

Feb 23, 2026

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

DOI

<https://dx.doi.org/10.17504/protocols.io.bp2l6em11gqe/v1>

Warwick B DUNN^{1,2,3}, David Broadhurst^{2,4}, Paul Begley², Eva 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 HUSERMET⁸

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

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.bp2l6em11gqe/v1>

Protocol Citation: Warwick B DUNN, David Broadhurst, Paul Begley, Eva 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 liquid chromatography coupled to mass spectrometry. **protocols.io**

<https://dx.doi.org/10.17504/protocols.io.bp2l6em11gqe/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 29, 2026

Last Modified: February 23, 2026

Protocol Integer ID: 242596

Keywords: LC-MS, Large-scale metabolic profiling, Serum, Plasma, Sample preparation, metabolomic study, metabolic profiling, scale metabolomic study, scale metabolic profiling of serum, mass spectrometry metabolism, based metabolic profiling, scale metabolic profiling, ultraperformance liquid chromatography, applying ultraperformance liquid chromatography, mass spectrometry, thousands of metabolite, metabolome, metabolic change, liquid chromatography, using liquid chromatography, chromatography, metabolite, thousands of human sample, multiple analytical batches over many month, multiple analytical batch, human sample, serum, sample preparation, based robust loess signal correction, robust loess signal correction

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 ultraperformance liquid chromatography–MS (UPLC-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.

! CAUTION Formic acid is corrosive and volatile, and should be handled in a fume hood.

! CAUTION Nitrogen is an asphyxiant.

☒ Leucine Enkephalin acetate salt hydrate **Merck MilliporeSigma (Sigma-Aldrich) Catalog #L9133**

☒ 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)

d-fructose ¹³C₆ (Cambridge Isotope Laboratories, cat. no. CLM-1553-0)

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)

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

UPLC-TOF-MS—Autosampler and Ultra-Performance Liquid Chromatograph (Waters Acquity system, Waters) coupled to a TOF mass spectrometer (Waters LCT mass spectrometer range, Waters). All instruments are controlled through a single software package (Waters MassLynx v3.x or greater).

UPLC-LTQ-Orbitrap-MS—Autosampler and Ultra-Performance Liquid Chromatograph (Waters Acquity system, Waters) coupled to a hybrid linear ion-trap-Orbitrap XL mass spectrometer (ThermoFisher Scientific LTQ-Orbitrap range, ThermoFisher Scientific). All instruments are controlled through a single software package (ThermoFisher Scientific XCalibur, v2.x or greater).

UPLC column—ACQUITY UPLC BEH C18 2.1 mm × 100 mm, 1.7 μm (Waters, cat. no. 176000864) with precolumn in-line filter (Waters, cat. no. 700002775)

UPLC vials—total recovery screw cap glass vials (Waters, cat. no. 186000385C)

Waters MassLynx (v3.x or greater for instrument control and data acquisition; Waters)

ThermoFisherScientific XCalibur (v2.x or greater for instrument control and data acquisition; ThermoFisher Scientific)

XCMS software (v1.10 or greater) for UPLC-MS data processing

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. UPLC conditions. Chromatographic separations are performed on an ACQUITY UPLC BEH column (C18, 2.1 × 100 mm, 1.7 μm) operating at 50 °C and applying a binary mobile phase system. Mobile phase A is water containing 0.1% (vol/vol) formic acid and mobile phase B is methanol containing 0.1% (vol/vol) formic acid. Different gradient elutions are performed for positive and negative ion mode detection, with flow rates of 0.36 and 0.40 mL.min⁻¹ for positive and negative ion mode detection, respectively.

B. UPLC system preparation. At the start of an experimental block, prepare a fresh set of UPLC mobile phase and needle wash solutions. Prime the system, wash the autosampler needle and sample syringe. Install a fresh, unused UPLC column and pre-line filter, and then condition the column by operating with 100% mobile phase B and starting the flow at 0.05 mL.min⁻¹; next, leave for 3 min and then increase to 0.1 mL.min⁻¹; next, leave for 3 min and then increase to 0.20 mL.min⁻¹; next, leave for 3 min and increase to the operating flow rate (0.36 or 0.40 mL.min⁻¹). Set the column temperature to 50 °C. Leave for 10 min. Change the mobile phase composition to 50% mobile phase A/50% mobile phase B and monitor the operating pressure. This should be < 12,000 p.s.i. If this is > 12,000 p.s.i., check the column, pre-line filter and post-UPLC tubing for blockages. If < 12,000 p.s.i., change the mobile phase composition to 100% mobile phase A and monitor the system pressure. This should be < 6,000 p.s.i. when stabilized. Allow the column to equilibrate for a minimum of 30 min.

