

Feb 23, 2026

Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography coupled to mass spectrometry

DOI

<https://dx.doi.org/10.17504/protocols.io.eq2ly51npvx9/v1>

Warwick B Dunn^{1,2,3}, David Broadhurst^{2,4}, Paul Begley², Eva Zelena Zelena², Sue Francis-McIntyre², Nadine Anderson², Marie Brown², Joshau D Knowles⁵, Antony Halsall², John N Haselden⁶, Andrew W Nicholls⁶, Ian D Wilson⁷, Douglas B Kell², Royston Goodacre^{1,2}, The Human Serum Metabolome Consortium (HUSERMET)⁸

¹Manchester Centre for Integrative Systems Biology, Manchester Interdisciplinary Biocentre, University of Manchester, Manchester, UK;

²School of Chemistry, Manchester Interdisciplinary Biocentre, University of Manchester, Manchester, UK;

³Centre for Advanced Discoveries and Experimental Therapeutics, Manchester Biomedical Research Centre and School of Biomedicine, Manchester, UK;

⁴Department of Medicine, Katz Group Centre for Pharmacy & Health, University of Alberta, Edmonton, Alberta, Canada;

⁵School of Computer Science, The University of Manchester, Manchester, UK;

⁶Department of Investigative Preclinical Toxicology, GlaxoSmithKline, Hertfordshire, UK;

⁷Department of Clinical Pharmacology, Drug Metabolism and Pharmacokinetics, AstraZeneca, Cheshire, UK; ⁸X

Warwick B Dunn: Correspondence should be addressed to W.B.D.;

Metabolomics Journal



warwick.dunn

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.eq2ly51npvx9/v1>

Protocol Citation: Warwick B Dunn, David Broadhurst, Paul Begley, Eva Zelena Zelena, Sue Francis-McIntyre, Nadine Anderson, Marie Brown, Joshau D Knowles, Antony Halsall, John N Haselden, Andrew W Nicholls, Ian D Wilson, Douglas B Kell, Royston Goodacre, The Human Serum Metabolome Consortium (HUSERMET) 2026. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography coupled to mass spectrometry. **protocols.io**

<https://dx.doi.org/10.17504/protocols.io.eq2ly51npvx9/v1>

Manuscript citation:

Dunn, W., Broadhurst, D., Begley, P. *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).

<https://doi.org/10.1038/nprot.2011.335>

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: January 28, 2026

Last Modified: February 23, 2026

Protocol Integer ID: 242595

Keywords: Large-scale metabolic profiling, Gas chromatography, Mass spectrometry (MS), Data acquisition, Quality assurance, metabolomic study, metabolic profiling, scale metabolomic study, based metabolic profiling, scale metabolic profiling, mass spectrometry metabolism, scale metabolic profiling of serum, thousands of metabolite, metabolome, metabolic change, mass spectrometry, using gas chromatography, gas chromatography, metabolite, applying gas chromatography, chromatography, thousands of human sample, multiple analytical batches over many month, based robust loess signal correction, multiple analytical batch, robust loess signal correction, human sample

Funders Acknowledgements:

BBSRC

Grant ID: BB/C008219/1

Abstract

Metabolism has an essential role in biological systems. Identification and quantitation of the compounds in the metabolome is defined as metabolic profiling, and it is applied to define metabolic changes related to genetic differences, environmental influences and disease or drug perturbations. Chromatography–mass spectrometry (MS) platforms are frequently used to provide the sensitive and reproducible detection of hundreds to thousands of metabolites in a single biofluid or tissue sample. Here we describe the experimental workflow for long-term and large-scale metabolomic studies involving thousands of human samples with data acquired for multiple analytical batches over many months and years. Protocols for serum- and plasma-based metabolic profiling applying gas chromatography–MS (GC-MS) are described. These include sample preparation, data acquisition, data pre-processing and quality assurance. Methods for quality control–based robust LOESS signal correction to provide signal correction and integration of data from multiple analytical batches are also described.

