STATISTICAL APPROACHES TO HIGHLIGHT ROBUST BIOMARKERS FROM GC×GC-(HR)TOFMS DATA

N. DI GIOVANNI, J-F FOCANT

Quartier AGORA, Allée de la chimie, 11
B-4000 Liège (Sart-Tilman)
Belgique

Despite all analytical efforts seen in recent years in clinical applications of separation science – design of experiment, sample preparation and instrument optimization- many issues remain, especially in GC×GC, a powerful but still not perfectly implemented technique.

Among them, data preprocessing – alignment, normalization and correction (QA/QC) - and statistical treatment. In this biomarker research study, we tried to consider, understand and use the many ways to handle data in order, in the end, to compare them and find the best way to obtain high confidence results.

Univariate, multivariate and non parametric techniques were employed (F-ratio, PCA, CDA, Decision Trees, Clustering (HCA, K-Means), SVM, ROC/AUC curves, PLS-DA), along with multiple testing correction, confidence intervals, variable selection, re-sampling and test validation. Confounding factors (bias) and variable colinearity were considered as well.

A validation was made on inflammatory cohorts, where class-to-class quality statistical separation was achieved and significant biomarkers found, in order to highlight the interest of the method.