

Assessing reproducibility of proteomic data by the multivariate coefficient of variation

Adelin Albert

University of Liège, Belgium

The coefficient of variation $CV = \sigma/\mu$ is commonly used to measure the reproducibility of analytical techniques or equipment: the lower the CV, the better the analytical precision. To assess the reproducibility of methods used in proteomics or genomics, which yield a huge amount of correlated data, we propose to extend the univariate CV to the multivariate setting by the expression $CV_m = [\mu^T \Sigma \mu / (\mu^T \mu)^2]^{1/2}$. The CV_m was applied to four different sample prefractionation methods to select the best starting point for further differential proteomic experiments by mass spectrometry (MS): highly abundant proteins precipitation, restricted access materials (RAM) combined with IMAC chromatography, and peptide ligand affinity beads at two different pH levels (4 and 9), respectively. The reproducibility of the four methods (65-70 peaks) was evaluated with regards to mass and peak intensity (5-6 repetitions). CV_m was low for mass in all methods (range: 0.02 – 0.05%) but for intensity reproducibility was significantly lower for method 1 (highly abundant proteins precipitation) than for the other methods ($CV_m = 2.84\%$ vs. 7.34 – 10%). Unlike univariate CVs calculated for each peak separately, CV_m yields a global reproducibility measure accounting for correlations between peak characteristics. (Joint work with Marianne Fillet and Lixin Zhang)

Biography

Adelin Albert has a Ph.D. in statistics and a doctoral degree in biomedical sciences from the University of Liege (Belgium). He is currently professor and head of Medical informatics and Biostatistics at the University of Liege.

aalbert@ulg.ac.be