A4 A Random Forest Approach to False Start Analysis

Kirsty Alsop, MSc, Coventry, West Midlands CV47AL, UNITED KINGDOM; Waltraud Baier, PhD, WMG, Coventry CV47AL, UNITED KINGDOM; Danielle Norman, PhD, WMG, Coventry CV47AL, UNITED KINGDOM; Brian Burnett, University Hospitals Coventry and Warwickshire, Coventry CV22DX, UNITED KINGDOM; Mark A. Williams, PhD, WMG, Coventry CV47AL, UNITED KINGDOM*

Learning Overview: The goal of this presentation is to inform attendees of the accurate prediction capability of regression random forest models for false start data.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by presenting a novel method for saw mark analysis that is complementary to current methods.

Introduction: In an attempt to increase objectivity in forensic disciplines, statistical and quantitative analyses are used to complement traditional techniques. Micro-Computed Tomography (micro-CT) of false starts allows these devices to be applied to the forensic toolmark field, which commonly still utilizes human evaluation techniques. Micro-CT is an ideal method for quantitative analysis as measurements may be taken from the digital model without altering the sample during the evaluation. Furthermore, as micro-CT is non-destructive, traditional evaluation methods can be used complementarily.

Method: Three hundred forty experimental false starts were created on fleshed human cadaveric long bones to represent the variability of marks seen in forensic casework. A total of 38 saws, of multiple classes spanning hand saws, hacksaws, and power saws, were used. Micro-CT scans of each bone were taken and cross-sections of the false starts collected to generate a 2D profile image. Quantitative data, in the form of seven distinct measurements, were established from the cross-sections. The seven measurements used aimed to describe the false start profile quantitatively, and included three width measurements, convex height, and three profile floor angles. Two regression random forest models were built on the experimental data with the aim of predicting the saw blade thicknesses from the false start measurements. An Unknown Class Model (UCM) was created from one regression random forest built on the pooled experimental measurement dataset. A Known Class Model (KCM) was built on the data separated by class and hence consisted of three regression random forests. A second set of experimental false starts was similarly created and measured, but excluded from the training data, to test the accuracy of the random forest models.

Results: Micro-CT is a suitable tool for the quantitative analysis of false starts on bone. Regression random forest models can be successfully created and are able to predict the blade thickness of the saw used to create a false start with up to 100% accuracy, within two standard deviations of the measured mean, if the saw is of a known class. If the class of the saw is unknown, saw blade thickness prediction is less accurate, with 88.3% of thicknesses being correctly predicted to within three standard deviations of the measured mean.

Conclusion: Random forest models have been produced from experimental false start measurement data to accurately predict saw blade thicknesses quantitatively. Knowing the class of the saw enables a more accurate prediction of saw blade thickness from false start measurements. Application of this technique in combination with traditional analysis techniques may provide further information to toolmark analyses in casework.

Toolmark, False Start, Statistical Analysis

*Presenting Author