Reviewer's report

Title: Prediction of breast cancer by profiling of urinary RNA metabolites using SVM-based feature selection

Version: 1 Date: 30 January 2009

Reviewer: Carsten Denkert

Reviewer's report:

The paper by Henneges et al reports a new approach for profiling of urinary RNA that could be used to separate patients with breast cancer from a healthy control group.

The strength of this article is that it provides a promising strategy for integration of metabolomic techniques into new diagnostic approaches, the techniques used are highly innovative.

Minor essential:

1. The paper also offers an interesting bioinformatical strategy using the arctan function as a modified normalization approach. (I have discussed this with a bioinformatician in my group.) While this generally makes sense from a bioinformatical point of view, there is not enough information on this approach in the manuscript. The authors should give references and explain in more detail the effects of this normalization and compare their approach with other possibilities for normalization. It is also confusing that sometimes arctan and sometimes (e.g. figure 6 "atan" (which probably refers to the R-package) is used, this should be corrected. Information on the software used for analysis should be included.

2. The authors should discuss more clearly the limitations of their study. It should be emphasized that this technique is not ready for a clinical diagnostic test, because an independent validation is lacking. This is probably clear to the authors themselves, however, it should be made clear to the reader who reads this from a clinical perspective. The authors might also consider to slightly change the title, since their study does not "predict" breast cancer, but gives an interesting new approach for classification of biofluids, that might be used for a diagnostic test in the future.

The sensitivity and specificity are impressive, but a leave-one-out analysis of one cohort can not be the basis for a diagnostic test. The authors might consider to perform a phenotype permutation test to further validate the leave-one-out approach. It is clearly beyond the scope of this study to perform an additional validation, however, the discussion should be adjusted.

An additional statistical review could be considered, but is not essential from my point of view, in particular if the authors add additional information on the
bioinformatics approach.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**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.