Reviewer's report

Title: Association of High Obesity with PAM50 Breast Cancer Intrinsic Subtypes and Gene Expression

Version: 3 Date: 26 November 2014

Reviewer: Zuzana Kos

Reviewer's report:

The authors present an interesting study of breast cancer gene expression in association with level of obesity. Particularly interesting is the finding of decreased ESR1 expression (and corresponding higher expression of proliferation genes and association with more aggressive subtypes) in tumours from highly obese women as compared to their less obese counterparts. This finding runs contrary to the current dogma of estrogen-driven tumourigenesis in overweight postmenopausal women, and indeed seems to indicate a threshold of obesity above which increased plasma estradiol level is no longer the driving force in tumour progression as suggested by the authors.

I do have some comments and suggestions for the authors.

Minor Essential Revisions

1. Particularly relevant to this study are some of the biases introduced by pooling the LACE and Pathway studies. The authors acknowledge in their discussion that including women enrolled in the LACE study (average enrollment 2 years post diagnosis) could potentially bias against the inclusion of more aggressive tumours. Tumours with higher proliferation rates (and lower ER expression) are more likely to recur early and be excluded from the LACE study. Equally important, however, is data from the authors’ prior study [ref 46] that show that there were significantly more obese women in the Pathway study versus the LACE study. If there is also a disproportionate number of highly obese women represented in the Pathways study (average enrollment 2 months post diagnosis), there is certainly a concern that this may influence the proportion of tumours with lower ESR1 and higher proliferation gene expression in the final analysis as these tumours are more likely to be captured in the Pathway versus LACE study. The authors should address these confounding biases more directly and consider showing the data and analyses from the individual study cohorts to demonstrate their findings of lower ESR1 expression (and higher proliferation gene expression) in the highly obese is not a spurious result of combining these two particular study samples. Of note, the authors have clearly considered and performed this analysis and do state that restricting the analyses to the individual trial cohorts resulted in findings "similar" to the combined data. These biases are central to the study findings, however, and deserve to be addressed with greater detail within the manuscript.
2. The BMI for this study was computed from self-reported weight and height data. The biases around self-reporting of these variables and resultant skewing of obesity-related epidemiological studies are well known and studied and should be discussed as a limitation of the study.

Discretionary Revisions

1. I would recommend that the authors change the wording in the methods-study population section, 1st paragraph describing entry criteria into the LACE trial: “no breast cancer recurrence or other cancer diagnosis within five years of enrollment”. This can potentially be misinterpreted as meaning no breast cancer recurrence within 5 years of enrollment.

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: I declare that I have no competing interests