Poster 037: MSlineaR: a new tool to assess linearity, improving statistical robustness and quality assurance in untargeted metabolomics

Janine Wiebach[1], Álvaro Fernández-Ochoa[2], Jennifer A. Kirwan[1]
[1]Metabolomics Platform of the Berlin Institute of Health (BIH) at Charité, Berlin, Germany; [2]Department of Analytical Chemistry, University of Granada, Spain

Aim

- Automate and accelerate **dataset clean-up**
- Improve **statistical robustness**
- Produce **filtered data set** with high level of **quality control**

Introduction

A major challenge of mass spectrometry based untargeted metabolomics is to **distinguish signals which carry useful biological information** from those which do not.

Three broad categories of signal:

i. **Biological signals** which show a **linear response with increasing concentration**
ii. **Non biological signals** (contamination, noise etc), mostly non-linear response
iii. Biological signals with **non-linear response** (unsuitable for quantitative statistical tests).

Analysing a dilution curve of a QC sample enables signals with linear responses to be detected: **challenging and slow to do manually**.

Materials and Methods

MSlineaR performs 7 steps to classify a feature as linear or not linear which are depicted in figure 1.

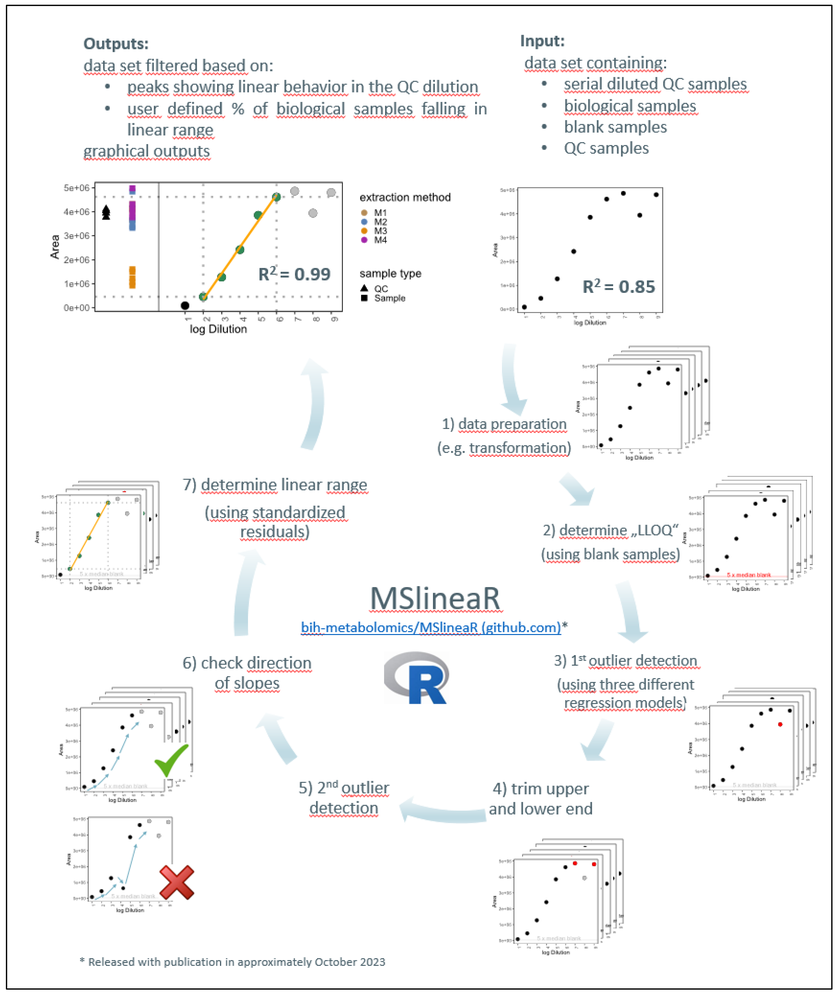

Figure 1. Schematic workflow of MSlineaR.

Results and Discussion

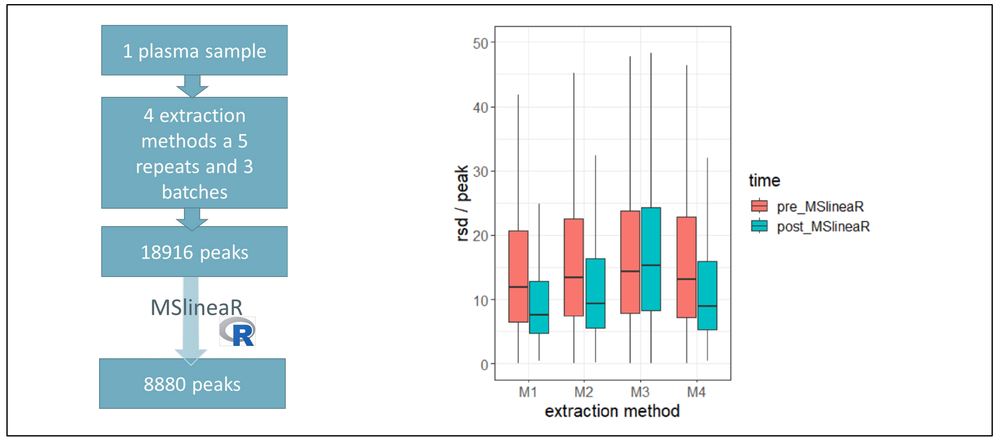

Figure 2. RSD per peak for the unprocessed data set (red) and after processing with MSlineaR (blue) for 5 different extraction methods. Removing non-linear peaks led to a reduced data set size of 88 %. The median RSD and the variation could be reduced for most of the extraction methods after using MSlineaR.

Table 1. Specificity of MSlineaR using an untargeted data set for benchmarking. As negative control, 470 heavy labeled C13 metabolites from the IROA internal standards were added per batch with constant concentration across all samples, which should lead to a non linear behavior. MSlineaR correctly classified all detected features.

	Linear	Non linear	Specificity
Peaks	0	470	**100%**

Table 2. Sensitivity of MSlineaR using a targeted data set for benchmarking. As positive control, 68 standard calibration series were used for which a linear range was expected. MSlineaR classified 6 features as non linear, which were confirmed by eye, leading to a sensitivity of 100%.

	Linear	Non linear	Sensitivity
Peaks	62	6*	**100%**

* Also assessed by eye as non linear

Conclusion

- **MSlineaR improves quality assurance by cleaning up data sets quickly in an automatized manner by removing noise and low quality peaks.**
- **It is the only accessible tool which includes an outlier detection and determines not only the linearity but also the linear range.**
- **It can be used for untargeted and targeted metabolomics and can be easily inserted in an existing preprocessing workflow in R.**