Author’s response to reviews

Title: A specific gene expression signature for visceral organ metastasis in breast cancer

Authors:
Cemile Savci Heijink (c.d.savciheijink@amc.uva.nl)
Hans Halfwerk (j.b.halfwerk@amc.uva.nl)
Jan Koster (jankoster@amc.uva.nl)
Hugo Horlings (h.horlings@nki.nl)
Marc van de Vijver (m.j.vandevijver@amc.uva.nl)

Version: 2 Date: 20 Dec 2018

Author’s response to reviews:

Manuscript: BCAN-D-18-00915R1

Title: A specific gene expression signature for visceral organ metastasis in breast cancer

Savci-Heijink CD, Halfwerk H, Koster J, Horlings HM, van de Vijver MJ.

Point by point reply to issues raised by reviewer 1:

Reviewer reports:

Reviewer #1:

Comment: This reviewer can understand that it was (and is) not possible to validate the signature in an independent cohort (e.g., no samples, no money, no people). However, the question remains what we learn from a signature with 80% sensitivity and 40% specificity. This has to be explained in much more detail.

And if the authors decided not to perform any wet-lab work I would expect comprehensive state-of-the-art in silico work (i.e., high-end statistics and thorough discussion of the pros and cons of the different statistical approaches), as proposed by reviewer#2. So far, the statistics are a bit too basic for a pure in silico manuscript about already published samples.
Answer:

We acknowledge this concern. Yet we believe that associations between the gene expression profiling of primary breast carcinomas /herein identified gene expression profiler for visceral organ metastasis and survival outcomes in such a large series of patients are worth of reporting. As we have mentioned earlier, the reproducibility and the robustness of this signature need to be validated in other independent data sets of breast cancer patients with detailed information on metastasis pattern/behavior. This is addressed with our statement in the closing paragraph of discussion section (discussion section, lines 307-309, page 15).

For our study, we have chosen for k-means method, a simple but and intelligible approach.

As can be seen in our response to Reviewer #2, we completely understand that there are many ways in which a prediction model can be calculated, with varying degrees of ease to enable reconstruction by others. The method we chose for this study, is simple and can easily be utilized/reproduced. This analysis is also readily available in the R2 platform (which we used to generate the signature and separation).

We have set the goal for this signature as setting to separate a cohort and learn about covariates, not necessarily to classify a newly presenting single sample to either of the 2 groups. Therefore, the test set served to show that the procedure has worked and provided insight into the clinical associations. The real value of course comes from its (albeit limited) ability to reproduce in a completely independent set. To assess the robustness of the separation, we executed the k-means procedure 10 times and picked the solution that was mentioned most frequently.