

Supporting Information

1. Outlier Detection

Hotelling's $T^2 \times Q$ residuals are statistical techniques used in multivariate analysis to detect outliers and assess the quality of a statistical model. Hotelling's T^2 statistic is a multivariate generalization of Student's t-test. It measures how far the sample mean vector of a multivariate dataset is from the population mean vector, accounting for the correlation structure between variables. Q residuals are used to identify univariate outliers in a multivariate dataset. It measures the Mahalanobis distance of individual observations from the centroid (mean vector) of the dataset, considering the correlation between variables.

1.1 Hotelling $T^2 \times Q$ plot

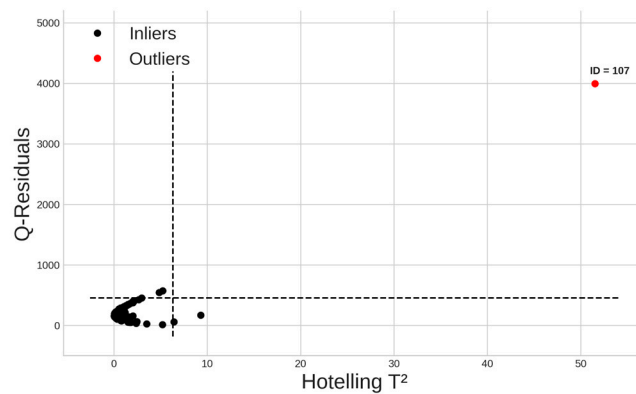


Figure S1. Hotelling $T^2 \times Q$ plot evidencing one outlier (patient id = 107).

1.2 Outlier spectra

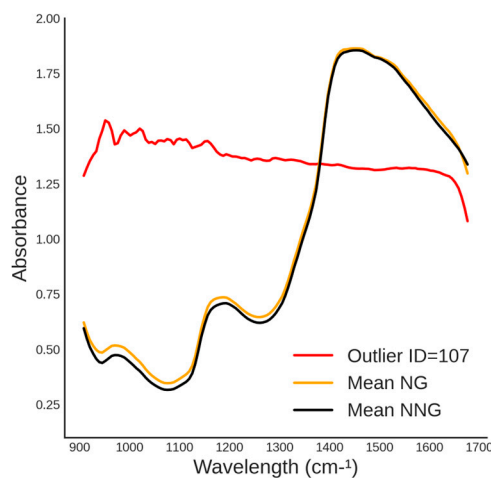


Figure S2. Outlier spectra versus average spectra per group (NG \times NNG).