

Supporting Information

MSlineaR: An R Package Assessing Linear Behavior to Improve Quality Assurance and Statistical Robustness in Untargeted Metabolomics

*Janine Wiebach^{*1,2}, Álvaro Fernández-Ochoa^{2,3}, Ulrike Bruning^{1,2}, Jochen Kruppa-Scheetz⁴, Maëlle Bonhomme⁵, Dominique-Marie Votion⁵, Jennifer A. Kirwan^{*1,2,6}*

1 Berlin Institute of Health at Charité Universitätsmedizin, BIH Metabolomics, Charitéplatz 1, Berlin 10117, Germany

2 Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Robert-Rössle-Str. 10, Berlin 13125, Germany

3 Department of Analytical Chemistry, University of Granada, Avenida de Fuentenueva S/N, Granada 18071, Spain

4 Hochschule Osnabrück University of Applied Sciences, Albrechtstr. 30, Osnabrück 49076, Germany

5 University of Liège, Faculty of Veterinary Medicine, Department of Functional Sciences, Pharmacology and Toxicology, Comparative Veterinary Medicine, FARAH, Avenue de Cureghem, 7A B41, Quartier Vallée 2, Liège 1 (Sart Tilman), Liège 4000, Belgium

6 Metabolomics Unit, University of Veterinary Medicine Vienna, Veterinaerplatz 1, Vienna 1210, Austria

Text S-1 : Analytical Analysis and Processing of Data Set 1 (Targeted Data Set)	3
Text S-2: Analytical Analysis and Processing of Data Set 2 (Untargeted Data Set).....	5
Text S-3: Analytical Analysis and Processing of Data Set 3 (Untargeted Biological Data Set).....	7
Text S-4: Evaluation of different metrics of ABL in detecting true positives assessed by MSlineaR using DS2	12
Figure S-1: Example feature which was removed after applying the class-based filter from DS2	14
Figure S-2: Examples of 4 non-linear features contributing to the separation by PCA between MeOH1.5 and other groups in DS2	15
Figure S-3: Median RSD comparison of DS2 without and with using MSlineaR	16
Figure S-4: Example of unexpected ion behavior at high dilutions	17
Figure S-5: Assessment of the dilution-dependent signal behavior of 142 IROA ¹² C features from DS2 across three batches (provided as a separate Supporting Information PDF)	
Figure S-6: Venn diagram depicting the overlap in linearity of 126 IROA standards across three batches from DS2	18
Figure S-7: Three dilution curves from DS2 illustrating examples of features which were removed by MSlineaR.....	19
Table S-1: MetaData template for running MSlineaR.....	20
Table S-2: SampleData template for running MSlineaR.....	21
Table S-3: Parameters for running MSlineaR	22
Table S-4: Mathematical Formulation of the MSlineaR Algorithm.....	27
Table S-5: Metrics assessed by MSlineaR.....	30
Table S-6: Sequence order of data set 1	31
Table S-7: Sequence order for one batch of data set 2	31
Table S-8: preparation of dilution series from pooled QC	32
Table S-9: Sequence order data set 3.....	33
Table S-10: Number of metabolites retained by MSlineaR for DS1 using different R ² thresholds	33

Text S-1 : Analytical Analysis and Processing of Data Set 1 (Targeted Data Set)

Samples:

The first data set consists of 101 EDTA plasma samples consisting of 62 dilution curves of a single calibration standard from an unpublished data set. The original biological samples have not been analyzed here.

LC-MS Analysis:

The analysis consisted of a targeted analysis of tryptophan metabolites based on a previously validated method¹. Liquid chromatography – mass spectrometry (LC-MS) analysis was performed with a 1290 Infinity 2D HPLC system (Agilent Technologies, USA) combined with a TSQ Quantiva triple quadrupole mass spectrometer with a heated ESI source (Thermo Scientific, USA). Before starting, an extraction solvent was prepared comprising 90% methanol (LC-MS grade, HiPerSolv (VWR)), 0.15 µg/mL mixed internal standards, 0.02% ascorbic acid, and 0.2% formic acid and cooled to -20°C. For each sample, 80 µL pre-chilled extraction solvent was added to 40µL of EDTA plasma. Plasma samples were held at 4°C and shaken for 10 min at 800rpm (Eppendorf ThermoMixer C) before being stored in the -20°C for 60 min. The samples were centrifuged for 10 min at 11000g and 4 °C. The supernatants were transferred to brown LC-MS vials for LC-MS/MS analysis. 20µL of each plasma sample was pooled, and the pooled plasma was also extracted using the same protocol to make quality control (QC) samples. These QC samples were run every sixth samples.

The standards were mixed and serially diluted to achieve a concentration range of 0.045ng/mL to 888ng/mL for most metabolites except for a minority of metabolites with particularly high or low abundance in plasma (tryptophan, formylkynurenine, methyltryptamine, methoxytryptamine, kynurenine). These standard dilutions in solvent were subsequently used to extract charcoal stripped plasma (Fetal Bovine Serum Charcoal stripped, Sigma-Aldrich) in the same procedure as described above. The charcoal stripped plasma was designed to mimic the matrix effects of real plasma and reduce analytical differences between the samples and standard solutions. A medium range dilution (6th dilution) of the standard curve was used as a second QC sample type and injected every 10th sample throughout the sequence (Table S-6).

LC-MS analysis was conducted with a 10-min gradient. A reversed-phase column was used (VisionHT C18 Basic; L × I.D. 150 mm × 4.6 mm, 3 µm particle size) and held at a constant temperature of 30°C. The mobile phase consisted of 0.2 % formic acid in H₂O (solvent A) and 0.2 % formic acid in methanol (solvent B). The following gradient was run with a constant flow rate of 0.4 mL min⁻¹: A/B

97/3 (0 min), 70/30 (from 1.2 min), 40/60 (from 2.7 to 3.75 min), 5/95 (from 4.5 to 6.6 min) and 97/3 (from 6.75 to 10 min). Before starting, the LC-MS was conditioned by running 10 blanks (solvent A) and 3 QC samples for 130 min total run time. The molecular ion and at least two transitions were monitored for up to 32 metabolites that comprise the tryptophan pathway.

Data Pre-processing:

Data was exported into Skyline v.19.1 - 64 bit to identify and quantify peak intensity and area. Transition settings in the Skyline search were: isotopic peaks included: count; precursor mass analyzer: QIT (Quantifier Ion Transition); acquisition method: targeted; product mass analyzer: QIT. Method match searching tolerance was 0.6 m/z, and data were manually checked to ensure the correct peaks were selected. After internal standard normalization by dividing the area of each compound with the area of the corresponding internal standards to correct for technical variances, a cubic spline drift correction using the R notame package (version 0.3.2)² was applied per metabolite and to all sample types using the pooled QC samples as references to fit the splines³. The first and the last QC samples used to fit the cubic splines are the most critical to the resulting fit. In this case, the first pooled QC sample after conditioning, and the first of two pooled QC replicates at the end of the analytical run were used as the first and last QC respectively in the batch correction. QCs analyzed before or after these timepoints were disregarded for further analysis to avoid excessive leverage of the cubic spline. Background subtraction was performed on calibration standards and standard derived QCs using charcoal stripped plasma as the background reference. Subsequently, standards of increasing concentration were used to construct calibration curves for each metabolite by fitting a linear regression model. All compounds with an $R^2 > 0.96$ and 5 or more dilution points were defined as linear. Concentration values less than or equal to zero were declared as missing. Precision (%RSD less than 15% in standard QC samples) was adequate for all compounds.

