**Supporting Information**

**Expert Algorithm for Substance Identification (EASI) using Mass Spectrometry: Application to the Identification of Cocaine on Different Instruments using Binary Classification Models**

Samantha A. Mehnert,a,b,# J. Tyler Davidson,a,$ Alexandra Adeoye,*a* Brandon D. Lowe,b Emily A. Ruiz,b Jacob R. King,b Glen P. Jacksona,b\*

*aDepartment of Forensic and Investigative Science, West Virginia University, Morgantown, WV 26506, USA.*

*bC. Eugene Bennett Department of Chemistry, West Virginia University, Morgantown, WV 26506, USA.*

*#Current Address: Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA.*

*$Current Address: Department of Forensic Science, Sam Houston State University, Huntsville, TX, 77340, USA.*

\*glen.jackson@mail.wvu.edu;*304-293-9236*

**Table S1.** Summary statistics for 20 ion abundances of different groups of compounds.

**![Table

Description automatically generated]()**

**Table S2**. Summary of spectral measures of similarity and dissimilarity for the consensus approach and EASI approach to predicting ion abundances in different data sets.

|  |
| --- |
|  |

\*Using Stein’s original weightings of *x*=0.6 and *y*=3 and equations 1 and 2.1

**Table S3.** Summary of binary classification figures of merit for various measures of spectral similarity for EASI and the consensus model for cocaine identification in the absence of retention time information relative to the cocaine training set. Only the 20 most abundant peaks in a spectrum are compared in each model, and the threshold is set to allow one false positive so that a likelihood ratio (TPR/FPR) can be calculated. The likelihood ratio is vastly underestimated because of the limited number of known negatives in the database. Mahalanobis distances are calculated for every spectrum relative the cocaine training set from Lab 1.



(1) Stein, S.; Scott, D. R., Optimization and testing of mass spectral library search algorithms for compound identification, *J Am Soc Mass Spectrom* **1994**, *5*, 859-866.