Reviewer's report

Title: Inter-observer agreement according to three methods of evaluating mammographic density and parenchymal pattern in a case control study: impact on risk prediction

Version: 1 Date: 22 December 2014

Reviewer: steven allen

Reviewer's report:

This is a well written paper which is largely scientifically sound although I have a few but important points to make.

In looking at 3 subjective methodologies, this paper is of interest but not scientific importance as newer automated methodologies have been shown to be far superior in reliability of density assessment. This reference is a case control study to this effect "

Digital mammographic density and breast cancer risk: a case control study of six alternative density assessment methods."

Eng, A. ; Gallant, Z. ; Shepherd, J. ; McCormack, V. ; Li, J. ; Dowsett, M. ; Vinnicombe, S. ; Allen, S. ; Dos-Santos-Silva, I. ;

Another major flaw in this paper is that there has been no assessment made of non breast density breast cancer risk other than age. It is alluded to in the discussion that BMI was not obtainable and clearly this is an important example of this. However there are many other questionnaire type breast cancer risk variables such as menopausal status, contraceptive use etc. all of which are not controlled for in your study.

With the above in mind I therefore cannot accept in anyway your results/conclusions about breast cancer risk in relation to breast density. I would consider this a major flaw in the study and therefore the only result of interest to me is the comparative analyses between the different methodologies.

More specific points are as follows:

With regards automated analyses there are several papers including the above that cite these techniques and show their accuracy. this should be mentioned more extensively in relation to both the introduction and discussion.

Your follow up data concludes at 3yrs which is a short timescale. Why was that used?

December 2010 was a long time ago. Why have you not published prior to this?
Within the discussion you allude to the masking effect (p17 line 8). Please explain this further discussion page 18, line 12. I think it is an assumption to assume Tabar classification familiarity is affected by reader experience especially when you describe that none of the readers had used this technique prior to the study discussion page 19, line 10. 4.7% of women would be reallocated on cumulus. This is a very good agreement and better than many other studies using this technology. Any explanation for this?

Level of interest: An article of limited interest

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'