Text S-2: Analytical Analysis and Processing of Data Set 2 (Untargeted Data Set)

Samples:

This data set was a validation data set originally designed to compare different published untargeted methods⁴⁻⁶. The test samples comprised of the commercial human plasma sample (GTX73265, Genetex) extracted with different solvents and/or sample-solvent ratios (method 1 “MeOH3”: plasma to methanol ratio 1:3; method 2 “MeOH1.5”: plasma to methanol ratio 1:1.5; method 3 “MeOH9”: plasma to methanol ratio 1:9; method 4 “MeOH_EtOH3”: plasma to methanol:ethanol mixture (1:1) in a final ratio of 1:3). 5 replicates of these samples were analyzed in each batch (Table S-7). It is important to note that the final injection concentration of MeOH9 samples was approximately eight-times lower than for the other extraction methods. Intra-batch pooled technical QCs were also prepared using the same method as for MeOH3 samples.

A dilution series consisting of 9 dilution points was prepared (with ratios of 0.05, 0.10, 0.25, 0.50, 0.75, 1, 1.5, 2, 3 to the pooled QC concentration) (Table S-8). 20 μ L of each sample was extracted with methanol (ratio 1:3) for protein precipitation. Subsequently, the samples were vortexed for 15 seconds and then centrifuged at 15800g. at 4 °C for 15 minutes. The supernatants were taken and fully dried. Prior to analysis, samples were reconstituted in 50 μ L of water containing IROA-C13 internal standards (TruQuant Yeast Extract Semitargeted QC Workflow Kit, IROA Technologies). These internal standards have distinctive isotopic 95% C13 labels (U-95% C13).

Three identical analytical batches (sequence of one batch in Table S-7) were analyzed in this study. Each contained its own dilution series. Additionally, these batches were made up of equilibration samples (pooled QC samples used only for equilibration and excluded from subsequent data analysis), technical pooled QC samples, long term reference sample (LTRS) from IROA (TruQuant Yeast Extract Semitargeted QC Workflow Kit, IROA Technologies), blank samples and a series of commercial human plasma samples prepared as described below.

LC-MS Analysis:

An adaptation of a previously published method was used for the analysis of the samples based on UPLC-ESI-Q-Orbitrap-MS⁶. Briefly, these analyses were performed in a UPLC system 1290 Infinity II (Agilent, Santa Clara, CA, USA) coupled to a Q-Exactive Plus (ThermoFisher Scientific, Waltham, MA, USA) with electrospray ionization mode. 5 μ L of sample was injected onto a reverse phase column (RP C18, ACQUITY BEH, 2.1x100 mm, 1.7 μ m, Waters

Corporation) at 50 °C. The following solvents were used as mobile phases: Phase A: Water + 0.1% Formic acid; Phase B: Methanol + 0.1% Formic Acid. A flow rate of 0.4 mL/min was used with the following elution gradient: 0.0 [100 % A], 1.0 min [100 % A], 16.0 min [0% A], 20.0 min [0% A], 22.0 [100 % A] and 24 min [100 % A]. MS acquisition was performed in heated electrospray positive ionisation (HESI) mode with a spray voltage of +4kV. The data were acquired in full scan using an m/z range of 60 to 900 m/z.

Data Preprocessing:

The acquired data were firstly converted to .mzML format using MSConvertGUI software (<https://proteowizard.sourceforge.io/>). Once converted, the data was further processed with R (version 4.2.0) in RStudio environment (version 2023.6.2.561).

XCMS package⁷⁻⁹ (version 4.6.4) was used to carry out the peak picking, RT alignment and grouping stages. The following optimized parameters were used in this package: peak width: 3.35-50; ppm: 10; mzdif: 0.00175, snthresh: 10, noise: 100000; gapInit: 0.78; gapExtend: 2.5; Response: 14; binsize-adjustRT: 0.41; minFraction: 0.5; bw: 2; binsize-group: 0.0055 and maxFeatures: 50.

For batch correction the first and last pooled QC are essential, therefore we ran duplicates of these as backups. Following manual inspection of the sum of area the duplicated QC samples and the equilibration samples were removed from further processing. For processing the rest of the data set, features were filtered out if: it was below a signal-to-noise ratio of 3:1 compared to the median of blank samples or it was not detectable in all three batches and in at least 80% of samples or it was not detectable in less than five pooled QCs, (to ensure good drift correction was possible). Next MSlineaR was applied with parameters set as follows: signal-to noise = 3, min_feature = 5 and min_R2 = 0.9. Due to the extreme differences in pre-analytic concentrations, we have omitted a post analytical normalization step. The notame package² (version 0.3.2) was used to correct for analytical drifts within batch using a cubic spline method¹³. To correct for inter-batch deviations a groupwise median normalization strategy was used. Technical variability was assessed by calculating the RSD per feature over the pooled QCs (RSD_{QC}). All features were removed which had technical variability above 15 (RSD_{QC} > 15%) to ensure only reproducible features were kept.

Text S-3: Analytical Analysis and Processing of Data Set 3 (Untargeted Biological Data Set)

Samples:

DS3 is derived from a longitudinal field study conducted on horses from three different locations over an 18-month period. Over the study period, venous blood samples were collected every two months. The study aimed to identify metabolomic biomarkers associated with the development of a specific condition (unpublished data, Bonhomme et al., in preparation). Fourteen horses were initially included (n = 7 developing the condition over time, n = 7 matched-controls). Two horses (one case and one control) originating from a third location were excluded from the statistical analysis, as they were the only ones from that site. All samples were collected in EDTA tubes under real-world conditions, then processed and analyzed using LC-MS. Samples collected during the 0–2, 2–4, 4–6, and 6–8-month intervals preceding the onset of the condition (or at matched intervals for the controls) were analyzed. These intervals reflect the regular bimonthly sampling visits rather than precise timepoints relative to the event. Immediately after collection, plasma samples were centrifuged at 1000g, aliquoted in 1.8mL polypropylene Eppendorfs (CryoTubes™, Nunc™, Thermo Fisher Scientific, USA) and immediately frozen and stored at –80 °C until further analysis.

To control stable environmental contamination, one “field blank” per stable was prepared using ultrapure water processed in the field identically to blood samples. Extraction blanks were also included.

Various quality control (QC) sample types were regularly injected after each 10th sample to assess instrument performance, monitor signal drift and correct for analytical variability. These included: intrastudy QC pooled samples (pooled from all biological samples of the batch) and class-specific intrastudy QC samples. Intrastudy QC pool and class-specific QC pools were used for MS/MS analysis. Additionally, intra-study QC pools were used to perform a calibration curve, consisting of 9 dilutions (0.05 to 3-fold) to filter nonlinear peaks during data processing.

Long Term Reference Samples (LTRS), comprising a commercial 1:1 IROA mixture of labeled U-5% and U– 95%-13C metabolites (Sigma-Aldrich, USA, IROA TruQuant Yeast Extract IQQ Workflow kit, supplied by IROA Technologies) were injected at the beginning and end of each sequence for quality control, identification, instrument performance monitoring and comparison across studies.

Internal standards were used to assess chromatographic drift and intensity variation over time, as well as to enhance metabolites identification. IROA internal standards, composed of several hundred labeled 95% U-13C metabolites (Sigma-Aldrich, USA, IROA TruQuant Yeast Extract IQQ Workflow kit, supplied by IROA Technologies) were added to QC samples, procedure blanks, field blanks and biological samples (except solvent blanks).

The LC-MS system was equilibrated using QC pools before the analysis of plasma samples.