Guidelines

The protocols described here were developed in 2009–2010. Some procedures related to instrument setup and operation should be assessed as whether they are appropriate for newer GC-MS instruments for application of these protocols.

Materials

REAGENTS

Human blood sample for serum or plasma or plasma

! CAUTION Adhere to all relevant ethical regulations and guidelines for the collection and use of human blood.

! CAUTION To avoid potential contact with bloodborne pathogens, perform all work with appropriate personal protection equipment including gloves and glasses.

 Methanol suitable for HPLC $\geq 99.9\%$ **Merck MilliporeSigma (Sigma-Aldrich) Catalog #34860**

! CAUTION Methanol is toxic and highly flammable and should be handled in a fume hood.

 Water HPLC Plus, suitable for HPLC, suitable for SM 4500 - NH₃ **Sigma aldrich.com Catalog #34877**

 Acetonitrile **Merck MilliporeSigma (Sigma-Aldrich) Catalog #34851**

! CAUTION Acetonitrile is harmful and highly flammable and should be handled in a fume hood.

 n-Hexane, 96+%, Extra Dry, AcroSeal™ **Acros Organics Catalog #326920010**

! CAUTION Hexane is harmful and highly flammable and should be handled in a fume hood.

Pyridine (Acros Organics, cat. no. 131980010)

! CAUTION Pyridine is harmful and highly flammable and should be handled in a fume hood.

Formic acid (VWR, cat. no. 450122M)

! CAUTION MSTFA is an irritant and flammable and should be handled in a fume hood.

O-methoxylamine hydrochloride (Acros Organics, cat. no. 21049025C)

! CAUTION O-methoxylamine hydrochloride is corrosive and harmful.

 Helium (99.9999%, BOC Specialty Gases, cat. no. 285365) **! BOC Catalog #285365-L**

! CAUTION Helium is an asphyxiant.

Nitrogen (99% + , supplied by a nitrogen generator)

! CAUTION Nitrogen is an asphyxiant.

 Malonic acid-d₄ **Merck MilliporeSigma (Sigma-Aldrich) Catalog #175854**

 Succinic acid-2,2,3,3-d₄ **Merck MilliporeSigma (Sigma-Aldrich) Catalog #293075**

 Glycine-d₅ **Merck MilliporeSigma (Sigma-Aldrich) Catalog #175838**

Citric acid d₄ (Cambridge Isotope Laboratories, cat. no. DLM-3487-0)

L-tryptophan d₅ (Cambridge Isotope Laboratories, cat. no. DLM-1092-0)

L-lysine d₄ (Cambridge Isotope Laboratories, cat. no. DLM-2640-0)

L-alanine d₇ (Cambridge Isotope Laboratories, cat. no. DLM-251-0)

Stearic acid d₃₅ (Cambridge Isotope Laboratories, cat. no. DLM-379-0)

Benzoic acid d₅ (Cambridge Isotope Laboratories, cat. no. DLM-122-0)

Octanoic acid d₁₅ (Cambridge Isotope Laboratories, cat. no. DLM-619-0)

Sodium chloride (VWR, cat. no. 452142B)

Docosane (Sigma-Aldrich, cat. no. D4509)

! CAUTION Docosane is an irritant and should be handled in a fume cupboard.

 Nonadecane **Sigma aldrich.com Catalog #N28906**

! CAUTION Nonadecane is an irritant and should be handled in a fume cupboard.

 Decane **Sigma aldrich.com Catalog #D90-1**

! CAUTION Decane is an irritant and should be handled in a fume cupboard.

 Dodecane **Sigma aldrich.com Catalog #D22110-4**

! CAUTION Dodecane is an irritant and should be handled in a fume cupboard.

Pentadecane (Sigma-Aldrich, cat. no. P-7385)

! CAUTION Pentadecane is an irritant and should be handled in a fume cupboard.

Dodecanamide (Sigma-Aldrich, cat. no. S408344)

 N-Benzoyl-L-leucine **Sigma aldrich.com Catalog #75813**

 21-Hydroxyprogesterone **Sigma aldrich.com Catalog #D6875**

 Cortisone **Sigma aldrich.com Catalog # C2755**

 L-Thyroxine **Sigma aldrich.com Catalog #T2376**

 Epitestosterone **Sigma aldrich.com Catalog #E5878**

! CAUTION Epitestosterone is harmful.

