

Analyzing Spectroscopic Datasets

Brian T. Luke (lukeb@ncifcrf.gov)

This section presents the results of multiple analyses of different datasets.

[Published Prostate Dataset](#)

Adam and coworkers [[Ada-02](#)] produced mass spectra for individuals in four different groups; prostate cancer (T₁/T₂), prostate cancer (T₃/T₄), benign prostate hyperplasia (BPH), and individuals with healthy prostates. Several classification studies using subsets of these four datasets are presented.

[Datasets Containing No Biological Information](#)

To examine the classification accuracy of various methods, and therefore the [significance of each method](#), on datasets with no biological information, 30 artificial datasets with random peak intensities have been created. All datasets contain 300 “peaks” and have the same number of Cases and Controls (30, 42, 60, 90, 150, and 300 of each category). For each number of Cases and Controls, five independent datasets are available for [download](#) and analysis.

(Last updated 4/29/07)