Independent acquisitions were performed for both positive and negative ionization mode. 2 μ L of each sample was injected. Each acquisition started with the injection of three solvent blanks to flush out residual metabolites from the column, followed by five equilibration samples to condition the system (plasma QC pool without IROA internal standards), two QC pool samples, one LTRS, three extraction blanks, nine dilutions of QC pools and one solvent blank. Samples, field blanks, extraction blanks and the different QC samples were injected in a randomized order, with a QC pool sample run every ninth injections. The batch concluded with the injection of one LTRS, two QC pool samples and five solvent blanks to ensure no residue was left on the column (Table S-9).

LC-MS Analysis:

The samples were analysed following a protocol described by Southman et.al 2020¹¹ using UPLC-ESI-Q-Orbitrap-MS. Briefly, these analyses were performed in a UPLC system (Vanquish Flex UHPLC systems, Thermo Fisher Scientific, USA) coupled to a using a high-resolution mass spectrometer Orbitrap Exploris 240 (Thermo Fisher Scientific, USA) with electrospray ionization mode. System and mass calibration were conducted for both ionization modes before each acquisition run, using a mass calibration solution (Thermo Fisher Scientific, USA, Pierce™ FlexMix™ Calibration Solution, Reference A39239, Lot YJ382143, including a mixture of 16 high-purity ionizable components with a mass ranging from 50 to 3000 m/z, for positive and negative ionization calibration). 2 μ L of sample was injected onto a hydrophilic interaction liquid chromatography (HILIC) column (ACQUITY Premier, BEH Amide Vanguard FIT, 1.7 μ m, 2.1 mm x 150 mm, Waters Corporation, USA) at 35 °C. The following solvents were used as mobile phases: Phase A: 5% water + 95% acetonitrile + 10 mmol of ammonium formate + 0.1% of formic acid; Phase B: 50% acetonitrile, 50% water + 10 mmol of ammonium formate + 0.1% of formic acid. A flow rate of 0.4 mL/min was used for 14 minutes with the following elution gradient: 0.0 [99 % A], 1.0 min [99 % A], 3.0 min [85% A], 6.0 min [50% A], 9.0 [5 % A], 10.0 [5 % A], 10.5 [99 % A] and 14 min [99 % A]. MS acquisition was performed in heated electrospray positive ionisation (HESI) mode with a spray voltage of 3.5kV for

positive mode and 2.7kV for negative mode. The data were acquired in full scan using an m/z range of 70 to 1000 m/z . Mass spectrometer parameters were: sheath gas = 53 arbitrary units, aux. gas = 14 arbitrary units, sweep gas = 3 arbitrary units, vaporizer temperature = 438 °C for positive mode and 320 °C for negative mode, capillary temperature = 269 °C for positive mode and 320 °C for negative mode, resolution = 15 000, RF lens = 70%. Data-dependent tandem mass spectrometry was performed on one global QC pool, one case-class specific QC pool and one control-class specific QC pool over 100 – 1000 m/z scan range, using an isolation window of 3 m/z and normalized collision energies of 25, 60, 100% with a Orbitrap resolution of 15 000.

Data Preprocessing:

Data acquisition was performed using FreeStyle ThermoFisher Mass Spectrometry software (Version 1.8, Thermo Fisher Scientific, USA). Subsequent data processing was carried out using ClusterFinder (version 4.2.67, IROA Technologies, USA) and R Studio (Version 4.2.1). Instrument performances were evaluated through the monitoring of pressure curves, manual inspection of total ion chromatogram of individual acquisitions, assessment of QC samples metrics and signal drift visualization.

The acquired data were first converted to .mzML format using Proteowizard MSConvert (<https://proteowizard.sourceforge.io/>).

Peak detection, retention time correction, chromatogram alignment, feature annotation, and putative annotation (based on accurate mass) were conducted using XCMS Online (<https://xcmsonline.scripps.edu>, version 3.7.1)^{12,13}. The following parameters were used: $\Delta m/z$ = 5 ppm, minimum peak width = 5 s (10 s for negative mode), maximum peak width = 20 s (60 s for negative mode), signal/noise threshold = 4, $mzdiff$ = 0.01, Integration method = 1, prefilter peaks = 3, prefilter intensity = 100. For retention time correction, the obiwarp method was used with a $profStep$ of 1, ensuring precise alignment across samples. Chromatogram alignment parameters were set as follows to maintain consistency in feature detection across the data set: $minfrac$ = 0.5, $mzwid$ = 0.025 (0.015 for negative mode), bw = 5, $min\ samples$ = 1 and $max\ samples$ = 100.

For batch correction the first and last pooled QC are essential, therefore we ran duplicates of these as backup samples. Following manual inspection of the sum of area the duplicated QC samples were removed. Features were filtered out if they were below a signal-to-noise ratio of 5:1 compared to the median of blank samples. A feature also needed to be present in all three batches and in at least 80% of samples. Additionally, features were excluded where less than five pooled QCs measurements for that feature were present, to ensure good drift correction.

Data normalization was performed using probabilistic quotient normalization (pqn)^{20,21} with the R package pmp¹⁰ (version 1.20.0). The notame package² (version 0.3.2) was used to correct for analytical drifts within batch deviations using cubic spline method¹³. Technical variability was assessed by calculating the RSD and D-ratio per feature over the pooled QCs (RSD_{QC}, D-ratio_{QC}). All features were removed which had technical variability above 20% (RSD_{QC} > 20%), a D-ratio above 40% (D-ratio_{QC} > 40%) or which had a biological variability (RSD_{Sample}) lower than 1.2* technical variability (RSD_{QC}) to ensure only reproducible features with biological relevance were kept. Compound discoverer (Thermo Fisher, 3.3 SP2)

References

1. Holle, J. *et al.* Inflammation in Children with CKD Linked to Gut Dysbiosis and Metabolite Imbalance. *J. Am. Soc. Nephrol. JASN* **33**, 2259–2275 (2022).
2. Klåvus, A. *et al.* “Notame”: Workflow for Non-Targeted LC–MS Metabolic Profiling. *Metabolites* **10**, 135 (2020).
3. Kirwan, J. A., Broadhurst, D. I., Davidson, R. L. & Viant, M. R. Characterising and correcting batch variation in an automated direct infusion mass spectrometry (DIMS) metabolomics workflow. *Anal. Bioanal. Chem.* **405**, 5147–5157 (2013).
4. De Paepe, E. *et al.* A validated multi-matrix platform for metabolomic fingerprinting of human urine, feces and plasma using ultra-high performance liquid-chromatography coupled to hybrid orbitrap high-resolution mass spectrometry. *Anal. Chim. Acta* **1033**, 108–118 (2018).
5. Kouassi Nzoughe, J. *et al.* A Nontargeted UHPLC-HRMS Metabolomics Pipeline for Metabolite Identification: Application to Cardiac Remote Ischemic Preconditioning. *Anal. Chem.* **89**, 2138–2146 (2017).
6. Dunn, W. B. *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–1083 (2011).
7. Tautenhahn, R., Boettcher, C. & Neumann, S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics* **9**, 504 (2008).

8. Benton, H. P., Want, E. J. & Ebbels, T. M. D. Correction of mass calibration gaps in liquid chromatography-mass spectrometry metabolomics data. *BIOINFORMATICS* **26**, 2488 (2010).
9. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**, 779–787 (2006).
10. Jankevics, A., Lloyd, G. R. & Weber, R. J. M. *Pmp: Peak Matrix Processing and Signal Batch Correction for Metabolomics Datasets*. (2022).
11. Southam, A. D. *et al.* Assessment of human plasma and urine sample preparation for reproducible and high-throughput UHPLC-MS clinical metabolic phenotyping. *The Analyst* **145**, 6511–6523 (2020).
12. Gowda, H. *et al.* Interactive XCMS Online: Simplifying Advanced Metabolomic Data Processing and Subsequent Statistical Analyses. *Anal. Chem.* **86**, 6931–6939 (2014).
13. Tautenhahn, R., Patti, G. J., Rinehart, D. & Siuzdak, G. XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Anal. Chem.* **84**, 5035–5039 (2012).