 N-BENZOYL-D-PHENYLALANINE **Sigma aldrich.com Catalog #S782971**

EQUIPMENT

- GC-TOF-MS—autosampler (Gerstel MPS-2L, Gerstel) and gas chromatograph (Agilent 6890 GC with split/splitless injector, Agilent) coupled to a TOF mass spectrometer (LECO Pegasus III, LECO). All instruments are controlled through a single software package (LECO ChromaTOF software, v2.x or greater)
- GC column (VF-17MS column, 0.25 mm ID × 30 m × 0.25 µm film thickness or similar, Varian, cat. no. CP8982)
- Low pressure drop liner with wool GC liner (Thames Restek, cat. no. RE20994)
- GC vials and inserts (2 ml vials with screw caps; Thames Restek, cat. nos. RE21142 and RE21723) and 200 µl vial inserts (Fisher Scientific, cat. no. VGA-100-504D)
- LECO ChromaTOF software (v2.x or greater for instrument control, data acquisition and GC-MS data processing; LECO)
- Heater block (Techne Dri-Block heaters, Thermo Fisher, cat. no. BLD-715010G) and aluminum alloy blocks (Thermo Fisher, cat. no. BLD-715-010G)
- Vortex mixer (Thermo Fisher, cat. no. MPR-558-010F)
- Centrifuge (Thermo Fisher, cat. no. CFA-114-010G)
- Refrigerator (4 °C)
- Freezer (– 80 °C)
- Centrifugal vacuum evaporator (Fisher Scientific, cat. no. DTF-600-010L)
- Centrifuge tubes (15 ml, Fisher Scientific, cat. no. FB55951)
- Centrifuge tubes (2 ml, Fisher Scientific, cat. no. TUL-150-370W)
- Microcentrifuge tubes (Eppendorf or equivalent)
- R statistical scripting language (version 2.6.0)

Before start

A. During long-term studies in which a single block of 120 samples is analyzed every week, a representative indication of the quality of data acquired from an instrument can be observed from the data acquired in the previous set of sample injections. Load data for five QC samples from the start, middle and end of the last day of the previous week and check that peak widths, heights, retention times and chromatographic resolution do not vary significantly. Check the instrument's log file and confirm that no malfunctions (indicated by red bullet points) have been recorded. Also check the instrument log book for any observed changes or errors in performance or operation. If any are found, either rectify or defer analysis until the instrument is serviced.

B. If the data are reproducible, replace the septum, inlet liner and gold seal. Remove 2–5 cm from the front of the GC column and re-install. These components collect the contaminating components of the sample upon injection and therefore require regular replacement. Check the autosampler syringe and needle for damage or jamming, and clean, replace or repair if necessary. With the 'clean syringe' command on the Gerstel MPS-2L control panel, check that the syringe draws solvent correctly without entraining air bubbles. Confirm that the vacuum is at 3×10^{-7} Torr or lower; if not, find and rectify leaks in the mass spectrometer housing and column connections. Using the full diagnostics procedure, you should obtain a combined air leak and calibration gas mass spectrum, and then confirm that the nitrogen peak height at m/z 28 is < 15% of the m/z 69 peak height, and oxygen < 5% of the m/z 69 peak height.

C. On the first day of a block analysis, the instrument should be tuned. For the LECO Pegasus III mass spectrometer described in this protocol, perform an 'Acquisition System Adjust', 'Filament Focus', 'Ion Optics Focus' and 'Mass Calibration' set of tuning and mass calibration operations. Set the detector voltage in the mass spectrometer method to a value 50 V greater than the tune file detector voltage. Replace the wash solvents (pyridine) and dispose of waste solvent.



