LC–MS method validation.
QC and QA

Georgios Theodoridis
Aristotle University Thessaloniki, Greece

Acknowledgment

Collaborators
Prof. Ian Wilson, ICL
Assist. Prof. M. Witting, Helmholtz
Assist. Prof. A. Agapiou, Cyprus Univ.

Prof. V. Mougios
Prof. G. Arsenos, Vet.
Assoc. Prof. N. Raikos, MD
Assoc. Prof. K. Sarafidis, MD
Assoc. Prof. G. Sianos MD
Assist. Prof. G. Giannakoulas, MD
Prof. M. Rolildis, MD

Biomic–AUTH

FoodOmics GR
Prof. N. Thomaidis, Athens Univ
Prof. M. Mikros, Athens Univ
Prof. A. Spyros, Crete Univ.
G. Tsangaris, Athens Academy and the FoodOmicsGR_RI Consortium

AUTH Group
Assist. Prof. Helen Gika
C. Virgiliou, PhD
Q. Deda, PhD
A. Pechlivanis, PhD
A. Orfanidis, PhD
C. Zisi, PhD

PhD students
O. Begou
D. Diamantidou
A. Lioupi
M. Marinaki
A. Pesiroidou
A. Krokos

All MSc students
BIOMIC_AUTh

- Interdisciplinary Effort AUTh
- 17 Staff from 8 schools
- >25 young researchers
- Omics with emphasis on metabolomics
- New laboratory facility 250 m²
- http://biomic.web.auth.gr/

LC–MS/MS
GC–MS/MS
And NMR
Databases
Spectra libraries
Statistical software
Pathway analysis
>100 publications in metabolomics

FoodOmicsGR–RI

- Open Research Infrastructure
- 2019–2021
- 7 Greek Universities and Institutes
- AUTh, Athens Univ., Crete Univ, Ioannina Univ.,
  Aegean Univ., Agric. Univ. Athens, Intern. Univ., Academy Athens

Scope
Map regional foods: comprehensive characterization of foods
Support nutritional studies, studies on food authenticity and control
of geographical origin.
Provide hard molecular data on nutritional value
>450 man months for young researchers
>70 Staff from >25 Disciplines (plant growth, animal husbandry, milk
Science, pharmacognosy, toxicology, analytical chemistry, food chem.
Food technology, informatics, biochemistry—proteomics, NGS,
Elemental metabolomics
The elements...

What is metabonomics?
“the quantitative measurement of the dynamic multiparametric response of living systems to pathophysiological stimuli or genetic modification”

J. Nicholson et al. 1999

Metabonomics is the “systematic study of the unique chemical fingerprints that specific cellular processes leave behind” - specifically, the study of their small-molecule metabolite profiles”

B. Daviss, 2005

Metabonomics/Metabolomics

• The determination of the whole metabolic complement of a cell, tissue, organism, sample
• Correlation with physiology, stress, origin, other factors
• It is a holistic approach
• Why?
• “The whole is greater than the sum of its parts”
Targeted Metabolomics

Key challenges
- Ion suppression issues
- No blanks available
- No agreed strategy for validation
- Metabolites are there! Unlike in pesticide analysis: Urine >140 determined, cell culture >100 metabolites ...
- How to quantify? Surrogate matrix, standard addition method, external calibration?

Sample Preparation Strategies for the Effective Quantitation of Hydrophilic Metabolites in Serum by Multi-Targeted HILIC-MS/MS.
Tsakelidou et al Metabolites, 2017
Signal spans to several orders of magnitude

More issues

• Labelled internal standards not always available
• If available, $$$
• Not easy to handle
• Not easy to organise the calibrants
• Wide range of linear ranges depending on the matrix and condition