C. TOF-MS system preparation 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 by > 20%. If the observed shift in retention time from the data acquired at the beginning and the end of the block vary significantly (by > 0.2 min), check the UPLC system for leaks. Check whether the vacuum pressures are as typically observed and check that all voltages are as set in the Tune page of MassLynx. The instrument logbook should be checked for any observed changes or errors in performance or operation. If any are found, either rectify immediately or defer analysis until the instrument is serviced. At the end of each analytical batch, the source is cleaned to remove residues that can reduce the instrument sensitivity. Sonicate the sample cone in a 50:50 (vol/vol) methanol/water solution containing 1% (vol/vol) formic acid for 15 min and dry with nitrogen before replacing. Infuse leucine-enkephalin solution into the mass spectrometer through the reference probe and check the MS response. The response should be < 300 counts in continuum mode for the ¹²C isotope peak in attenuated mode and for the ¹³C isotope peak in normal mode. Tune the instrument manually (using ion source parameters first and detector voltages second) until the required ion count is observed. The MS system should be calibrated according to the manufacturer's instructions with a sodium formate solution. The intensities of all calibrant peaks should be around 300 counts per scan in continuum mode (if the intensity of any of the calibrant peaks is substantially higher (by > 50%), this peak should be manually rejected from the calibration before calibration acceptance). The residual error of an acceptable calibration must be within 10 p.p.m. Switch the photo diode array lamp on for at least 1 h before data acquisition. Load the appropriate MS tune method, setting the appropriate values for the parameters on the ES – or ES + page and wait for the read-backs to reach the set values. The source and desolvation

temperatures should be set to 100 and 200 °C, respectively, for nominal mass correction and calibration. Set these temperatures to 140 and 480 °C, respectively, for sample analysis

D. LTQ-Orbitrap-MS system preparation. 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 (by > 20%). If the observed shift in retention time from the data acquired at the beginning and the end of the block vary significantly (by > 0.2 min), check the UPLC system for leaks. Check whether the vacuum pressures are as typically observed and check that all voltages are as set in the Tune page of LTQTune and green ticks are in position against each parameter. If any errors are found, either rectify immediately or defer analysis until the instrument is serviced. At the end of each analytical batch, the source is cleaned to remove residues that can reduce the instrument sensitivity. This involves sonication of the sample cone and transfer lens in a 50:50 (vol/vol) methanol/water solution containing 1% (vol/vol) formic acid for 15 min. Tune the mass spectrometer according to the manufacturer's specifications, applying the peaks at m/z 514.2 (taurocholate, negative ion) and 524.2 (MRFA peptide, positive ion). Switch the photo diode array lamp on for at least 1 h before data acquisition. The MS system should be calibrated according to the manufacturer's instructions, with a calibration solution containing SDS, sodium taurocholate, MRFA peptide, Ultramark 1621 and caffeine. If the instrument is being switched into operate mode from standby, wait a minimum of 2 h before performing a mass calibration. The intensities of all calibrant peaks should be between 104 and 107. Set the scan time to 100 ms and average three microscans. The error for all calibrant peaks should be within 5 p.p.m. If the mass error is greater, perform a second mass calibration. If the mass calibration error is still > 5 p.p.m., either rectify immediately or defer analysis until the instrument is serviced. Load the appropriate MS tuning method, setting the appropriate values for the parameters, and then wait for green ticks to be in position against each parameter.

Prepare Internal Standard solutions

- 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 $0.167 \text{ mg. mL}^{-1}$. All solutions are stored at 4°C and must be prepared fresh every week.

Prepare serum and plasma samples: 2–3 h

2h

- 6 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

- 7 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.
- 8 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 .



- 9 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

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



- 11 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 1200 μL methanol.



- 12 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 .



- 13 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

- 14 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.





- 15 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.

Prepare mobile phases: 10 min

- 16 Mobile phase A is prepared by addition of 1.0 mL of formic acid to 1,000 mL of HPLC-grade water and mixing thoroughly.

Mobile phase B is prepared by addition of 1.0 ml of formic acid to 1,000 mL of HPLC-grade methanol and mixing thoroughly.

Prepare Needle Wash solutions: 10 min

- 17 Weak needle wash solution: A weak needle wash solution is prepared by adding 50 mL of HPLC-grade methanol to 950 mL of HPLC-grade water and mixing thoroughly.

Strong needle wash solution: A strong needle wash solution is prepared by adding 800 mL of HPLC-grade methanol to 200 mL of HPLC-grade water and mixing thoroughly.