Text S-4: Evaluation of different metrics of ABL in detecting true positives assessed by MSlineaR using DS2

IROA ^{12}C isotopes in data set 2 were used as true positives to assess different metrics for evaluating diluting-dependent signal behavior. There were 142 C12 isotopes detected in 3 batches (total of 426 dilution curves), of which 94 were manually assessed as not exhibiting a dilution-dependent signal response in a dilution curve (likely due to poor ionization in these conditions). This left 332 isotopes as our ground truth to assess metrics. R^2 successfully detected a high level of true positives.

Retained IROA ^{12}C	Retained IROA ^{12}C in %	Metric
330	99.40	$R^2 > 0.9$
305	91.87	Monotonicity whole range ($\rho = 1$)
320	96.39	Monotonicity inRange ($\rho = 1$)
280	84.34	deviation $\leq 20\%$
234	70.48	slope within 20% (proportional)

We also compared the metrics by using different thresholds for R^2 :

R^2	Nr curves	monotone	monotone in range	deviation $\leq 20\%$	slope within 20% (proportional)
>0.98	269	258	269	254	206
>0.96	302	284	299	267	224
>0.90	330	300	319	277	233

For example, there are 330 curves with an $R^2 > 0.9$ and 319 are monotone in the trimmed range. In the figures below we show in two examples where a subjective decision to retain some features may be justified even when they are not monotonic. In both examples, there is a generally consistent dilution-dependent response and a high R^2 , but due to a small deviation, the feature is

not strictly monotonic and may be excluded if other metrics of linearity are used. In hypothesis-generating metabolomics studies, scientific aims may allow more tolerance to retain such features than in validation studies where stricter criteria may be preferred.

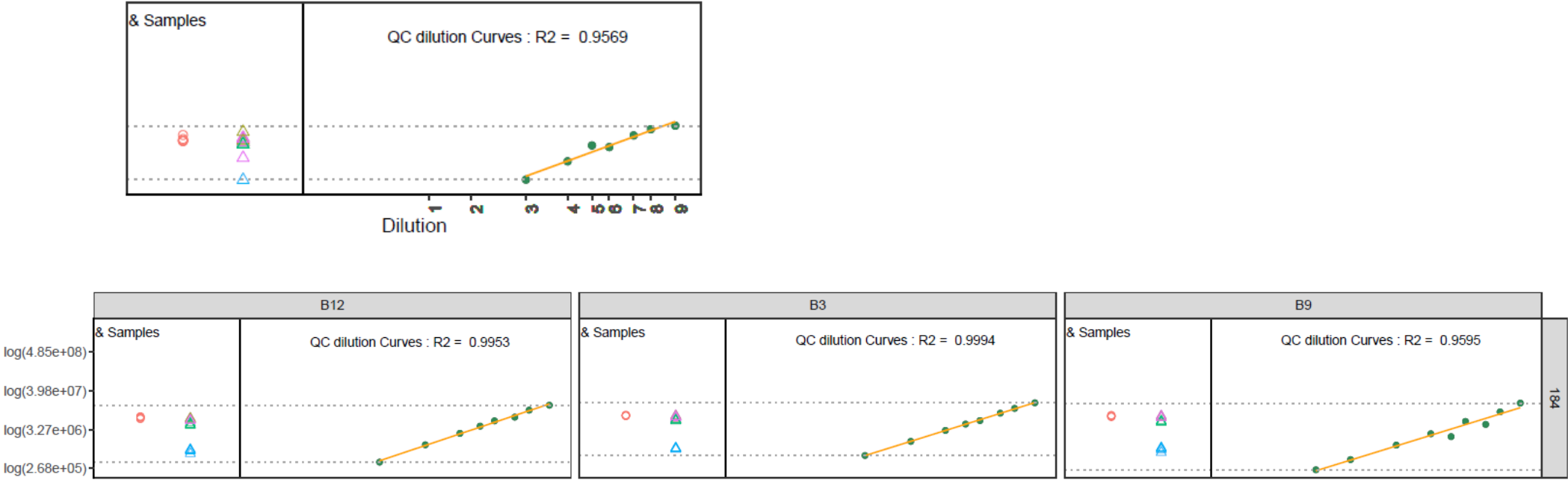


Figure S-1: Example feature which was removed after applying the class-based filter from DS2

Dilution curves illustrating examples of features which were removed by MSlineaR. The dilution curve shows good linearity, but 70% of features across all sample groups are outside either the upper or the lower limit of linearity. S1 = MeOH3, S2 = MeOH1.5, S3 = MeOH9, S6 = MeOH_EtOH3.

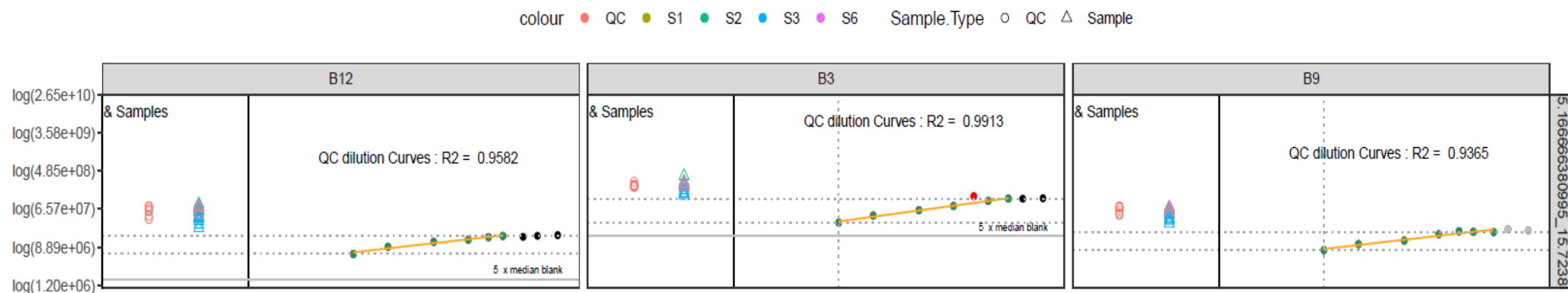


Figure S-2: Examples of 4 non-linear features contributing to the separation by PCA between MeOH1.5 and other groups in DS2

Linear assessment of 4 features which have the highest loadings in PC 2 and are responsible for separating the Group SMet1.5 in DS2. These features were removed from the data set by MSlinearR because of their non-linearity. S1 = MeOH3, S2 = MeOH1.5, S3 = MeOH9, S6 = MeOH_EtOH3.

