UNIVERSITY OF COPENHAGEN
DEPARTMENT OF NUTRITION, EXERCISE AND SPORTS

Poster!

Pre-print!

QC4Metabolomics: Real-time and Retrospective Quality Control of Metabolomics Data

Jan Stanstrup, Lars Ove Dragsted

Why do I need this?

1) Untargeted LC-MS data is too complex to evaluate the data quality fully during analysis
2) Spot-checking is not sufficient, and many issues may be overlooked (loss of calibration, contaminants, ion suppression).
3) Mathematical corrections are often not sufficient to fix critical problems that remain unnoticed until the data analysis step.

What does `QC4Metabolomics` do?

- Automatically monitors user-specified ions in (almost) real-time and calculates quality metrics (fig. 1).
- Sends you an e-mail alert when a quality parameters is outside a given threshold.
- Monitors ~800 common LC-MS contaminants (fig. 2).
- Provide automatic conversion of raw data to mzML.

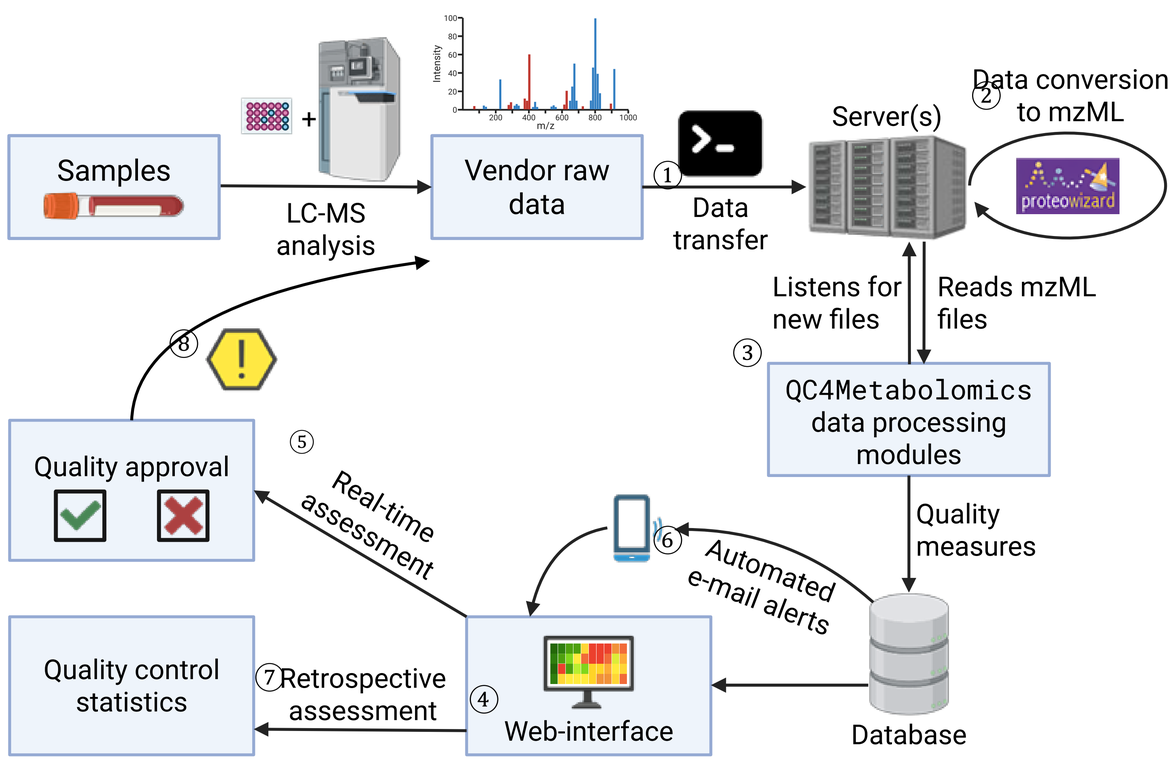

Figure 1. Workflow of `QC4Metabolomics`. Created in BioRender. Stanstrup, J. (2025) https://BioRender.com/2jlx43q.