• Whatever the cost: **worth the effort**
Workflow

Collect samples

LC-MS analysis:
Short: 10-12 min/ sample
Full scan 100-900 amu

Multivariate statistics
PCA, PLS-DA, O-PLS

Investigate ions that contribute to the group separation

Export as peak list files:
SAMPLE - Mass/Rt

Identify/validate biomarkers

NMR, GC-MS
Orbi TRAP-MS

Multivariate statistics
PCA, PLS-DA, O-PLS

Pipeline for untargeted LC–MS based Metabolomics

Biological question

Study protocol

Sample collection– storage– treatment

LC–MS Analysis

Data preprocessing
Deconvolution, normalisation

Multivariate statistics

Data interpretation
### Obvious sources of error

- Changes in chromatographic properties (retention, peak shape, resolution, selectivity).
- Changes in mass accuracy
- Changes in sensitivity (e.g. source contamination)
  
  **Danger** if instrument variation is misinterpreted as being due to biological process

- Extraction, inappropriate. Extraction edits the sample and thus the metabolic profile

---

### Control Methods

- Internal standards?
- Test Mixtures (external standards ca. 40 metabolites of different classes).
- Quality controls (QCs)?
- Replicates?

---

### Intervention

- How long can a batch be? 24h–48h–72h?
- When to stop and clean the source?
- When to change pre or analytical column?
- How can I put data together? Extract all samples together?
Quality controls/internal standards

- For LC-MS in metabon/lomics internal standards (IS) are not as useful as in conventional quantitative methods. Suggestion: 5 or more Deuterated IS representative of important classes.

- Test Mixture. Data (XIC) used to control the stability of the LC and MS (rt and intensity). This is used to chop bad data sets not to prove that a dataset is useful. Data NOT to be used in PCA.

- QCs – use of a standard/representative BIOLOGICAL sample:
  - an aliquot of each test sample to make a "master mix" or pooled sample. For very large sample sets QCs can be made by aliquoting a number of test samples only (not all).
  - QC should be used for its own batch only for PCA and XIC. Peak Tables from QCs stored and used for comparison.
  - QC samples run before, during and after analysis to assess system suitability.

- Replicate injections. All? Selected samples? For large datasets 1 inj–2inj– or 3inj? Suggestion: First screen one injection of all samples and repeats (1x) of up to 20%. Repetitions of the whole batch in second stage to verify the finding (including the data processing).

QC approach

Data scrutiny: Conventional XIC and MVA.
Gika, JPR 2007, 2012, Guy 2008, Kamleh 2012...
Sample Analysis
UPLC-MS

Data scrutiny: Conventional XIC and MVA.
Gika, JPR 2007, 2012, Guy 2008, Kamleh 2012...

Sample Analysis
UPLC-MS

- Know your method
  - Use methods from literature
  - Run authentic standards

- Unbiased tandem MS acquisition
  - Run Data Dependent Acquisition (DDA)
  - MS2
  (Use pooled/QC samples – pooled groups/extremes)

> Run Pos and Neg modes with the same gradient

- Run a test mix
  - Check RT
  - Accurate mass
  - Frequent adducts

- Run sample preparation blanks
Retention time

Peak height

PCA scores plot

QC cluster
Note: PCA plot can be misleading
In case where test samples have small differences principal components represent small proportion of the variability in the data and QCs appear more scattered.
Plot profile of ions

QC Peak Table: Check ions with high CV% values

<table>
<thead>
<tr>
<th></th>
<th>QC6</th>
<th>QC7</th>
<th>QC8</th>
<th>...</th>
<th>QCx</th>
<th>CV%</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1_t1</td>
<td>1045</td>
<td>1280</td>
<td>1147</td>
<td></td>
<td>1321</td>
<td>8</td>
</tr>
<tr>
<td>m2_t2</td>
<td>876</td>
<td>757</td>
<td>120</td>
<td></td>
<td>850</td>
<td>35</td>
</tr>
<tr>
<td>mx_tx</td>
<td>2378</td>
<td>2395</td>
<td>2547</td>
<td></td>
<td>2367</td>
<td>6</td>
</tr>
</tbody>
</table>