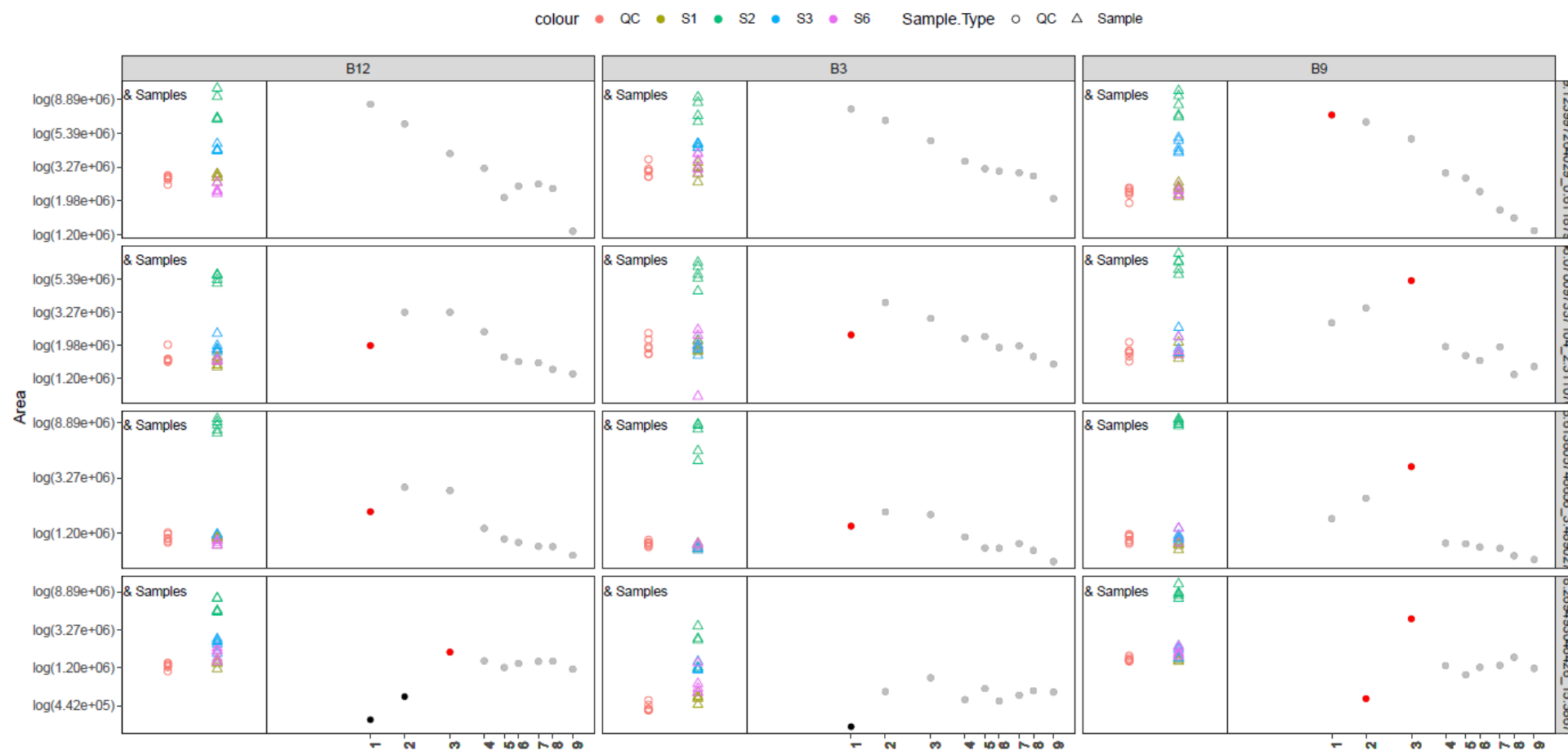


Figure S-3: Median RSD comparison of DS2 without and with using MSlineaR
Median RSD comparison of DS2 without and using MSlineaR colored according to extraction methods.

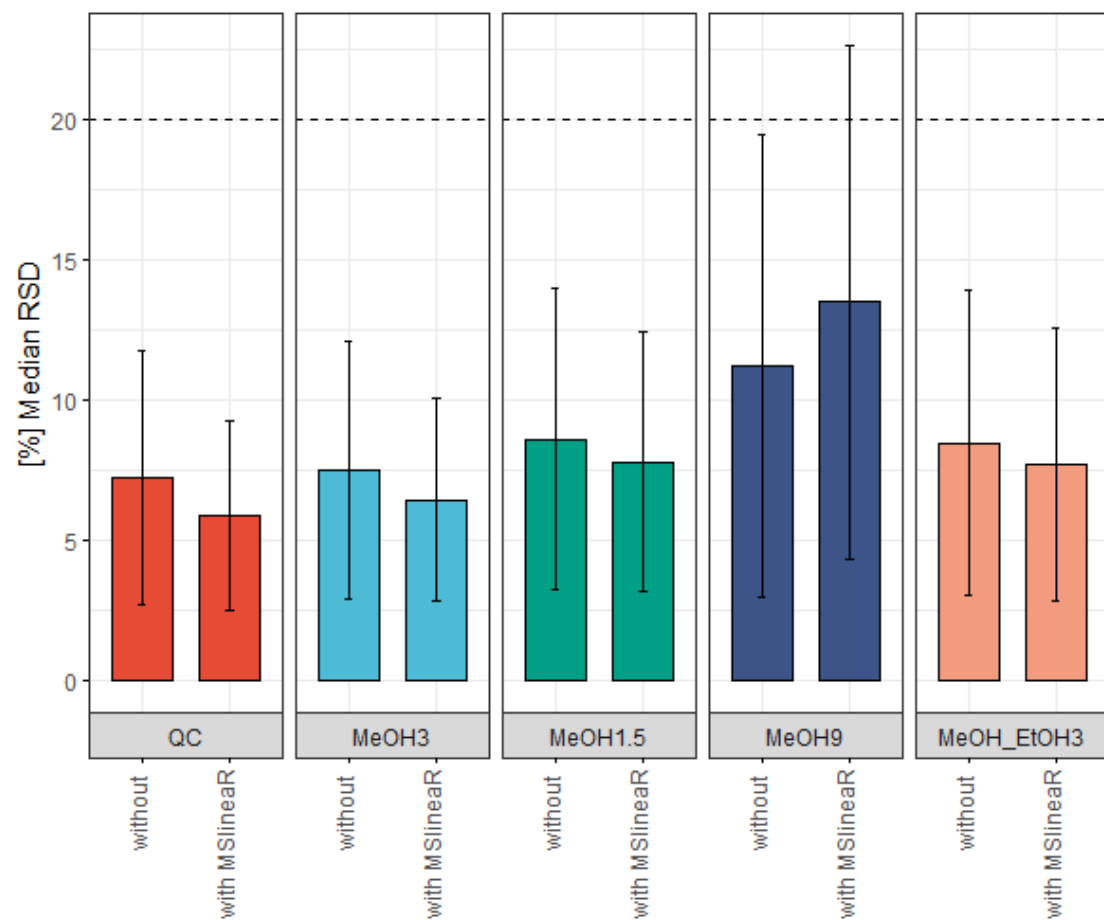


Figure S-4: Example of unexpected ion behavior at high dilutions

High dilution rates can occasionally cause unexpected behavior in ions. Internal standards added at the same concentration to all samples should have a uniform signal across the dilution range. In this example, the C12 ion increases as expected in concentration regardless of whether it is in a diluted or very diluted (1:3) matrix. By contrast, the C13 internal standard behaves as anticipated at standard dilution rates (with some ion suppression in the upper ranges), but shows an unexpected slope when dilution is very high. The reason for this phenomenon is still being explored, but may be explained by some mild ion enhancement being afforded by other ions in the matrix.

