:

Reviewer #1: This is an interesting paper on a study that examined the probability of immune recovery post-adjuvant therapy and the potential impact of adjuvant therapy type and cancer stage on immune recovery in newly diagnosed breast cancer patients. Although the topic is interesting and timely, there are several concerns that need to be addressed for publication.

In the abstract, please specify the study design used in the study and include 2-3 sentences on the data analysis methods used in the study.

At the end of the introduction, please clarify the hypotheses. Rewriting the hypotheses in a statistical way would help the audience follow the paper more easily. For example, the first hypothesis is stated as "immune responses would return to pretreatment baseline values by the 12-month time point." This hypothesis could be rewritten as "there is a significant difference in the immune responses between the pretreatment baseline measures and the 12-month time point measures."

Under the method section, please clarify the study design.

Under the section on participants, please provide more information on the settings (e.g., the types of the clinics, how many potential participants in each clinic). Also, please provide the rationale for the sample size (e.g., power analyses).

Under the section on data collection procedures and immune measurements, please separate the data collection procedures from the immune measurements (instruments) and provide more details on the data collection procedures (e.g., how the participants were approached through the settings, how the informed consents were obtained, where the samples were obtained in what environments by whom, etc.).

A generalized linear mixed models (GLMM) approach was used for the data analysis, which seems to be a sound approach. Yet, please link the data analysis methods to the hypotheses that were tested in the study.

The discussion section tends to be well written with supporting literature.

Reviewer #2: This is a very nicely written paper with a well developed logic. There is so much interest in PNI that this is a good addition to the literature. I fully agree that more long term studies are needed to see how these immune measures associate with recurrent disease and other clinical problems, or if interventions can accelerate the return to baseline (or improve) immunological function. The one suggestion I would make to

increase clarity of presentation is to remove all discussion of the association between immune markers, outcomes, and advanced disease in the Discussion section. In two places the authors note that there were no advanced cases in the study. A sentence referring to an association between immune markers and survival in persons with advanced disease, with all the references listed, would be sufficient. Otherwise, the reader has to keep sorting out data relevant to advanced versus earlier (curable) stages 1-3.

Reviewer #3: The purposes of this study were to "(1) examine how immune responses recover post-cancer adjuvant therapy over the first year of cancer diagnosis and treatment and (2) determine what effects the type of adjuvant therapy and cancer stage have on immune recovery in early stage (I-III) BC patients."

This manuscript intended to predict the immune recovery post-adjuvant therapy by type of therapy and stage of breast cancer. The first step before conducting a prediction study is to understand the actual situation of the phenomenon to be studied. Neither literature review nor the results section has addressed or reported on the actual percentage of immune recovery response at different time points after the adjuvant therapy. Therefore, a recommendation is to add the actual percentages of subjects who have reached or exceeded the baseline value within each adjuvant therapy groups under different time points in the results section. Such basic descriptive statistics are important to understand the actual situation.

The authors briefly described the significance of the study, variables related to the outcome measures, purposes of the study and hypotheses in the introduction section. There was no evidence or conceptual framework to support the two proposed hypotheses. It is not clear whether the intervention in the parent study has any effect on immune recovery response. The authors need to address this issue and make sure such confounding factor has been taken care of.

There were several issues in this prediction model. First, the authors did not clearly specify what model they were used. Generalized linear mixed model was used to estimate regression coefficients for predicting the probability of the immune responses recovery. However, there was no information about how the prediction models were set up. If the authors can provide one example, it will be helpful for readers to understand better how their data were analyzed. Most of the readers may not be interested in the model equations, but such equations can help others researchers duplicate similar studies. In Tables 3, 4 and 5, readers see many probabilities, since the model was not clearly presented, it is difficult to know how those probabilities were derived. For example, the authors provide an equation about how probabilities were calculated but the information of the model was not sufficient, without knowing what does the beta (Greek) represent, the

equation is less meaningful. The readers need to know how the beta(Greek)was derived?

The next question is how to assess the accuracy of those probabilities? This relies on different factors such as sample size, how good is the model fit, any missing values and how missing values were handled. The authors need to give more information on these issues. With such detailed information, interpretation of the findings will be more accountable.

The results section needs to be strengthened. The authors only used approximately one page to report the findings. One third of the results were related to demographic characteristics. A majority of the participant's characteristics were related to regiments of chemotherapy. If dosage of chemotherapy has impact, then this should be well addressed and controlled in the prediction model. The findings were based on Tables 3, 4 and 5. There is a concern about reporting probabilities without differentiating whether they were derived from individual therapy group or combined all groups. When there was no time and group interaction, the probabilities were based on combined groups, if the interaction exists, then probabilities were based on individual group. When the authors reported the probabilities, they did not differentiate them. It is not clear whether the sample size for each group under CD4, IFN-Gama, LAK are the same as the other combined groups. By examining the standard errors, it seems that the sample sizes were different. If this is the case, shouldn't the authors present the range of probabilities separately? In order to make the table presentation clearer, the word "combined" under CD4, IFN-?, and LAK can change to "Chemo+radio" in order to avoid the confusing with the word "Combined" which means to combine the three adjuvant therapy groups.

CHECKLIST FOR STYLE

Title Page: Provide the professional title for J. Carpenter