QC Peak Table: Check ions with high CV% values

<table>
<thead>
<tr>
<th>m/z</th>
<th>0</th>
<th>100</th>
<th>200</th>
<th>300</th>
<th>400</th>
<th>500</th>
<th>600</th>
<th>700</th>
<th>800</th>
<th>900</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (min)</td>
<td>0.5</td>
<td>2.5</td>
<td>4.5</td>
<td>6.5</td>
<td>8.5</td>
<td>10.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Metabonomic Study Diseased vs Control population**

Raw data, TIC across all samples
- QC samples

**PCA**

Check for run order effects
Metabonomic Study Diseased vs Control population

QC Run order

Outlier? Problem in the system?

System stabilization?

Raw data, Log (intensity), 4 randomly selected peaks

Peak 113:040:150

Peak 249:360:999

Peak 402:240:5:99

Peak 197:120:4:50
Other proposed filters

- 80% rule
- 50% rule
- System suitability testing
  (i) m/z error of 5 ppm compared to theoretical mass, (ii) retention time error of <2% compared to the defined retention time, (iii) peak area equal to a predefined acceptable peak area ±10% and (iv) symmetrical peak shape with no evidence of peak splitting
- System suitability blank and process blank samples

Not all researchers agree
Day to day variation

Over 20h run (60 human urine samples) repeated for 5 days.

How reproducible the data are?

Obvious day-to day effect appears vertically in t2 although only 8% of total variance in data

11 shows sample to sample variation horizontally
Day to day variation

PCA “control model”

- both day-to-day and within day variation

Day-to-day variation

- Start up variation each day
- 4th and 5th batch higher reproducibility
QC data offers ways to correct batch effect

Pre-processing of data may have a profound effect on the results

- LC-MS raw data, software parameters
  - Unfolding and alignment issues
  - Mass matching and time windowing

- Normalisation (to total signal, to constant peak to baseline noise etc): danger, may reverse the direction of trends!
  
  *Eg. If major peak absent then other peaks increase in apparent importance*

- Scaling

- Improve reproducibility in the analytical method

When we cannot repeat experiment (ethical issues, small sample volume cost) we may ‘rescue’ the data by

- OPLS-DA to “See through” unwanted variation
- spectral filters

BUT overall it can be GIGO
MQACC
https://epi.grants.cancer.gov/Consortia/mQACC/

• The consortium’s mission is to engage the metabolomics community to communicate and promote the development, dissemination and harmonization of best QA/QC practices in untargeted metabolomics through the following objectives:
  • To identify, catalog, harmonize and disseminate QA/QC best practices for untargeted metabolomics.
  • To establish mechanisms to enable the metabolomics community to adopt QA/QC best practices.
  • To promote and support systematic training in QA/QC best practices for the metabolomics community.
  • To encourage the prioritization and development of reference materials applicable to metabolomics research.

MQACC
Reference and Test Material Working Group
Dissemination of Current QA/QC Practices Working Group

• Think Tank on Quality Assurance and Quality Control for Untargeted Metabolomics Studies (October 2017)
• Workshops: Metabolomics 2018 and 2019 Workshop: QA and QC in Untargeted Metabolomics
• Think Tank on Quality Assurance & Quality Control for Untargeted Metabolomics Studies
• Quality Control in Untargeted Metabolomics
• Quality Assurance Panel
• Conferences
• Webex meetings
MQACC
https://epi.grants.cancer.gov/Consortia/mQACC/

Synopsis

- Metabolomics is coming of age for the regulators (best practice guidelines, discussions, few papers)
- LC–MS is going to be a major part of the solution (and the problem!)
- We need to be very cautious with data (instrument variation dependant)
- Always validate data before statistical analysis (Easy to produce statistical artifacts based on bad quality data!)
Reading


Quality control for plant metabolomics: reporting MSI-compliant studies.


Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample, Godzien et al Metabolomics 2015


Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies, D. Broadhurst et al 2018