Automated data analysis workflow

① During data acquisition, the raw data is transferred by a simple batch script to a server (optional) accessible by `QC4Metabolomics`.
② The data is then automatically converted to mzML by `MSconvert` from `ProteoWizard`.
③ When a new mzML file is created, `QC4Metabolomics` starts processing the file and writes the resulting quality parameter data and contaminant levels to a database.
④ This data is then available on a web-interface.
⑤ The user can decide whether or not the system is sufficiently performant and take actions as necessary.
⑥ An e-mail alert is send if a quality parameter is outside a given threshold
⑦ Retrospectively the data can also be reviewed for possible issues to be addressed during pre-processing.
⑧ Eventually, a decision can be made to re-analyze the samples when the system is again deemed performant.

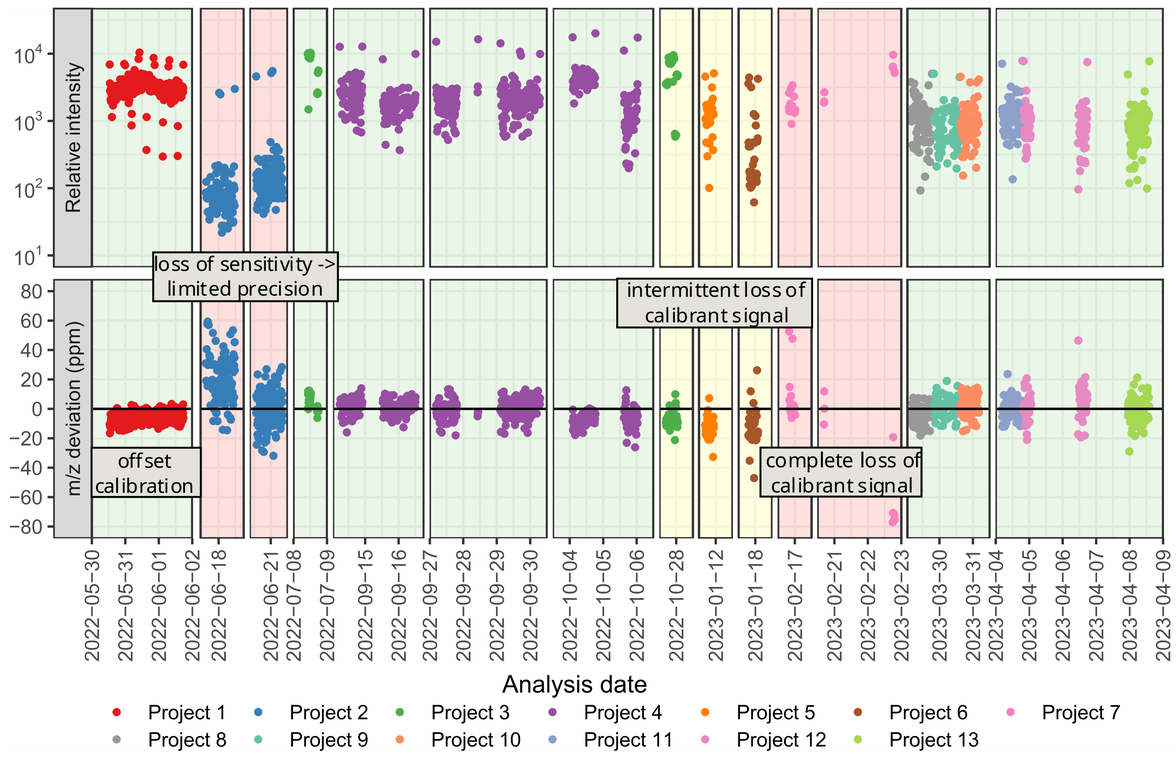

Figure 2. Example of performance monitoring. The response of the ion corresponding to tryptophan from all injections between May 2022 and April 2023 are shown. The upper panel shows the intensity while the lower panel shows the relative *m/z* deviation. The plot is not a screenshot from `QC4Metabolomics` but constructed based on data extracted from our `QC4Metabolomics` DB.

