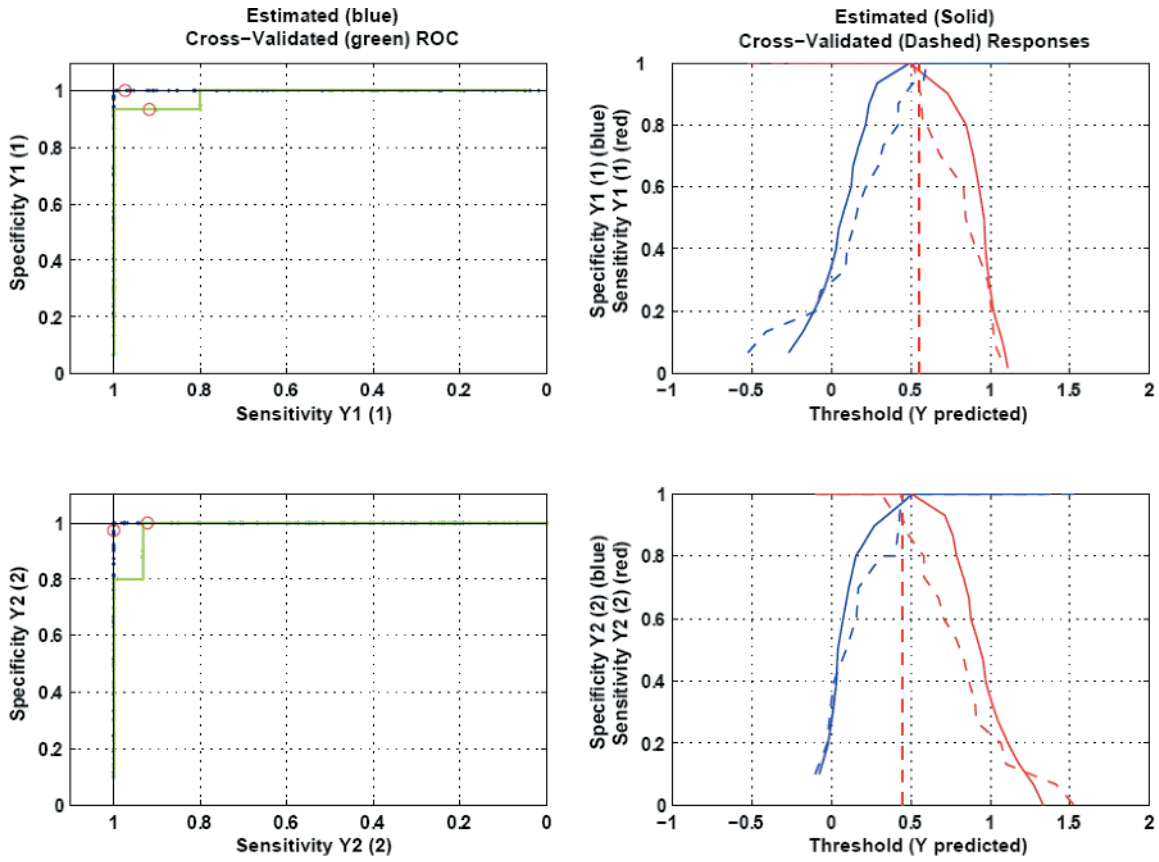

Figure W1. Receiver operating characteristic curves of sensitivity *versus* specificity (the plots on the left) and threshold *versus* specificity/sensitivity (the plots on the right) calculated for cross-validated PLS-DA applied to NMR spectra of serum samples from healthy controls (class 1) and patients affected by oral cancer (class 2). A high degree of specificity and sensitivity is observed.

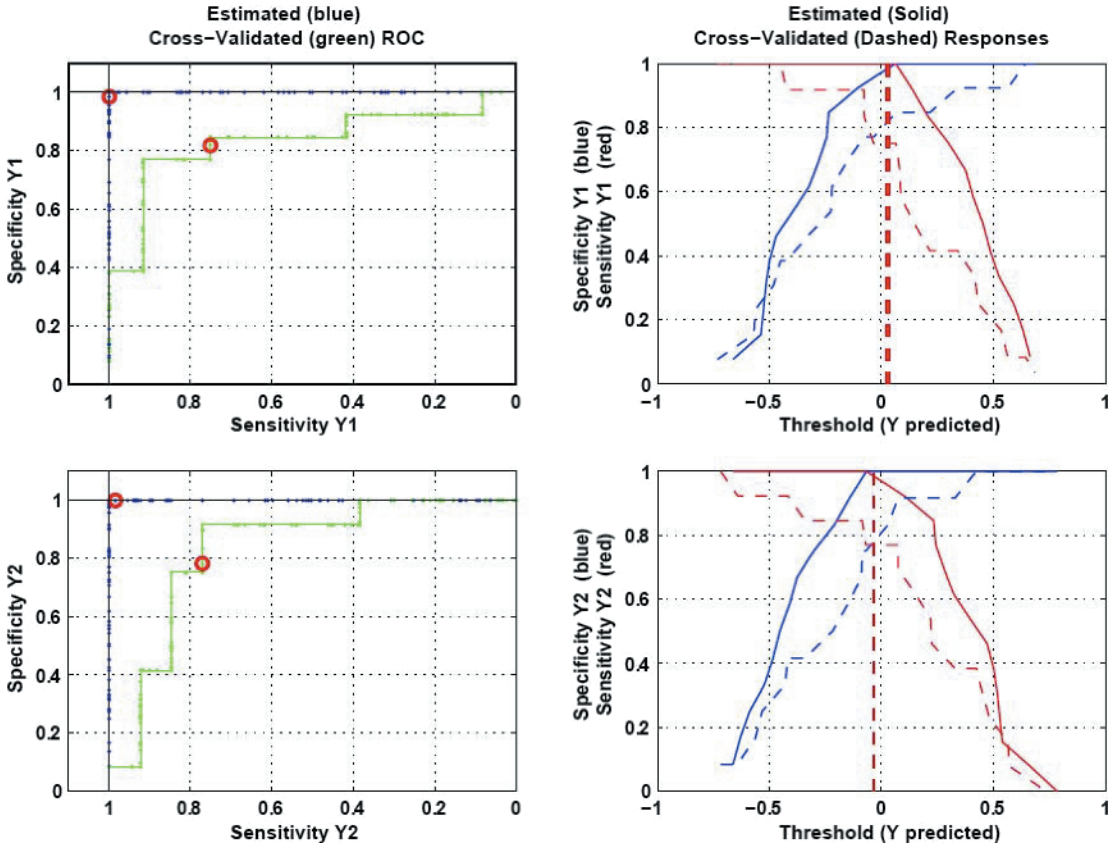

Figure W2. Receiver operating characteristic curves of sensitivity *versus* specificity (the plots on the left) and threshold *versus* specificity/sensitivity (the plots on the right) calculated for cross-validated PLS-DA applied to NMR spectra from healthy controls below (class 1) and above (class 2) the mean age. Compared with Figure W1, a lower degree of specificity and sensitivity is observed.

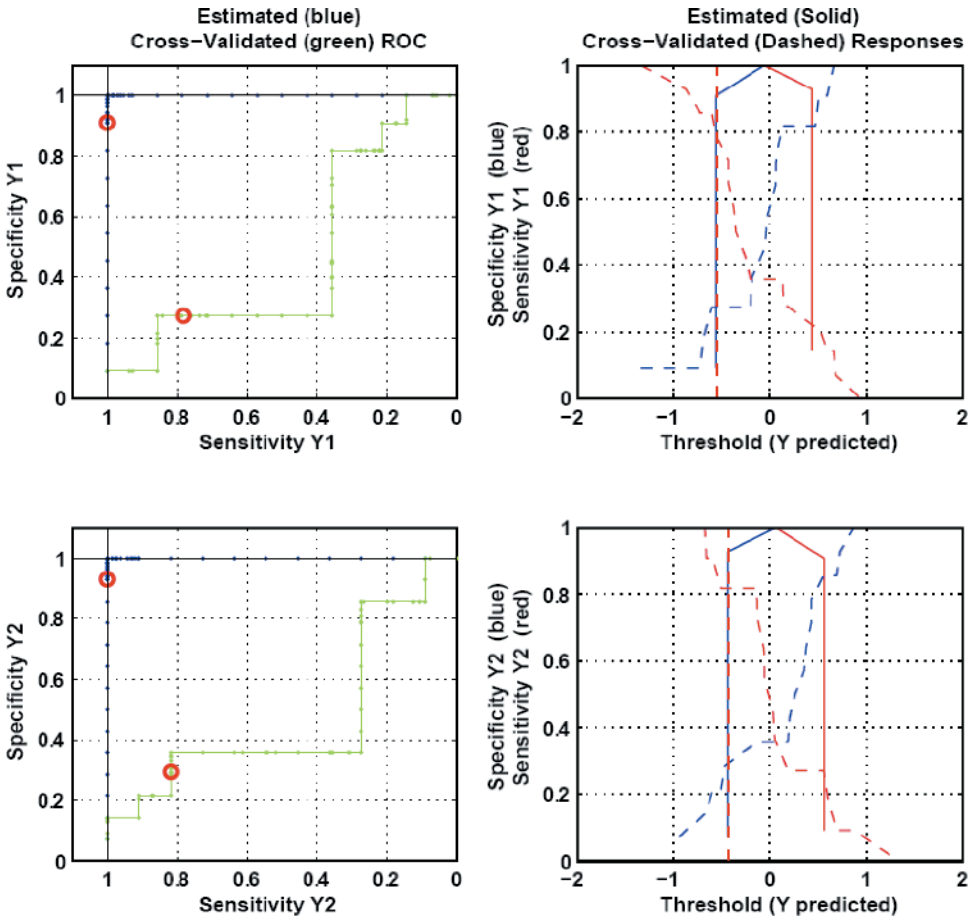

Figure W3. Receiver operating characteristic curves of sensitivity *versus* specificity (the plots on the left) and threshold *versus* specificity/sensitivity (the plots on the right) calculated for PLS-DA applied to NMR spectra with groups for sex (male: class 1 and female: class 2). Compared with Figure W1, a lower degree of specificity and sensitivity is observed.

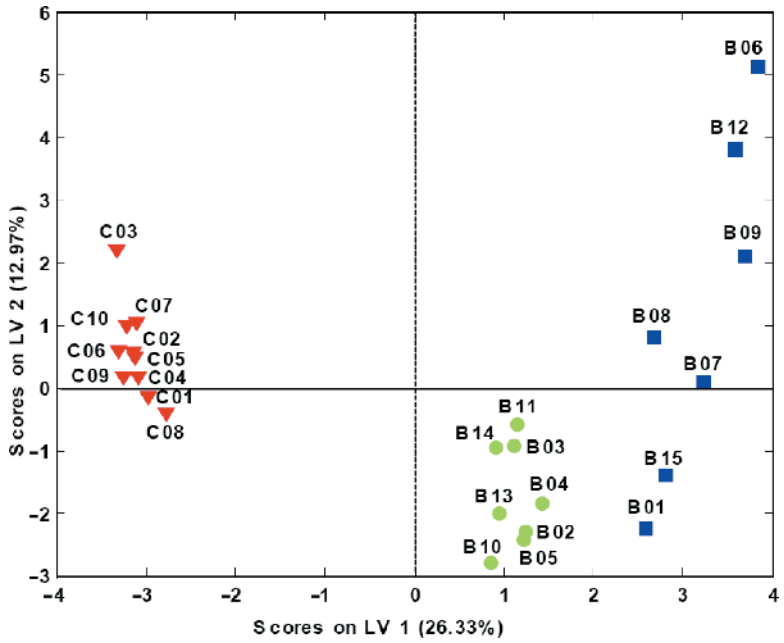

Figure W4. PLS-DA scores plot of ^{1}H NMR spectra of human blood sera using three different classes for patients with ● early and ■ late stage diseases and healthy controls (▼). The application of this multivariate chemometric model shows a clear separation between the samples from healthy controls and the patients with disease. Very narrow groupings are observed for the control samples and the patients with early stage disease. On the contrary, the sera of patients with late stage disease are more widely spread in the scores plot, highlighting the higher inhomogeneity of their metabolic profile.