Prepare chromatography check solution: 15 min

15m

- 18 Accurately weigh 10 mg of each chemical (± 3 mg) of dodecanamide, benzoyl leucine, 11-deoxycorticosterone, cortisone, thyroxine, epitestosterone and N-benzoyl-d-phenylalanine into a 15-mL centrifuge tube. Dissolve in 10.0 mL of solvent (50:50 (vol/vol), methanol/water) to prepare a 1.0 mg.mL^{-1} solution and label the tube as CCS1.

- 19 Aliquot  2.0 mL of CCS1 solution into a 15-mL centrifuge tube, add  8.0 mL of solvent (50:50 (vol/vol) methanol/water) and thoroughly mix to prepare a solution of $200 \text{ }\mu\text{g.mL}^{-1}$. Label this as CCS2.



Note

PAUSE POINT: Store the samples at 4 °C for up to 2 weeks.

Sample reconstitution for UPLC-MS analysis: 30–45 min

30m

- 20 Add  100 µL (TOF-ES – and Orbitrap ES + and ES –) or  200 µL (TOF-ES +) of water to dried samples, vortex for  00:00:15 and centrifuge at  15800 x g, 00:15:00 .   
- 21 Transfer  90 µL (TOF-ES – and Orbitrap ES + and ES –) or  180 µL (TOF-ES +) of supernatant to low-recovery-volume 2-ml vials and seal with screw caps. 
- 22 Tap the bottom of each vial to release air bubbles present at the bottom.
- 23 Place in UPLC autosampler/sample manager operating at  4 °C . 

UPLC analysis: 22 or 24 min per sample

- 24 Inject  10 µL of reconstituted sample on to the UPLC column from sample vials stored at  4 °C . 

TOF-MS analysis: 22 or 24 min per sample

- 25 If the UPLC system is coupled to a Waters LCT TOF-MS instrument, perform the following procedures. If a different manufacturer's TOF-MS system or a different Waters TOF-MS model is used, please develop and validate the methods described, taking into account the manufacturer's instructions.
- 26 Acquire accurate mass data in 'V mode' as centroid data in the m/z range of 50–1,000 and with dynamic range enhancement activated. A scan time of 0.4s (0.35 s and 0.05 s dwell time) is applied. Half (50%) of the UPLC eluent is directed to the mass spectrometer and the other 50% is diverted to waste.
- 27 The UPLC gradient elution programmes to be applied for positive and negative ion mode are defined in the table below.

**TABLE 1** | Waters Acquity UPLC gradient elution program applied for UPLC-MS analysis for ES+ and ES− modes.

Time (min)	Flow rate (ml min ^{−1})	Mobile phase A (%)	Mobile phase B (%)	Curve
<i>Positive ion mode (ES+)</i>				
Initial	0.4	100	0	5
1.0	0.4	100	0	5
16.0	0.4	0	100	5
20.0	0.4	0	100	5
22.0	0.4	100	0	1
<i>Negative ion mode (ES−)</i>				
Initial	0.36	100	0	4
2.0	0.36	100	0	4
17.0	0.36	0	100	4
22.0	0.36	0	100	4
24.0	0.36	100	0	1

- 28 At the end of each analytical batch, assess seven metabolites (dodecanamide, benzoyl leucine, 11-deoxycorticosterone, cortisone, thyroxine, epitestosterone and N-benzoyl-d-phenylalanine). Ensure that the peak shapes, peak heights and retention times are reproducible with no systematic drift. The typical analysis order is composed of QC samples, subject samples, chromatography check solution and saline blank samples, analyzed in a predetermined order.

Note

PAUSE POINT Archive raw analytical data for future use.

LTQ-Orbitrap-MS analysis: 22 or 24 min per sample

- 29 If the UPLC system is coupled to a ThermoFisher LTQ-Orbitrap XL instrument, perform the following procedures. If a different model of the Orbitrap system

(ThermoFisher Scientific) is applied, please develop and validate the methods described, taking into account the manufacturer's instructions.

30 Acquire accurate mass data in the Orbitrap mass analyzer in the m/z range of 50–1,000 (in centroid mode), with a mass resolution of 30,000 at m/z 400 (FWHM), and with a scan time of 0.4 s. The UPLC eluent is split, with 50% directed to the mass spectrometer and 50% diverted to waste.

31 The UPLC gradient elution programmes to be applied for positive and negative ion mode are defined in the table below.

TABLE 1 | Waters Acquity UPLC gradient elution program applied for UPLC-MS analysis for ES+ and ES– modes.