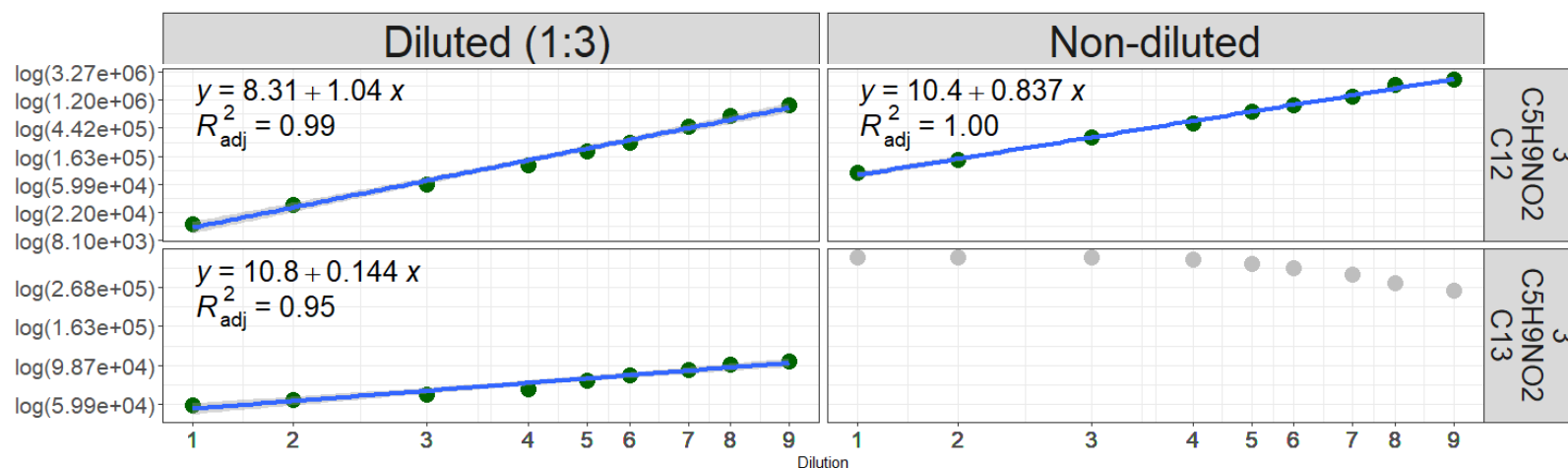


Figure S-6: Venn diagram depicting the overlap in linearity of 126 IROA standards across three batches from DS2

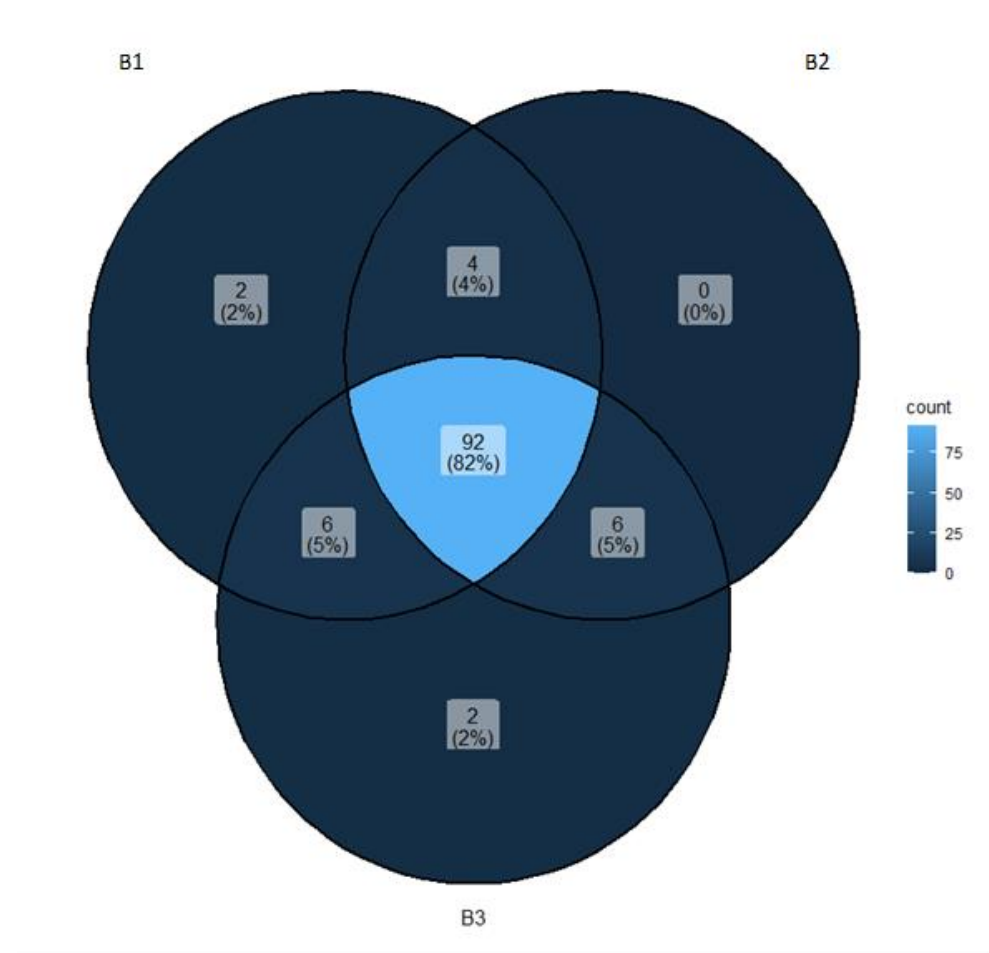


Figure S-7: Three dilution curves from DS2 illustrating examples of features which were removed by MSlineaR.

Examples of features which were removed by MSlineaR. Linearity is prevented by (A) high noise level, (B) non consistent dilution curve shape and (C) steep gradient, leading to more dilution points undergoing saturation. S1 = MeOH3, S2 = MeOH1.5, S3 = MeOH9, S6 = MeOH_EtOH3.

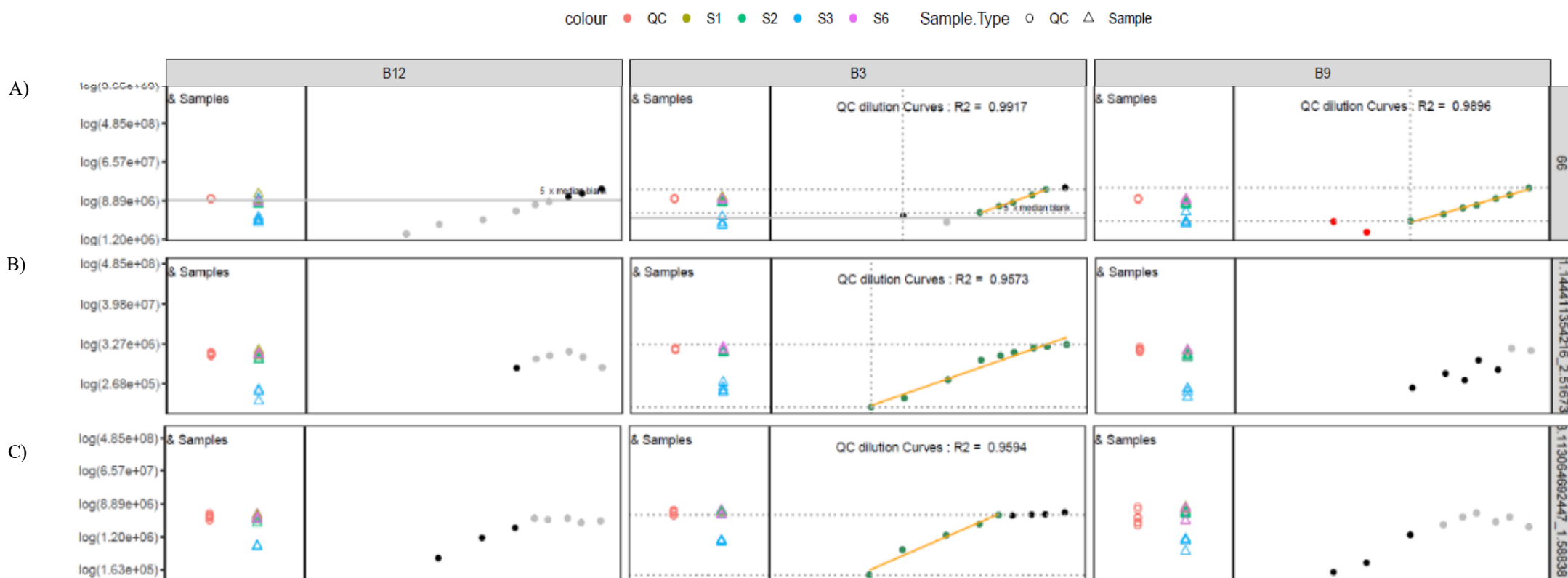


Table S-1: MetaData template for running MSlineaR

Green columns are required.

