Confidence in metabolomic data obtained by innovative GC × GC-(HR)TOFMS

N. Di Giovanni[1], J.-F. Focant[1]

[1]CART, Organic and Biological Analytical Chemistry, Department of Chemistry, University of Liège, Belgium

Université de Liège

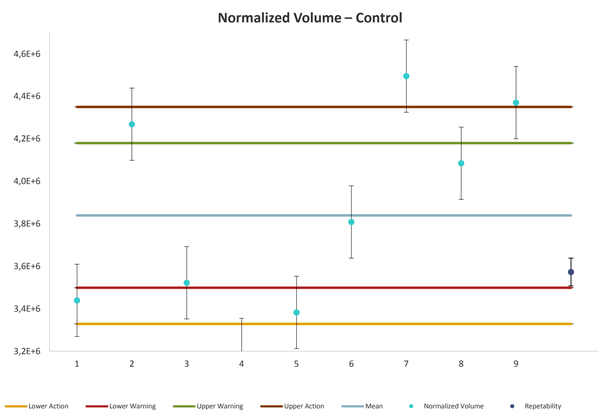
Introduction

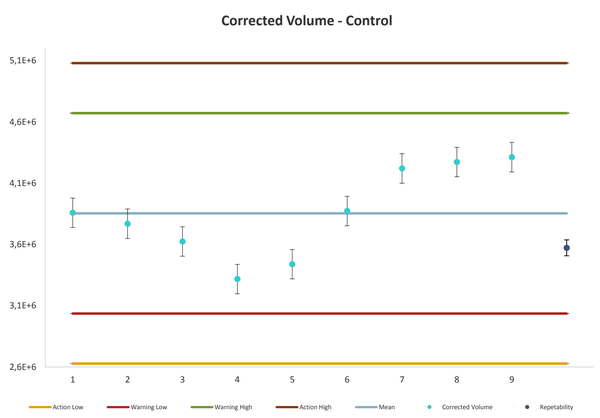
- GC- (TOF)MS powerful for (semi)-volatile metabolites separation and identification ;
- GC × GC-(HR)TOFMS proven to be of added value for untargeted analysis but at the cost of more analytical efforts ;
- Need for further implementation of robust criteria and harmonization of procedures ;

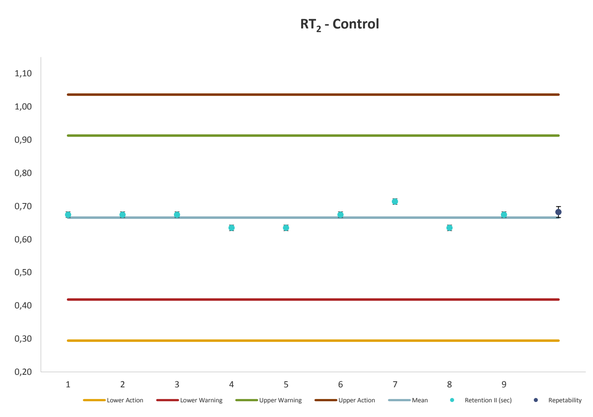
➔ **CONFIDENCE in results is of the utmost importance and still the most difficult thing to achieve.**

Analytical Method – untargeted analysis of human serum

- **Efficiency** through litterature starting point followed by extensive development (Design of Experiment) of :
  - GC × GC (columns phase, length, diameter, flow, temperature ramp, modulation period and hot jet, inlet design and temperature) ;
  - MS (TOF, ion source temperature, voltages).
- **Validation** for accuracy, precision, recovery, sensitivity (LOD, LOQ) and linearity on :
  - NIST certified material (SRM 1950 – metabolites in plasma) ;
  - In-house QC samples representative of the matrix of interest.





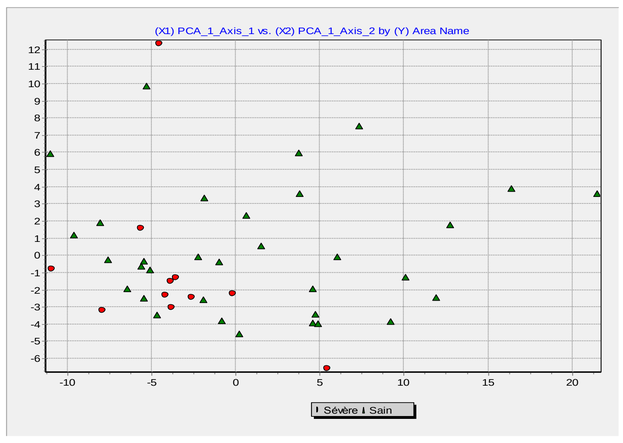


Stability over time and analytical events (maintenance) monitored by visual inspection of chromatographic resolution and by control charts for retention times and peak volume of selected representative compounds.

