

## Kirjaamo

---

**From:** Marko Sysi-Aho <marko.sysiaho@gmail.com>  
**Sent:** 10. kesäkuuta 2016 13:55  
**To:** Kirjaamo

**Subject:** Re: FW: The Process for handling allegations of the responsible conduct of research

Dear recipient,

Today I read through the reports and draft decision on the investigation. Overall the reports summarize what can be concluded by reviewing the material available upon article submission. The reports do have some deficiencies that would become apparent if the results in the investigation were tried to be reproduced from the mass instrument raw files.

In order to genuinely evaluate whether good scientific practice has been followed reproduction of the results from the raw mass instrument files and interviews of the researcher conducting the study would be necessary.

Since this is not the case in the current investigation, as an alternative I suggest that the processed data set used in the publication would be sent as a supplementary online material to the journal publishing the article. The processed data owned by VTT and other owners is the same data that I have in close collaboration with Matej produced and it does not contain any personal data on the patients. Neither does it have any further IPR value as the publication and patent have been already sent out, and the processed data is shared by various collaborators in Turku, Tampere, Aalto and Zora during the years 2006-2011. Maria Saarela requested that data from me late 2013 for VTT wanted to investigate claims in the publication and funding applications.

That processed data is an end result of many manual selection steps inducing uncertainties and limitations that were thoroughly discussed by me, Matej Oresic, Tapani Suortti and also in part by Mikko Katajamaa. These uncertainties and limitations were meant to be explained in the publication but they were considered to be out of focus and unnecessary by the corresponding author. Also, contrary to the beliefs of the reviewers in the investigation reports, sophisticated and complicated statistical methods were applied to investigate the longitudinal aspects of the data and to evaluate the predictive performance of lipid and metabolite profiles to predict onset of T1D. None of them showed prognostic potential for early detection/prediction of T1D on a patient level. These analyses and poor results were thoroughly known and discussed by me and Matej and they were meant to be described in the publication, but were kept unnecessary and out of focus by the corresponding author.

Personally, I am not aware of gross falsification or fabrication related to the published article. However, the article contains intentional selection of what has and has not been presented. It is very hard for a layman to tell whether this is in accordance with a good scientific practice or whether it is, or should it be, a common practice or not in the scientific community.

In my opinion, sharing the anonymized processed data as a supplementary material to the article would be best and most transparent solution allowing anyone to base their claims on the actual data and putting an end to speculations and discussions going around the publication.

Sincerely,  
Marko Sysi-Aho