Prepare internal standard solution

- 1 MSG IS1: Accurately weigh and record 10.0 ± 0.5 mg quantities of malonic acid d_2 , succinic acid d_4 and glycine d_5 into a single 15 mL centrifuge tube, add 10 mL of water and vortex mix for 1 min to provide full dissolution. Label as MSG IS1.
- 2 CFT IS1: Accurately weigh and record 10.0 ± 0.5 mg quantities of citric acid d_4 , d-fructose $^{13}C_6$ and L-tryptophan d_5 into a single 15 mL centrifuge tube, add 10 mL of water and vortex mix for 1 min to provide full dissolution. Label as CFT IS1.
- 3 LA IS1: Accurately weigh and record 10.0 ± 0.5 mg quantities of L-lysine d_4 and L-alanine d_7 into a single 15 mL centrifuge tube, add 10 mL of water and vortex mix for 1 min to provide full dissolution. Label as LA IS.
- 4 BO IS1: Accurately weigh and record 10.0 ± 0.5 mg quantities of stearic acid d_{35} , benzoic acid d_5 and octanoic acid d_{15} into a single 15 mL centrifuge tube, add 10 mL of methanol and vortex mix for 1 min to provide full dissolution. Label as SBO IS1.
- 5 IS2 Solution: A working internal standard solution 'IS2' is prepared fresh each day by combining 2 mL aliquots of each of the four IS1 stock solutions (MSG IS1, CFT IS1, LA IS1 and SBO IS1) and adding 4.0 mL of water to produce a final volume of 12.0 mL. The nominal concentration of each component is 0.167 mg.mL^{-1} . All solutions are stored at 4°C and must be prepared fresh every week.

Prepare Retention Index (RI) solution

- 6 RI1 solution: Accurately weigh 30 mg (± 3 mg) each of docosane and nonadecane into a 15 mL centrifuge tube. Add 40 μL each of decane, dodecane and pentadecane. The tube must be weighed after addition of all five alkanes and the weight of each component recorded, this should be 30 mg (± 5 mg) each. Add 10 mL hexane to form the retention index marker solution 1 (RI1).
- 7 RI2 solution (working solution): The working retention index solution 2 (RI2) is prepared by adding 2.0 mL RI1 to 8.0 mL pyridine. RI2 may be stored in a sealed container at 4°C for up to 4 weeks.

Prepare serum and plasma samples: 2–3 h

2h

- 8 Allow plasma/serum samples to thaw  On ice at  4°C for  00:30:00 –  01:00:00 . 



Prepare serum and plasma samples: 2–3 h

2h

- 9 Aliquot  400 μL of plasma/serum into a labeled 2.0-mL microcentrifuge tube and add  200 μL of internal standard solution (IS2) and then  1200 μL of methanol. 
- 10 Thoroughly mix on a vortex mixer for  00:00:15 and pellet the protein precipitate in a centrifuge operating at  Room temperature and at  15800 x g, 00:15:00 .   
- 11 Transfer  370 μL aliquots into four separate labeled 2.0-mL microcentrifuge tubes and dry down (lyophilize) each sample in a centrifugal vacuum evaporator for  18:00:00 . Apply no heating during the drying process.  

Note

PAUSE POINT: Store the samples at  4 °C for up to 12 weeks.

Prepare QC samples: 2–3 h

2h

- 12 Allow plasma/serum samples to thaw  On ice at  4 °C for  00:30:00 –  01:00:00 . 
- 13 Aliquot  400 μL of plasma/serum into a labeled 2.0-mL microcentrifuge tube and add  200 μL of internal standard solution (IS2) followed by 1,200 μL methanol. 
- 14 Thoroughly mix on a vortex mixer for  00:00:15 and pellet the protein precipitate in a centrifuge operating at  Room temperature and at  15800 x g, 00:15:00 .   
- 15 Transfer  370 μL aliquots into four separately labeled 2.0-mL microcentrifuge tubes and dry down (lyophilize) each sample in a centrifugal vacuum evaporator for ~18 h. Apply no heating during the drying process.   

Note

PAUSE POINT: Store the samples at 4 °C for up to 12 weeks.

