F12 Can We Handle It? Creating a Reference Database to Test the Limits of Current Forensic Software?

Kenneth W. Aschheim, DDS, 44 East 67th Street, New York, NY 10065*

The goal of this presentation is to familiarize the attendee with the methodology utilized in the creation of a reference database for use in dental research on victim identification.

As part of the mass disaster identification team, the forensic odontologist has become a key player in the identification process. However, recent mega disasters such the Asian tsunami of 2005, and the Haitian earthquake of 2010, has shown that the odontologist's role is often limited by the scarcity of dental records and the extremely large size of the victim pool. Although disasters such as Hurricane Katrina and the World Trade Center Attacks have shown that the current modalities work adequately for a victim population in the low thousands, a key unanswered question is what would happen if the number of victims that needed to be identified dentally were 10 or even 100 times the number of victims seen previously? How is this affected by introducing body fragmentation?

In order to test these questions, a reference database needs to be created to objectively evaluate the limits of current forensic odontological software. This presentation will impact the forensic science community by describing the methodology involved in creating a reference database of dental data based on real world data. In addition, by characterizing previous dental disaster data and carefully controlling characteristics of the antemortem to postmortem changes in the reference databases the ability to improve matching algorithms could be explored.

Part of New York City's goal for disaster preparedness includes the ability to handle thousands and possibly tens of thousands of victims. Although current modalities have served well in the past, questions still arise as to what would happen if an even greater magnitude mass fatality incident were to occur. In addition, questions arise outside of the mass disaster realm as to how effective current dental matching algorithms are with large national databases of missing/unidentified individuals.

One of the dictates of evidence-based decisions is the ability to test these questions in a controlled, objective manner. Therefore, a realistic reference database needed to be created that would allow for benchmarking of current forensic odontology systems. Currently, New York City dental forensic reference databases, which are used for testing, contain up to 3000 entries. The goal of this project was to create a database of at least 10 times that size based on real world data and controlled manipulation of key parameters.

For the purposes of this project, a large sample of dental data was compiled from numerous National Health and Nutrition Examination Studies (NHANES) as well as the 1994 and 2000 Tri-Service Comprehensive Oral Health Survey (TSCOHS) military data. The NHANES data that was utilized for this study consists of approximately 33,000 records of adults aged 17 and above, in four databases of the U.S. population, compiled between the years of 1988-2004. The TSCOHS used for this study was compiled as part of a congressional directive to evaluate the dental care system. It consisted of standardized protocols developed by dental epidemiologists to assess the oral health of over 20,000 Army, Navy, and Air Force personnel. The NHANES and TSCOHS data provide over 50,000 records of dental data. The two databases are a useful source of "real world" data; however, since it gives only a single "snapshot" of tooth condition it could only serve as a starting point for database creation. The first step in the conversion process was to analyze the NHANES coding system. Unfortunately, the type of dental data collected and the coding methodology used changed over time. It was then necessary to find the "least common denominator" for the data to ensure that a single database could be created. Methodology was developed to create statistically valid dental data when the NHANES coding system lacked the information. Finally, conversion algorithms were created to translate NHANES coding to a more universally accepted coding system.

In order to create a transition from antemortem to postmortem, different methodologies were explored. A look-up table was created to define all possible "explainable" and "unexplainable" discrepancies based on the software's codes. Software was then created that could alter the data in order to create random control changes to allow for this transition. For example, if the software was set to place "1 explainable discrepancy per record" the selected antemortem codes for "M" would randomly be changed to either "MO," "MOD," "MODFL," "X," etc. on the postmortem side. If the software was set for a "1 unexplainable discrepancy per record" the antemortem code for "M" code would randomly be changed to either a "OD," "V," "FL" etc. on the postmortem side.

The result of this project has been to build software that creates an infinite number of scenarios to mimic mass fatality incidents of various parameters. The creation of a very large reference database that can be manipulated in a controlled manner is a vital first step in the testing of forensic odontology software and our ability to utilize evidence based techniques. In addition, an added benefit of the software was its' ability to

 ** Presenting Author*

quantify the degree of explainable and non-explainable discrepancies that occur in “real world data,” which will also be presented. Finally, having a reference data set is a vital first step in testing the performance of new coding formats and ranking algorithms in order to advance the field.

Multiple Fatality Incidents, Reference Database, Forensic Odontology Software

 ** Presenting Author*