Interoperability and multi-source data integration in metabolomics for the identification common Metabomic Syndrome phenotypes

***Elfried Salanon*[1]**

***Direction: Blandine Comte*[1]*, Julien Boccard*[2]*,Estelle Pujos-Guillot*[1]**

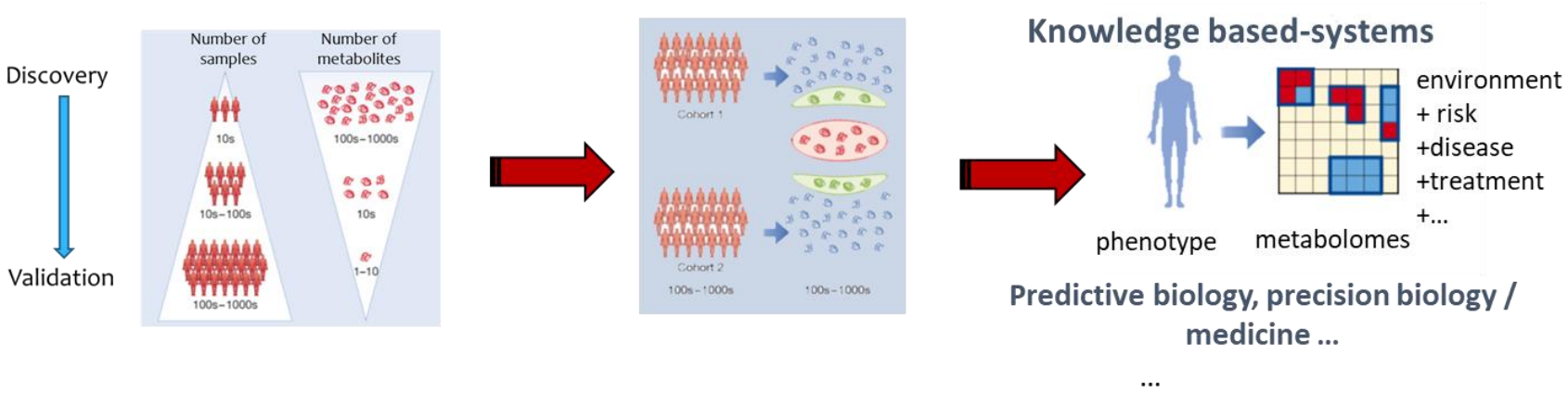
[1]University of Clermont Auvergne, INRAE, UNH, Metabolism Exploration Platform, MetaboHUB Clermont, Clermont-Ferrand, France

[2]School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland

Publications: 5 ***(****1 Published, 1 Under review, 3 in prep)*
Communications: 5 *(2* ***oral com****; 3* ***Posters****)*
Invention reports: 2
Startup in incubation: 1
ANR submitted: 1 (Ongoing)

DIGIT-BIO INRAE metaprograme

SCIENTIFIC CONTEXT



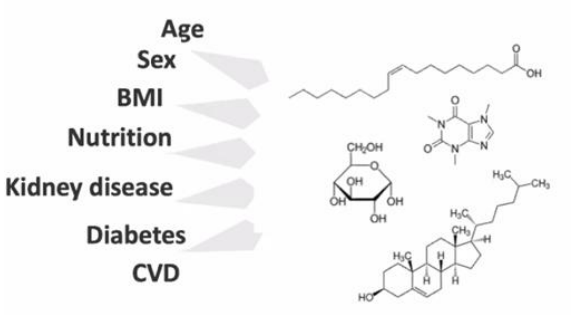
BETTER PREDICT PHENOTYPES

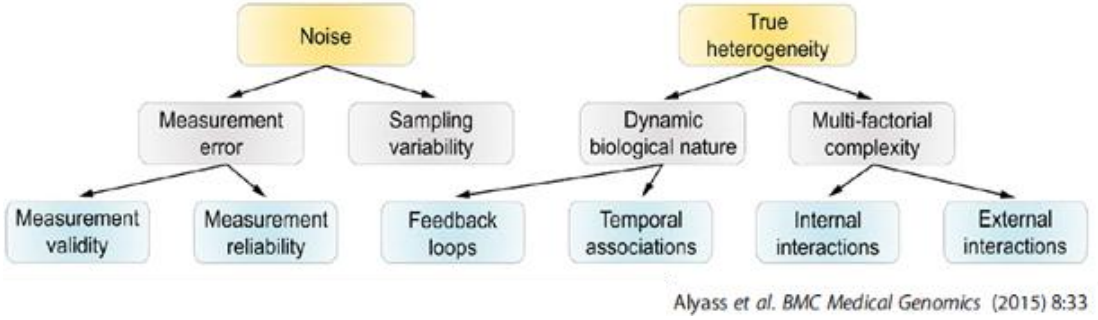
Metabolomics:

- Study of all the small molecules present in a biological matrix. *(Fiehn et al, 2000).*
- Powerful phenotyping tool (*Hajjar et al., 2023*)

METABOLOMICS

Metabolomics Data complexity

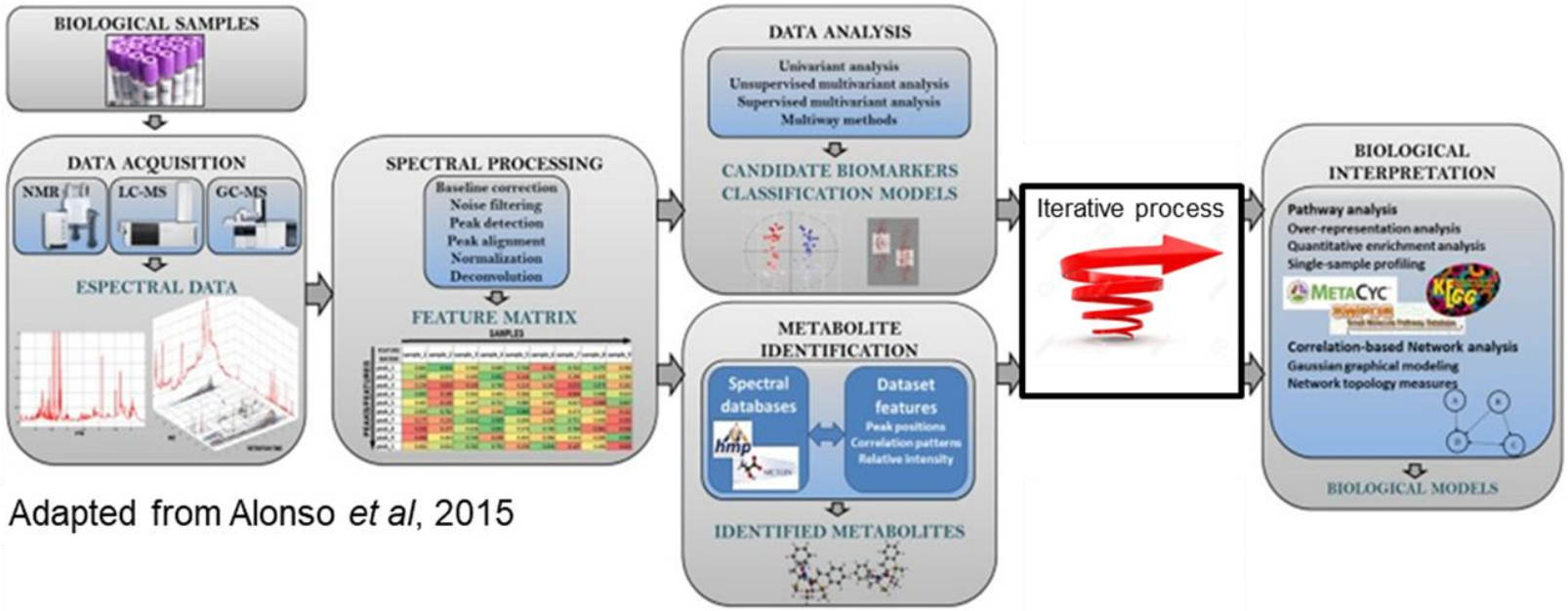




Lack of metabolomics data interoperability, preventing intercomparisons across studies and limiting their impact in precision biology *(Hajjar et al., 2023)*

General objective: To investigate interoperability between metabolomics data from individuals analysed in different health status

Obj 1: Reproducibility of metabolomic data → Obj 2: Cross-study integration → Obj 3: Application : identifying common and early phenotypes of metabolic syndrome



Adapted from Alonso *et al*, 2015

INRAE UNIVERSITÉ DE GENÈVE

COLLÈGE DES ÉCOLES DOCTORALES UNIVERSITÉ Clermont Auvergne

Visualization of the bivariate dispersion structure for the robust assessment of the repeatability and reproducibility of analytical measurements.