Prepare saline blank samples: 15 min

15m

- 16 Aliquot  100 μL of 0.7% (wt/vol) sodium chloride into a 2.0-mL microcentrifuge tube and add  50 μL of internal standard solution followed by  300 μL methanol. 
- 17 Thoroughly mix on a vortex mixer for  00:00:15 and dry down (lyophilize) each sample in a centrifugal vacuum evaporator for ~18 h. Apply no heating during the drying process.  

Note

PAUSE POINT: Store the samples at 4 °C for up to 12 weeks.

Chemical derivatization for GC-MS analysis: 45–60 min

45m

- 18 Lyophilize dried biological, QC and saline blank samples for  01:00:00 ; switch on the Dri-Block heater and allow it to reach a set-point temperature of  80 °C . 
- 19 Add  50 μL of a 20 mg.mL⁻¹ O-methoxylamine in pyridine solution to the dried extract, thoroughly mix for  00:00:15 on a vortex mixer, and then heat in the Dri-Block heater at  80 °C for  00:15:00 .   
- 20 Remove samples from block heater, add  50 μL of MSTFA to each solution, vortex for  00:00:15 and heat in a block heater at  80 °C for  00:15:00 .   
- 21 Remove samples from the block heater and allow them to cool for  00:05:00 . To each sample, add  20 μL of working retention index solution (RI2) and vortex for  00:00:15 . 
- 22 Centrifuge each sample at  15800 x g, 00:15:00 and transfer  100 μL of the supernatant to a 200- μL vial insert placed in a 2-mL vial; seal with a screw cap.  

GC-TOF-MS analysis: 30 min per sample

- 23 Analyze samples applying the following instrument parameters. A volume of 1 μL of derivatized sample solution is injected through a split/splitless injector operating at a 

temperature of $280\text{ }^{\circ}\text{C}$, at a split ratio of 4:1 and with a helium carrier gas flow rate of $1\text{ mL}\cdot\text{min}^{-1}$ in constant flow mode. Chromatographic separations are applied as described below.

During sample analysis, chromatographic separations are performed on a Varian VF-17MS column. Gas saver flow ($25\text{ mL}\cdot\text{min}^{-1}$) is switched on 15 s after sample injection. The temperature program begins at $70\text{ }^{\circ}\text{C}$ with a hold time of 4 min, followed by a linear temperature ramp of $20\text{ }^{\circ}\text{C}$ per min up to $300\text{ }^{\circ}\text{C}$, followed by a hold time of 4 min. The oven temperature is then allowed to cool to $70\text{ }^{\circ}\text{C}$ before the next injection. The transfer line temperature is held at $240\text{ }^{\circ}\text{C}$. The mass spectrometer source is operated at a temperature of $250\text{ }^{\circ}\text{C}$ in EI mode, with an electron energy of 70 eV. Data are acquired over the range of m/z of 45–600, at an acquisition rate of 20 Hz. The detector is operated in the range 1,400–1,800 V, typically 50 V greater than the voltage determined during the LECO-defined tuning checks.

- 24 At the end of each analytical batch, assess six metabolites (lactic acid, alanine, glutamine, fructose, tryptophan and octadecanoic acid). Ensure that the peak shapes, peak heights and retention times are reproducible with no systematic drift.

Data preprocessing: MS 6–9 h

- 25 Preprocess the data by following the steps for GC-TOF-MS data.
 - **GC-TOF-MS data analysis [TIMING 4–6 h for target list generation, 2–3 h for raw data processing of data acquired over 5 d]**
- 26 Using LECO's terminology, perform a 'peak find' data processing method with a single QC sample injected in the middle of the block experiment. The data processing method should have 'Baseline', 'Peak Find', 'Calculate Area/ Height' and 'Retention Index' functions activated. Key parameters in this method are the baseline offset, data points to be averaged for smoothing, expected chromatographic peak width, maximum number of unknown peaks to find and the minimum signal-to-noise ratio for the (automatically selected) quantitation mass. All parameters are sensitive to the chromatographic performance obtained and must be selected to reflect this. From representative chromatograms acquired in the HUSERMET project, in which we analyzed thousands of human serum samples with GC-MS, baseline offset was set at 0.5, data points to be averaged for smoothing was set at automatic, peak width was set at 1.8 s and the maximum number of unknown peaks to find was set to 400. A signal/noise (S/N) threshold of 100:1 was used; this was an informed compromise between comprehensive reporting and the collation of spectra of sufficient quality to be reliably found subsequently. A retention index method is prepared in the software by compiling a method table containing the retention indices (1,000, 1,200, 1,500,

1,900 and 2,200), the observed retention time and the quantitation ions used to confirm the detection of each retention index compound.