Sample_ID	Injection_order	Sample.Type	Batch	Dilution	Age_diagnosis	Sex	Smoking status
CHP_Blank_01	1	Blank	Batch1	NA	NA	NA	NA
CHP_Blank_02	2	Blank	Batch1	NA	NA	NA	NA
CHP_Blank_03	3	Blank	Batch1	NA	NA	NA	NA
CHP_Blank_04	4	Blank	Batch1	NA	NA	NA	NA
CHP_Blank_05	5	Blank	Batch1	NA	NA	NA	NA
CHP_Blank_06	11	Blank	Batch1	NA	NA	NA	NA
QC_standard_01_01	17	Reference QC	Batch1	NA	NA	NA	NA
QC_standard_01_02	18	Reference QC	Batch1	NA	NA	NA	NA
QC_pool_01	19	Pooled QC	Batch1	NA	NA	NA	NA
S12	20	Calibration Standard	Batch1	1	NA	NA	NA
S11	21	Calibration Standard	Batch1	3	NA	NA	NA
S10	22	Calibration Standard	Batch1	9	NA	NA	NA
S09	23	Calibration Standard	Batch1	27	NA	NA	NA
S08	24	Calibration Standard	Batch1	81	NA	NA	NA
S07	25	Calibration Standard	Batch1	243	NA	NA	NA
S06	26	Calibration Standard	Batch1	729	NA	NA	NA
S05	27	Calibration Standard	Batch1	2187	NA	NA	NA
S04	28	Calibration Standard	Batch1	6561	NA	NA	NA
S03	29	Calibration Standard	Batch1	19683	NA	NA	NA
X193	35	Sample	Batch1	NA	72.7	male	exsmoker
A282	36	Sample	Batch1	NA	60.7	male	active smoker
X146	37	Sample	Batch1	NA	65.4	male	neversmoker
A312	38	Sample	Batch1	NA	57.9	male	active smoker
A897	39	Sample	Batch1	NA	54.2	male	exsmoker
X116	40	Sample	Batch1	NA	62.2	male	active smoker

Table S-2: SampleData template for running MSlineaR

Sample_ID	Feature_ID	Batch	Area	Sample.Type	Injection_order	expected_conc
X193	Compound_01	Batch1	487279	Sample	35	
X193	Compound_02	Batch1	1383062	Sample	35	
X193	Compound_03	Batch1	76189	Sample	35	
A282	Compound_01	Batch1	504242	Sample	36	
A282	Compound_02	Batch1	1177905	Sample	36	
A282	Compound_03	Batch1	67591	Sample	36	
A282	Compound_04	Batch1	11780	Sample	36	
S03	Compound_01	Batch1	281854112	Calibration Standard	29	888.9
S03	Compound_02	Batch1	7924753	Calibration Standard	29	666.5
S03	Compound_03	Batch1	61782371	Calibration Standard	29	444.4

required, must match column names in metaData

required

only necessary for targeted analysis

Table S-3: Parameters for running MSlineaR

The function provides the possibility to disable each step separately and to change the settings for each step.

mandatory	parameter	explanation
	Input data	
yes	<i>analysisType</i>	Choose “targeted” or “untargeted”
yes	<i>inputData_feature</i>	Data table or data frame with at least six columns, see Table S-2
yes	<i>inputData_sample</i>	Data table or data frame with at least five columns, see Table S-1
yes	<i>column_sampleType</i>	Column name in inputData_sample , which indicates the sample types of the data set (default = “Sample.Type”). At least sampleType_serial needs to be provided.
yes	<i>sampleType_serial</i>	Sample type in column_sampleType denoting serial diluted or concentrated samples (default = “Dilution”)
yes	<i>sampleType_sample</i>	Sample type in column_sampleType denoting measured (biological) samples (default = “Sample”, disabled = NULL)

yes	<i>sampleType_blank</i>	Sample type in column_sampleType denoting blank sample i.e. the sample which will be used for the signal to noise calculation (default = “Blank”, disabled = NULL). This parameter can include a vector with more sample types, e.g. c(“ProcessBlank”, “SolventBlank”). Each subtype of blank sample will then be used for signal to noise determination.
yes	<i>sampleType_QC</i>	Sample types in column_sampleType denoting QC samples (default = “QCpool”, disabled = NULL). This parameter could include a vector with more sample types. All sub-types will be only used for visualization.
yes	<i>column_sampleID</i>	Column name in inputData_feature and inputData_sample for sample identifier (default = “Sample.Identifiaction”). In this column unique identifier per sample and batch is required.
yes	<i>column_featureID</i>	Column name in inputData_feature indicating a unique feature identifier per Feature (default = “CompoundID”), e.g.in our column Compound_ID we usually use a unique Feature ID ,paste together with mz and rt of the feature e.g. F15_144.68_4.67
yes	<i>column_batch</i>	Column name in inputData_feature and inputData_sample indicating a unique Batch identifier (default = “Batch”)
yes	<i>column_X</i>	Column name in inputData_sample indicating the single dilution of the serial diluted samples (default = “Dilution”)

yes	<i>column_Y</i>	Column name in inputData_feature indicating the Instrument response signal, e.g. Area under curve, Intensity (default = “Area”)
yes	<i>column_sampleClass</i>	optional Column name in inputData_sample indicating the statistical class for the measured samples (default = “Group”, disabled = NULL)
yes	<i>column_injectionOrder</i>	Column name in inputData_sample indicating the sequence order of the samples (default = “Sequence.Position”). This is used for arranging the data for visualization. Sequence_order must be numeric, but does not need to be consecutive and does not need to start with 1.
	<i>n_core</i>	How many cores should be used for parallel processing (default = 1)
	Signal to noise	The signal to noise assessment will be performed on the non transformed data if the parameters signal_blank_ratio and sampleType_blank area are not NULL.
	<i>signal_blank_ratio</i>	The ratio which will be used for signal to blank assessment. (default = 5, disabled = NULL). According to EMA guidelines 2012, the signal should be at least 5 times the signal of a blank sample.
	First outlier detection	
	<i>first_outlier_detection</i>	Should the outlier detection be performed? (default = TRUE, disabled = FALSE)

	<i>FOD_model</i>	Which model should be used for the outlier detection? Currently supported are linear regression (“linear”), quadratic regression (“quadratic”) and logistic regression (“logistic”). Default is a vector with all three models (default = c(“logistic”, “linear”, “quadratic”).
	<i>FOD_sdres_min</i>	Minimum residual standard deviation of a statistical model. If the residual standard deviation is below this value, there will be no outlier detection performed for this signal. (default = 1).
	<i>FOD_stdres_max</i>	Maximum value for standardized residuals of a statistical model. If a standardized residual is above this value, this point will be considered as outlier and removed from further processing. (default = 2).
	Curve trimming	
	<i>trimming</i>	Should the curve trimming function be performed? (default = TRUE, disabled = FALSE)
	Second outlier detection	
	second_outlier_detection	Should the outlier detection be performed? (default = TRUE, disabled = FALSE)
	<i>SOD_model</i>	Which model should be used for the outlier detection? Currently supported are linear regression (“linear”), quadratic regression (“quadratic”) and logistic regression (“logistic”). Default is a vector with all three models (default = c(“logistic”, “linear”, “quadratic”).