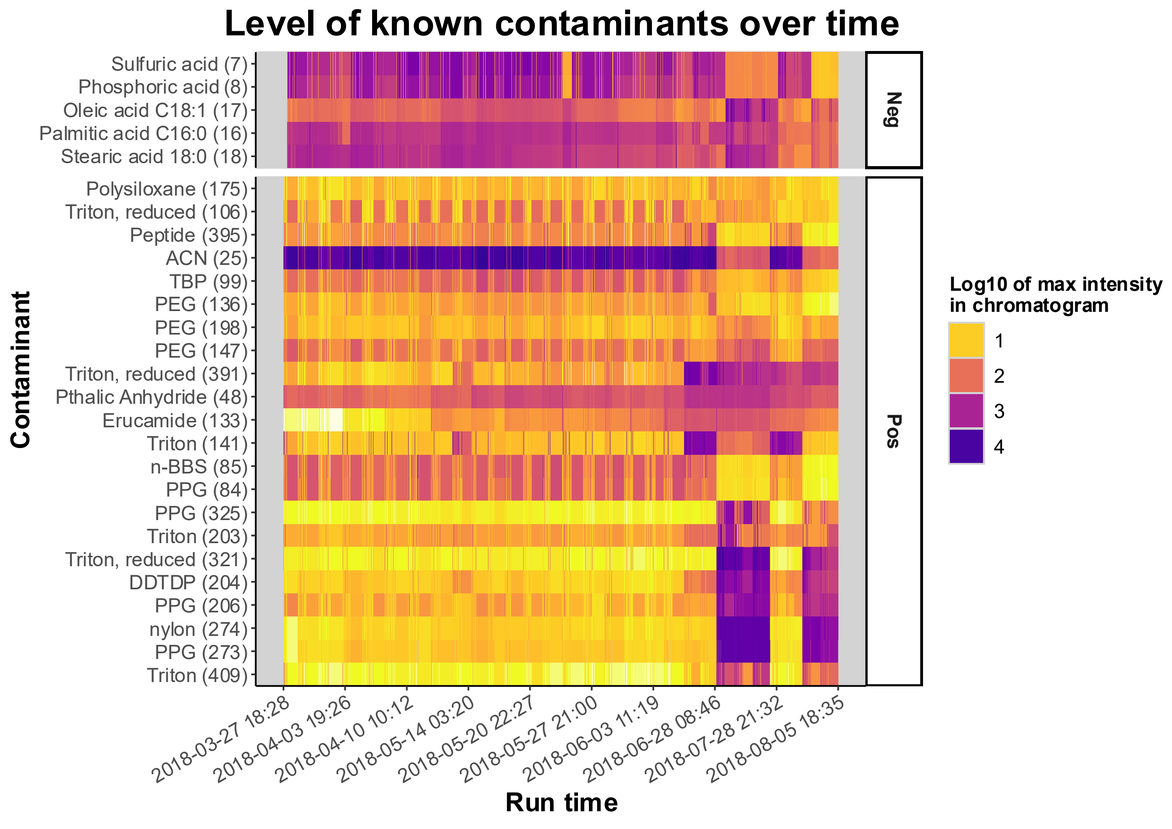

Figure 3. Screening of contaminants across injections. On the horizontal axis the detected contaminants are shown and on the vertical axis is the order the samples were analyzed in. The color scale indicates the intensity of the contaminant. It is evident that some contaminants appear periodically and disappear again. This might warrant investigation of the contamination source.

How do I get my hands on this?

You can download a docker-compose file from https://github.com/stanstrup/QC4Metabolomics, where you also find a demo and the documentation.

Acknowledgments: This work was supported by the Novo Nordisk Foundation (NNF19OC0056246; PRIMA—toward Personalized dietary Recommendations based on the Interaction between diet, Microbiome and Abiotic conditions in the gut)