27 Step 27 produces a table of potential candidates for inclusion in a reference table and annotated with a retention index, mass spectrum and single quantitation ion. From this table, delete candidates whose mass spectrum does not contain fragment ions expected for TMS derivatives at m/z 73 and 147, and whose quantitation ion chromatogram indicates that a single mass spectral feature has been reported as multiple features ('peak splitting'). In these cases, delete the features with lowest S/N while retaining the feature with the highest S/N. Manually edit the mass spectrum for the isotopically labeled internal standards to remove ions present in the unlabeled endogenous metabolite. Assess the automatically chosen quantitation masses for accuracy, a high S/N ratio and no interference to peak shape from co-eluting derivatized metabolite peaks. Amend the quantitation mass if necessary. The metabolite peaks are then exported to a reference file created before Step 27. Parameters in the reference table are set at 100,000,000 for tolerance (to ensure all peaks are matched and reported independent of peak area), 20 for RI deviation, 700 for match threshold, 2,500 for minimum area and 5.0 for S/N threshold.

28 A separate study sample can then be processed through the deconvolution software, as described in step 27, with the 'Compare' function also enabled. To do this, set the mass threshold setting at 50. Derivatized metabolic features uniquely detected in this sample are marked, the mass spectrum and quantitation masses are assessed as described above in Step 28 and then exported to the reference file. This process is performed for a range of samples from the study.

Note

CRITICAL STEP: In large-scale studies, we recommend performing Step 22 on samples from different experimental blocks to ensure that all derivatized metabolite peaks are present in the reference file.

29 Each peak in the reference file is named with a unique label (e.g., internal standard succinic d_4 acid, sample peak X). At this stage, definitive identification of each peak can be performed. To do this, compare the retention index and mass spectrum of each metabolite with those recorded for authentic chemical standards and present in in-house libraries (e.g., Golm metabolome database or MMD in-house library) or in commercially available mass spectral libraries (e.g., NIST or EPA libraries). If a match to a retention time/index (± 10) and mass spectrum (match > 70%) is observed, the identification can be described as definitive and the peak can be labeled metabolite name_definitive. If a match to only a mass spectrum is observed, the identification can be described as putative and the peak can be labeled 'metabolite name_putative'.





- 30 The final stage is used to define the most appropriate internal standard for each peak. This can be performed by analyzing 60 QC injections in a single block. Calculate the peak area ratio (peak area metabolite/peak area internal standard) for each metabolite peak associated with each internal standard and calculate the relative standard deviation (RSD) for each of these peaks for injections 6–60. The internal standard providing the lowest RSD is chosen as the internal standard for that metabolite.
- 31 Perform raw data processing using the reference table described above for all samples to reliably find and report the selected metabolic features in all samples. Process all the blocks using the appropriate set of parameters and internal standard selections. As noted, automatic feature detection and measurement achieves a high success rate (estimated to be in excess of 98%), which was further improved by manually inspecting the peak area measurements for each internal standard in each sample, and manually correcting where required. Further outlier rejection tests can be performed on a block basis before accepting data. This has led to the rejection of < 1% of the injections performed.

Note

PAUSE POINT: Archive processed data for future use.

Data processing, signal correction and QA procedures for multiple analytical blocks

6h

- 32 Perform data alignment and normalization for the complete data set, composed of multiple analytical blocks, as described below for GC-MS.