	<i>SOD_sdres_min</i>	Minimum residual standard deviation of a statistical model. If the residual standard deviation is below this value, there will be no outlier detection performed for this signal. (default = 1).
	<i>SOD_stdres_max</i>	Maximum value for standardized residuals of a statistical model. If a standardized residual is above this value, this point will be considered as outlier and removed from further processing. (default = 2).
	Determine linearity	
	<i>min_feature</i>	Minimal number of consecutive signals present per dilution/concentration curve, to consider this curve as linear (default = 6, according to EMA guidelines 2012)
	<i>R2_min</i>	Minimum coefficient of determination, to flag the feature as linear (default = 0.9).
	<i>LR_sd_res_factor</i>	maximum value for residuals of the linear model. If residuals above this value, this point will be considered as non linear. (default = 3)
	Output	
yes	<i>output_name</i>	name of the output file
yes	<i>output_dir</i>	directory where the output files will be saved

Table S-4: Mathematical Formulation of the MSlineaR Algorithm

The MSlineaR algorithm combines model fitting, residual diagnostics, and slope-based trimming to identify a stable response range for each feature. This framework integrates local and global criteria to characterize dilution-dependent signal behavior.

Notation and Data Structure	For each feature, a dilution series consisting of n observations is defined as: $S = \{(x_i, y_i), i = 1, \dots, n\}$ where x_i denotes the log-transformed dilution factor and y_i the corresponding log-transformed signal intensity.
Blank-Based Thresholding	A lower intensity threshold is defined based on blank samples: $T = \text{median}(y_{\text{blank}}) \times S:N$ Only observations satisfying: $y_i \geq T$ are retained for further analysis.
Model Fitting and Selection	Each dilution curve is approximated using three candidate models: Linear model: $f(x) = \beta_0 + \beta_1 \cdot x$ Quadratic model: $f(x) = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot x^2$ Log-logistic model: $f(x) = \frac{d}{\left(1 + \left(\frac{x}{e}\right)^b\right)}$ Model performance is evaluated using the root mean square error (RMSE):

	$\text{RMSE} = \sqrt{\left(\frac{1}{n} \times \sum (y_i - \hat{y}_i)^2\right)}$ The model with the lowest RMSE is selected
Outlier Detection	Residuals are defined as: $r_i = y_i - \hat{y}_i$ Standardized residuals are computed as: $r_i^* = \frac{r_i}{(\hat{\sigma} \times \sqrt{1 - h_i})}$ Observations are classified as outliers if: $r_i^* > \tau$, with default $\tau = 2$.
Plateau Detection via Local Slopes	Local slopes between adjacent dilution points are defined as: $m_i = \frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)}$ A reference slope is computed over a subset W: $m_{ref} = \frac{1}{ W } \sum_{i \in W} m_i$ Plateau regions are defined by: $m_i < \alpha * m_{ref}$, with $\alpha = 0.5$.
Iterative Trimming Procedure	Initial index set: $I^0 = \{1, \dots, n\}$ Iterative update: $I^{k+1} = I^k \setminus \{\text{indices failing criteria}\}$ Convergence condition: $I^{k+1} = I^k$

Final Linear Fit and Diagnostics	Linear model: $y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i$ Studentized residuals: $t_i = \frac{r_i}{(\hat{\sigma}_i \times \text{sqrt}(1 - h_i))}$ Retention criterion: $t_i \leq \tau_t$, with $\tau_t = 3$. Cook's distance: $D_i = \left(\frac{r_i^2}{(p \times \hat{\sigma}^2)} \right) \times \left(\frac{h_i}{(1 - h_i)^2} \right)$ Retention criterion: $D_i < \frac{4}{(n - p - 1)}$
Final Feature Acceptance Criteria	A feature is retained if: 1. Minimum number of points: $I \geq n_{min}$ 2. Coefficient of determination: $R^2 \geq R_{min}^2$

Table S-5: Metrics assessed by MSlineaR

Criterion	Metric	Interpretation	limitation
Monotonicity	Spearman correlation (ρ)	- Indicates whether the response is strictly monotonic across the dilution series - Reflects consistency in the direction of change across dilution levels - Captures rank-based trends independent of absolute signal magnitude	- Is sensitive to outliers - Does not capture magnitude of change (only rank order) - Can yield high values for features with a limited dynamic range
Residual based Linearity	Maximum absolute log-residual ($\Delta_{\max}$)	- Indicates the largest deviation from the fitted linear model - Reflects the maximum departure from expected model behavior - Captures worst-case inconsistency with an approximately linear response	- Can be dominated by a single outlier - May over-penalize low-intensity points, especially near limit of detection - Is sensitive to systematic curvature (e.g., ion suppression) - Depends on model specification (e.g., weighting, transformation)
	Relative deviation	- Indicates deviation from the model in relative (percentage) terms - Reflects the magnitude of error relative to expected signal intensity - Captures proportional differences between observed and predicted values	- Is sensitive to noise in data - Can inflate errors at low signal intensities, particularly near the limit of detection - Requires a reference or model-predicted value - Does not distinguish systematic bias from random noise - Is affected by heteroscedasticity
Proportionality	Slope (β_1)	- Indicates the proportional relationship between signal and dilution factor - Reflects the direction and steepness of the response - Captures overall scaling behavior across the dilution range	- Does not fully capture deviations from linearity - Depends on signal scaling and normalization - Can be biased if the regression does not pass through the origin
Goodness-of-fit	R^2	- Indicates the overall consistency of the data with the fitted model - Reflects the proportion of variance explained by the model	- Can remain high despite moderate curvature - Can be inflated by large dynamic ranges - Tends to be higher for features with greater signal variance - Ignores slope magnitude

Table S-8: preparation of dilution series from pooled QC

Dilution ID	Dilution fold of pool	Plasma volume (μL)	H2O volume (μL)	MeOH /ACN volume (μL)	Volume of supernatant to be dried in a new eppi (μL)
1	0.05	1	19	60	60
2	0.1	2	18	60	60
3	0.25	5	15	60	60
4	0.5	10	10	60	60
5	0.75	15	5	60	60
6	1 (actual QC concentration as analysed in full batch)	20	0	60	60
7	1.5	30	0	90	90
8	2	40	0	120	120
9	3	60	0	180	180

Table S-9: Sequence order data set 3

Brackets indicating repeated loops of samples and QCs.

system blank	system blank	system blank	Equilibration	Equilibration	Equilibration	Equilibration	Equilibration	QCpool	QCpool	LTRS	processed blank	processed blank	processed blank	d1	d2	d3	d4	d5	d6	d7	d8	d9	system blank	QCpool	Sample	Sample	Sample	Sample	Sample	Sample	Field_blank	Sample	QCpool_case	QCpool	Sample	...	Sample	Sample	Sample	Sample	LTRS	QCpool_control	QCpool	QCpool	system blank	system blank	system blank	system blank	system blank
--------------	--------------	--------------	---------------	---------------	---------------	---------------	---------------	--------	--------	------	-----------------	-----------------	-----------------	----	----	----	----	----	----	----	----	----	--------------	--------	--------	--------	--------	--------	--------	--------	-------------	--------	-------------	--------	--------	-----	--------	--------	--------	--------	------	----------------	--------	--------	--------------	--------------	--------------	--------------	--------------

Table S-10: Number of metabolites retained by MSlineaR for DS1 using different R^2 thresholds

N° metabolites \ $R^2 \geq$	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1
standard method	59	59	59	58	57	53	51	47	41	27	10
MSlineaR	63	63	63	62	61	57	55	49	43	28	12