Time (min)	Flow rate (ml min ⁻¹)	Mobile phase A (%)	Mobile phase B (%)	Curve
<i>Positive ion mode (ES+)</i>				
Initial	0.4	100	0	5
1.0	0.4	100	0	5
16.0	0.4	0	100	5
20.0	0.4	0	100	5
22.0	0.4	100	0	1
<i>Negative ion mode (ES–)</i>				
Initial	0.36	100	0	4
2.0	0.36	100	0	4
17.0	0.36	0	100	4
22.0	0.36	0	100	4
24.0	0.36	100	0	1

32 At the end of each analytical batch, assess seven metabolites (dodecanamide, benzoyl leucine, 11-deoxycorticosterone, cortisone, thyroxine, epitestosterone and N-benzoyl-d-phenylalanine). Ensure that the peak shapes, peak heights and retention times are reproducible with no systematic drift. The typical analysis order is composed of QC samples, subject samples, chromatography check solution and saline blank samples, analyzed in a pre-determined order.

Note

PAUSE POINT Archive raw analytical data for future use.

Data preprocessing: UPLC-MS 6–8 h

- 33 Preprocess the data by following the steps for UPLC-MS data.
- **UPLC-MS data analysis [TIMING 6–8 h for processing of data acquired over a 5-d period]**
- 34 Perform conversion of instrument-specific data format to NetCDF format. Data preprocessing is performed using the open-source XCMS software that requires data in a specific format. The NetCDF format is appropriate as many (but not all) mass spectrometer manufacturing companies provide software with instruments to convert to NetCDF format. For the Waters TOF instrument the software is called DataBridge, which operates in the MassLynx software, and for the ThermoFisher Scientific LTQ-Orbitrap instrument the software is called FileConverter, which operates in the Xcalibur software. All data can then be easily transferred as NetCDF files to the XCMS processing PC for data processing.
- 35 Perform XCMS data processing. In UPLC-MS data, semiautomated reference spectra selection and target list generation are not required. The deconvolution software performs this action in an automated manner. Data are deconvolved into a usable data matrix using XCMS. XCMS is run using the R statistical scripting language (version 2.6.0) and produces a matrix of features with associated retention time and accurate mass and chromatographic peak area calculated with a single accurate mass. The XCMS settings we applied for UPLC-LCT MS data are 'step' (0.10), 'S/N threshold' (3), 'bw' (10) and 'time limit' (15 s). All other parameters are set to 'default' settings. The XCMS settings we applied for UPLC-LTQ Orbitrap data are 'step' (0.02), 'S/N threshold' (3), 'mass limit' (0.05 amu), 'bw' (10) and 'mzwid' (0.05). All other parameters are set to 'default' settings. Preprocessed data are then exported as a .csv file for further data processing, univariate and multivariate data analysis procedures.

Note

PAUSE POINT: Archive processed data for future use.

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

- 36 Perform data alignment and normalization for the complete data set, composed of multiple analytical blocks, as described below for UPLC-MS.
- **Data processing, signal correction and QA procedures for UPLC-MS data [TIMING 6–8 h for processing of data acquired over a 5-d period]**
- 37 Remove data related to the first eight QC sample injections in each analytical batch. Perform signal correction for each analytical block using the QC-RLSC method.
- 38 Perform a QC procedure to remove metabolic features with poor repeatability. Data for all detected metabolic features for all QC sample injections from injection nine 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 > 20%.
- 39 An important consequence of separate analytical block experiments for UPLC-MS data is that metabolic features are not aligned across blocks; matching the same metabolic features across multiple blocks is required. Construct a reference database, composed of unique metabolic features—as defined by accurate mass and retention time — in a chosen subset of the analytical blocks. For example, for the HUSERMET data, three of ten analytical blocks were chosen. The reference database will contain all unique features detected in the biological experiment. The data describe the metabolite identifier, m/z and retention time, and the estimated median peak area (MPA) for the QC samples for that instrument. Subsequently, and for each block separately, match the features in each block to the reference peak database. Only match a metabolic feature if the error in m/z is less than a specific range (e.g., ± 5 p.p.m.) and the error in retention time is less than a specific range (e.g., ± 10 s). For matched metabolic features, include an identifier related to the reference peak database. For example, if a metabolic feature in block 1 matched to metabolic feature 76 in the reference database, then label this metabolic feature as 76. Following the matching of metabolic features, combine data from all blocks into a single data set.
- 40 As a secondary check, or validation, of the matching process, compare each of the peaks in the new 'matched' data set with the reference table with respect to the expected QC MPA (before QC-RLSC). Using the MPA tolerance formula: $\zeta_i = |(\text{MPA}_{b,i} - \text{MPA}_{\text{ref},i}) / \text{MPA}_{\text{ref},i}|$ where b is batch number and i is matched peak number; if a value of $\zeta < 4$ is found then the candidate peak is assumed to be within an acceptable peak area tolerance to be correctly matched.



Note

- Once the data are normalized using QC-RLSC, small differences in between-batch MPA will be corrected for. Any peak failing the final QA process is removed from the data set.

PAUSE POINT: Archive processed data for future use.

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.