Data processing, signal correction and QA procedures for GC-MS data [TIMING 6–8 h for processing of data acquired over a 5-d period]

- 33 Remove data related to the first three QC sample injections in each analytical batch. Perform signal correction for each data acquired in each analytical block using the QC-RLSC method to fit a LOESS polynomial curve to the QC data for each metabolic feature. In this implementation, the local polynomials that are fitted to the data are constrained to be either first or second degree (i.e., either locally linear or locally quadratic). The polynomial is fitted using weighted least squares with a standard tri-cubic weight function. To stop overfitting, use leave-one-out cross-validation over the integer range of $n\alpha$ for each degree of polynomial ($\lambda = [1,2]$), where α is the smoothing parameter. Once the LOESS curve is fitted to the QC data, construct a correction curve for the whole analytical run using cubic-spline interpolation, to which

the total data set for that metabolic feature is normalized. **Figure 2** illustrates the QC-RLSC procedure in practice for a metabolic feature in which signal drift across a given analytical batch was observed.

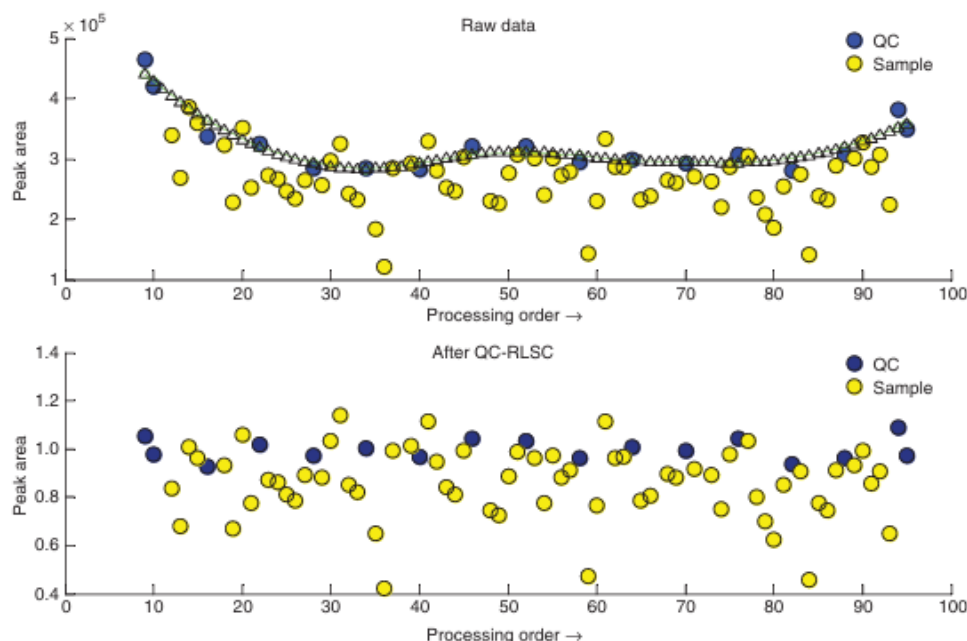


Figure 2 | The QC-RLSC protocol for a metabolic feature detected in UPLC-MS (ES +) with signal attenuation across a given analytical batch. A cross-validated LOESS curve (upper plot) is fitted to the QC samples, the correction curve interpolated (triangles), to which the total data set for that peak is corrected (lower plot).

- 34 Perform a QC procedure to remove metabolic features with poor repeatability. Data for all detected metabolic features for all QC sample injections from injection four to the last injection of the QC sample are applied. Remove all metabolic features that are detected in < 50% of QC samples and all metabolic features with a RSD, as calculated for the QC samples, of > 30%.
- 35 Combine data from the separate analytical batches for all blocks into a single data set. Include relevant information on blocks, subjects, sample types and injection order.



Acknowledgements

The human serum metabolome project (HUSERMET) is funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC) (BB/C008219/1), MRC, GlaxoSmithKline and by AstraZeneca. We thank the BBSRC and the Engineering and Physical Sciences Research Council for their financial support to The Manchester Centre for Integrative Systems Biology (BB/C008219/1). W.B.D. wishes to thank the UK National Institute for Health Research for financially supporting the Manchester Biomedical Research Centre.