Reviewer's report

Title: Breast cancer stage at diagnosis and area-based socioeconomic status: A multicenter 10-year retrospective clinical epidemiological study in China

Version: 2 Date: 10 January 2012

Reviewer: Mansoo Yu

Reviewer's report:

This study focuses on examining how individual demographic variables (e.g., education and occupation) are associated with breast cancer stage, and how these differ between lower and higher SES areas in a sample of 4211 breast cancer patients in China.

I. Major points:

1) In Table 3, the researchers used ordinal logistic regression models to examine how individual demographics predict breast cancer stages (early, middle, and late stages). For logit models with ordered categories (see Allison, 1997), it would be appropriate to report models such as Model 1 (early vs. middle), Model 2 (middle vs. late), and Model 3 (early vs. late). Then, in each model, they could examine how SES areas (high vs. low) moderate the effects of the different levels of individual demographic variables, particularly education, on breast cancer stages. Interaction plots would be useful as well.

2) In Table 3, the findings are counterintuitive. Results show that the effect of education on cancer stage is more evident among those living in high SES areas than among those living in low SES areas. This finding is not consistent with a reference cited in Discussion (Singh et al., 2003) indicating that breast cancer patients from lower income areas had lower survival rates than those from higher-income areas (p.10).

3) In testing the ordinal logistic regression models, it would be more interesting to test omitted variables from the study variables in Table 2 such as BMI and receptor status in predicting breast cancer stage. In addition to the variables that the authors noted in study limitations (p.10), other variables such as access to breast cancer screening services, alcohol use, and family history of breast cancer would be useful to test their effects on breast cancer stage in the future unless the variables are available to be tested in the current study.

4) Based on the study findings, education is critical to reduce breast cancer patients because the other variables were not significant in predicting breast cancer stage. In other words, the study suggests that promoting education is an important way to prevent for breast cancer. Could education be a proxy variable representing risk factors for breast cancer such as alcohol use, lack of exercise, access to health care services, health insurance coverage, etc.?
II. Minor points:

5) Sampling: purposive sampling would be more appropriate than convenience sampling because of three inclusion criteria for the hospital selection (page 4).

6) On page 5, the researchers noted that there are seven areas of economic and education status. However, there are only four described in the text.

7) In Table 1, range of the study variables would be useful. One way ANOVA analyses are needed for post hoc tests to indicate significance among multiple groups.

8) A number of grammatical errors and inappropriate words are evident in the manuscript. For example, standard variances should be changed to standard deviation (p. 5).

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.