B33   **Validating the Sexual Lubricant Database Using True Known and Unknown Samples for Forensic Analysis**

Santana A.L. Thomas, PhD, National Center for Forensic Science, Orlando, FL 32816-2367; Brooke R. Baumgarten, MS, Winter Park, FL 32792; Candice Bridge, PhD, National Center for Forensic Science, Orlando, FL 32816*

Learning Overview: The goal of this presentation is to outline the validation process of a sexual lubricant database. After attending this presentation, attendees will be informed of an established sexual lubricant database and its potential in the identification and classification of sexual lubricants.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by providing a database that can be incorporated into sexual assault cases. This database contains reference lubricants analyzed on both spectroscopic and spectrometric instruments that are common for analyzing trace evidence.

Efforts to analyze, identify, and classify sexual lubricant samples in the unfortunate cases of sexual assaults have led to the development of a sexual lubricant database established by the National Center for Forensic Science. Currently, the database is a compilation of over 100 reference sexual lubricants of various marketing types. The database provides characterization and instrumental data, as well as classification models for unknown samples. Research has demonstrated that sexual lubricants can be classified into distinctive groups based on chemical composition, some of which are linked to specific marketing types based on the presence of unique compounds (e.g., personal hygiene projects, bottled lubricants or condom lubricants). Therefore, the samples contained in the database were analyzed to provide classification models that can be utilized in classifying unknown sexual assault samples.

All lubricant samples in this study were analyzed in their neat form using Fourier transform infrared spectroscopy and direct analysis in real time-high resolution mass spectrometry in the case when bulk evidence is submitted to the laboratory (e.g,. condom wrappers). The samples were also extracted in methanol or hexane for analysis on gas chromatography/mass spectrometry. The extracts of lubricants were also necessary to simulate trace evidence scenarios, in the event that only a swab of a lubricant residue was submitted to the laboratory. Therefore, these extracts were also analyzed on the previously mentioned instruments to provide a comprehensive analysis of trace residues. The analysis of the data included supervised and unsupervised statistical techniques (e.g., hierarchical cluster analysis, Principal Component Analysis [PCA] and Linear Discriminant Analysis [LDA]). These statistical tools provided class assignments of the samples that were then used to calculate the percent error of the model using the projected classifications obtained in the LDA. Based on cross-validation of the parsed-out test set, all the models that were developed possessed classification error rates less than 5%. Further validation was essential for real-world scenarios.

The validity of the established method underwent two accuracy tests. First, 50 new samples were analyzed to evaluate the classification model using percent error. The theoretical "ground truth" classifications were determined by combining the data from the new 50 with the established classification model using PCA. Once these theoretical classes were determined, the projected classifications were then determined by using the 50 new samples as a test sample against the initial classification model. The theoretical and projected classifications defined the error rate of the classification model for a large set of new, "unknown" samples. Furthermore, a second validation method incorporated the use of real-world samples because reference samples can provide a chemical fingerprint but does not account for factors introduced during coitus. Lubricant residues were collected post-coitus from various points of contact on the sexual organs for analysis and subsequently subjected to classification using the initial classification model as well as the new classification model that incorporated the 50 new samples. This presentation expands on the evaluation of classifying known and unknown samples in the sexual lubricant database to operationalize it in real-world sexual assault cases.

Sexual Assault, Database, Trace Evidence

*Presenting Author