The evaluation of standard and advanced preprocessing methods for the univariate analysis of blood serum $^1$H-NMR spectra

Tim De Meyer, Davy Sinnaeve, Bjorn Van Gasse, Ernst-R. Rietzschel, Marc L. De Buyzere, Michel R. Langlois, Sofie Bekaert, José C. Martins, Wim Van Criekinge
**Fig. S1: Introduction of artificial correlations by normalization** for both AI-binned (Electronic Supplementary Material Fig. S1a) and equidistant binned (Electronic Supplementary Fig. S1b) spectra. Frequency polygon of R²-values for correlations between triglyceride concentration and each variable in the different preprocessed datasets: integral normalization, maximum (-) and summed (--) intensity variable definition; PQ-normalization (post binning), maximum (····) and summed (··-) intensity variable definition. The frequency polygons demonstrate that integral normalization results in the introduction of artificial dependencies between variables. This also appears to be the case for PQ-normalization, but to a clearly lower extent.

Fig. S1a: