Multivariate regression techniques and NMR spectroscopy to quantitatively determine falsification of pharmaceutical and food products

Yulia B. Monakhova[1,2], *Bernd W.K. Diehl* [2]

[1] *Institute of Chemistry, Saratov State University, Saratov, Russia*

[2] *Spectral Service AG, Köln, Germany*

There is no doubt that multivariate analysis is required to authenticate complex pharmaceutical samples and food products and, especially, to quantitatively predict the amount of adulteration in blends. In this report we consider two projects aimed at authentication of porcine heparin and sunflower lecithin by NMR spectroscopy. For multivariate modelling several statistical methods such as partial least squares regression (PLS), ridge regression (RR), stepwise regression with variable selection (SR), stepwise principal component regression (SPCR) were utilized for modeling NMR data of in-house prepared blends (n>80 in each case).

In case of porcine heparin blended with bovine species PLS and RR showed the best performance with the limit of detection (LOD) and root mean square error of prediction (RMSEP) below 2% w/w and 1% w/w, respectively. Reproducibility expressed in coefficients of variation was estimated to be below 10%. The developed method was found to be applicable also to heparinoid matrix (not purified heparin).

The models based on phospholipid (^{31}P NMR), fatty acid (^{1}H NMR) and saccharide (^{1}H NMR) distributions were validated to distinguish pure sunflower lecithin from that blended with soy species. PLS based on saccharide composition is able to estimate lecithin falsification regarding its vegetable origin with the sensitivity and RMSEP below 1% w/w and 3.5% w/w, respectively. Prediction error was improved by fusing three available NMR profiles.

The developed approaches allowed to evaluate the composition of lecithin and heparin blends on the basis of multiple chemical components in parallel with official NMR tests.