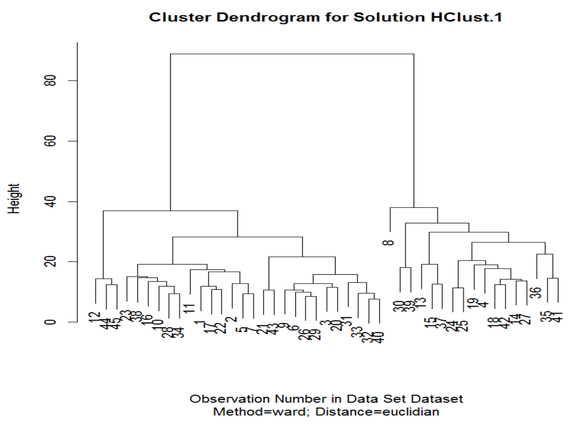
Confirmatory study – inflammatory cohorts – biomarker research

- **Analysis** of samples spiked with internal standards (IS) throughout sample preparation along with daily QC samples analysis to monitor the system and its maintenance needs ➔ **94** injections over 4 weeks including **70** study samples (**8** reinjections) and **16** QC samples.
- **Data preprocessing** : centroidization and alignment of the peaks followed by template creation based on cumulative image of all chromatograms (GC Image software) → definition of **524** significant areas of good chromatographic and mass spectral resolution. Normalisation on IS, correction for operational and instrumental variations – LOESS method on QC samples – and final selection of all areas of RSD < 30% → **178** and **192** areas selected for amino acid and organic acid normalisation, respectively.
- **Data processing (statistics) :** Confidence through multiplication of techniques and control/optimization procedures :

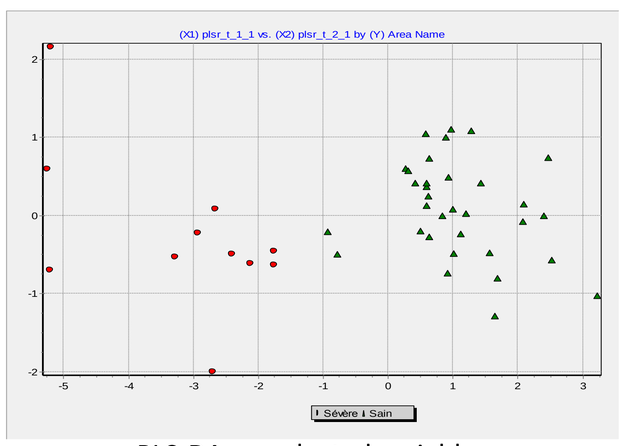
Univariate / Multivariate including supervised learning – Non parametric – Robust statistics



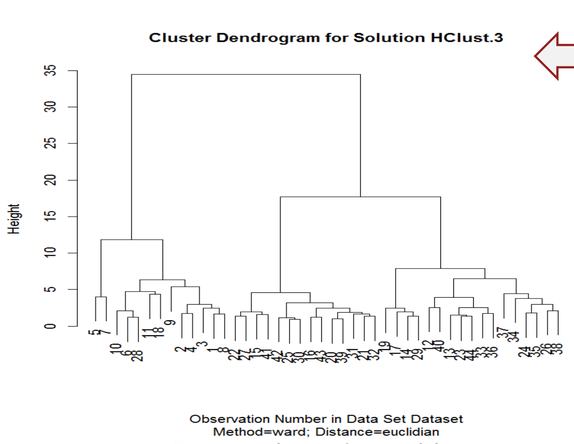
PCA on initial dataset



HCA on initial dataset



PLS-DA on selected variables



HCA on selected variables

One-way ANOVA	CVA / CDA
PCA / AFP	Decision Trees
Clustering (HCA, K-Means)	ROC – AUC
PLS-DA	SVM

Dimensionality reduction	Confidence intervals (> p-value)
Colinearity consideration	Variable selection
Confounding Factors (Bias)	Re-sampling / test validation
Bonferroni correction	Graphical visualization

Conclusion

In order to obtain high confidence in the results, untargeted metabolomic analysis need to be optimized and controlled at every step, from sampling to analytical procedures, data treatment and interpretation.

Organic and Biological Analytical Chemistry Group

OBIACHEM
Organic and Biological Analytical Chemistry Group