***Elfried Salanon*[1]**

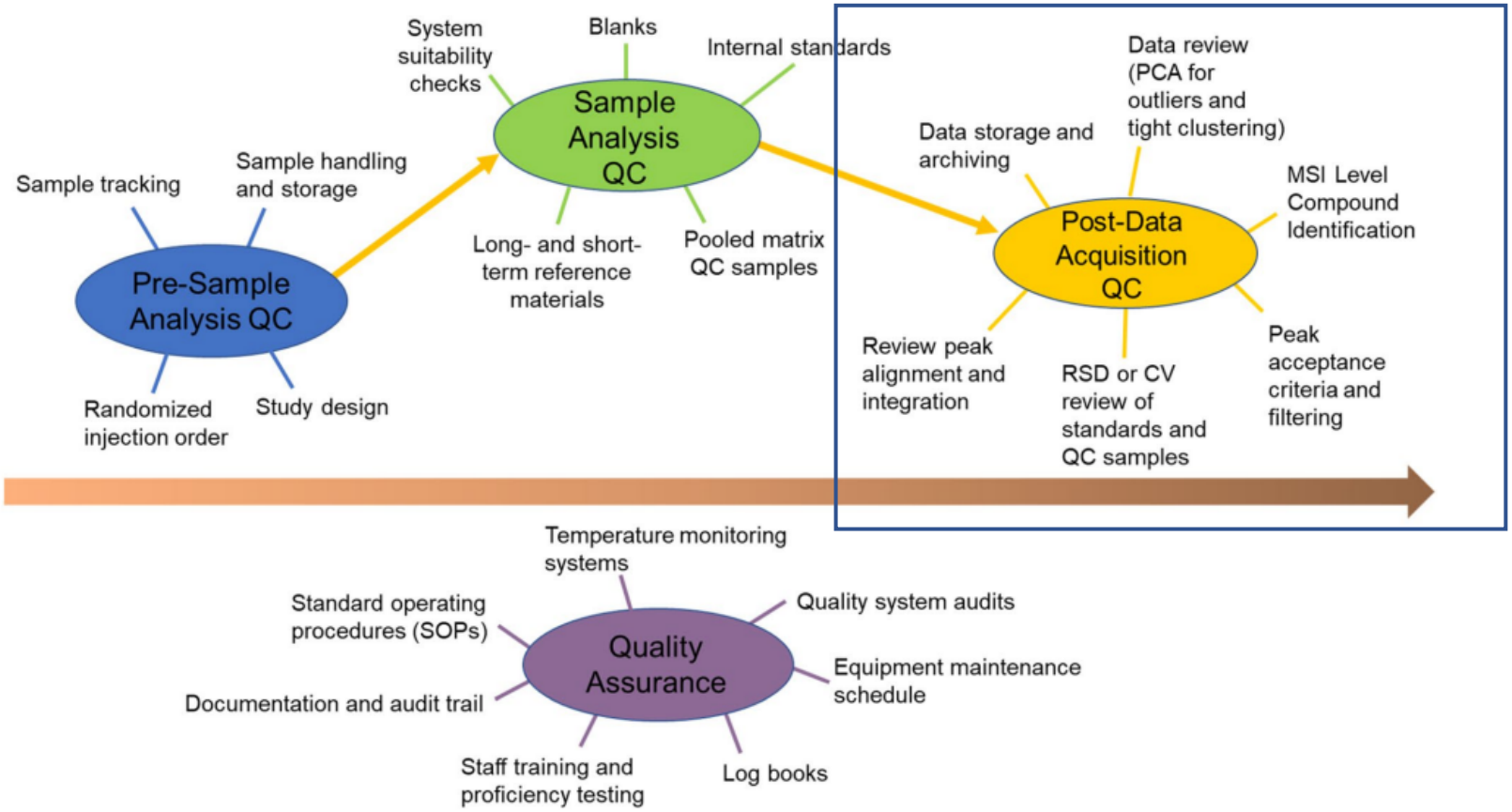
***Direction: Blandine Comte*[1]*, Julien Boccard*[2]*,Estelle Pujos-Guillot*[1]**

[1]University of Clermont Auvergne, INRAE, UNH, Metabolism Exploration Platform, MetaboHUB Clermont, Clermont-Ferrand, France

[2]School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland

DIGIT-BIO INRAE metaprograme

ASSESSMENT OF MEASUREMENTS ERRORS



Challenges:

- Analytical variability comes from multiple sources.
- Classical indicators poorly describe these issues.

