Reviewer's report

Title: Intratumor cholesteryl ester accumulation is associated with human breast cancer proliferation and aggressive potential: a molecular and clinicopathological study

Version: 5 Date: 27 April 2015

Reviewer: John Pearson

Reviewer's report:

A very interesting and timely exploration of the involvement of tumor lipids and prognosis markers. Previous reviewers have been thorough, reviewing the statistics: the sample size is modest however the authors have clearly demonstrated beyond statistical error that proliferation and hence poor outcome are associated with higher intra-tumor CE levels. Methods are appropriate and the conclusions are systematically drawn from the data, the following adjustments are all minor in nature.

Major
1. Figure 2
   Please use the same scale for A and D, B and E, C and F.
   For A, B, C use boxplots (SPSS does these), they are standard and give a better picture of distribution.
2. For D, E, F the dotplots do not look very linear. D (LDLR by CE) is strikingly exponential. In the statistical methods it is stated that Spearman's correlation will be used (this is usually "rho =", however in the figures r= and rho= are reported where the rho's look like p values and r is usually Pearson's correlation coefficient. Spearman is definitely better than Pearson here as the relationship is clearly not linear. I suspect that this can be fixed by changing r to rho and rho to p.

Figure 4
3. Please use same scales on A, B and C. Show outliers in A as these are highly influential for the correlations shown in B and C and are included in the significant p value shown in A.

4. Minor
I 93 missing "to" our instit....
I 111 eluded eluted
I 1614, so are the p values reported the actual p values or Bonferroni adjusted? My preference is to report the unadjusted and quote an adjusted threshold (say 0.05/3 if there are 3 comparisons).
I 266 delete "of"
I 273 coherently? maybe consistently.  
I 283 "such as LDLR" maybe "similar to LDLR levels"

5. Discretionary

For figure 2, fitting a 4 or 5 parameter logistic should give some idea of the kinetics and would be a better indicator of the relationships than the line shown. Alternatively, a line through the log expression data would fit better. For CE, it looks like low CE is showing no appreciable relationship with LDLR, while at high CE the relationship is logistic, same for SCARB1. CD36 is somewhat less clear, but there is a hint of a low magnitude positive linear relationship at high CE levels.

**Level of interest:** An article of outstanding merit and interest in its field

**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.