(Anne Evans et al, 2020)

See you around the Poster

Visualization of the bivariate dispersion structure for the robust assessment of the repeatability and reproducibility of analytical measurements.

Elfried Salanon[1], Blandine Comte[1], Delphine Centeno[1], Stéphanie Durand[1], Estelle Pujos-Guillot[1], Julien Boccard[2]

[1]Université Clermont Auvergne, INRAE, UMH, Plateforme d'Exploration du Métabolisme, MetaboHUB Clermont, Clermont-Ferrand, France

[2]School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland

Introduction

Challenges:

- Analytical variability comes from multiple sources.
- Classical indicators poorly describe these issues.

Assessment of measurements errors:

- Repeated measurements of QC samples.
- Calculation of quality indicators for each detected variable.
- Use of dispersion indicators

Objective: Develop a visualization method to better capture and understand the dispersion structure.

Methods

The bivariate dispersion using convex hull

- The convex hull of a set of points in plane is the shape taken by a rubber stretched around the nails pounded into the plane at each point
- An Euclidean space is defined by a plane made of the "measurement order" (X-axis) and the "Intensity" (Y-axis) of the analyte signal.
- The proposed visualization methods is made of convex hulls based on the intensities of repeated QC measurements by batch.

Quantitative Indicators

Intrabatch dispersion

$$IntraD = \text{median}\left(\frac{A_{i+1} - A_i}{x}\right)$$

with $i = 1, \ldots, n$ and $n+1$

Interbatch dispersion

$$InterD = \frac{A_{\text{relative order}} * A_{\text{intensity}}}{2}$$

Dispersion index

$$D - Index = \frac{IntraD}{1 + InterD}$$

Metabolomics case study

Metabolomics data obtained from the analysis of a human plasma reference material (SRM1950) made available by the National Institute of Standards and Technology (NIST) as Standard Reference Materials®, analyzed over one year, using a mass spectrometry-based untargeted approach. Deproteinized plasma extracts were injected in triplicates at the beginning of each analytical sequence of various metabolomics studies. Raw data were processed with the XCMS R-package using a Galaxy web-based platform.

Software

The test were conducted using R based code

Conclusion:

- The visualization method effectively evaluates variability structure.
- It allows capturing within and between batch/group variations.
- It allows revealing potential issues in measurement sequence.
- It allows assisting in diagnosing and selecting correction methods.

References

Elfried Salanon, Blandine Comte, Delphine Centeno, Stéphanie Durand, Estelle Pujos-Guillot, et al.. An alternative for the robust assessment of the repeatability and reproducibility of analytical measurements using bivariate dispersion. Chemometrics and Intelligent Laboratory Systems, 2024, 250, pp.105148. ff10.1016/j.chemolab.2024.105148

All metabolomics analyses were funded and performed within the MetaboHUB French infrastructure (ANR-11-INBS-0010). E. Salanon is recipient of a doctoral fellowship from the INRAE DIGIT BIO metaprogramme and Geneva University.

Corresponding author:
Elfried SALANON
elfried.salanon@inrae.fr

Results

Visualisation of analytes with same CV but different structure

L-Tryptophan | Caffeine

Visualisation of analytes of different qualities

L-Tryptophan (CV=0.14) | Sphingosine-1-phosphate (0.32)

The visualisation of real experimental data involving non-constant bias revealed different types of structures with various shapes.

Finding common behavior between analytes

This visualisation tool can also be used to automatically highlight common behaviors between analytes.

Batch effect assessment in metabolomics

Spectra | Before batch effect correction (CV=0.17) | After batch effect correction (CV= 0.09)

INRAE UNIVERSITÉ DE GENÈVE

COLLÈGE DES ÉCOLES DOCTORALES UNIVERSITÉ Clermont Auvergne

DIGIT-BIO