The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	Enhancing Foundational Validity of Forensic Findings in Medico-Legal Strangulation Examinations
Author(s):	Kathryn S. Laughon, Ph.D., RN, SANE-A, FAAN; Karen Kafadar, Ph.D.
Document Number:	308456
Date Received:	January 2024
Award Number:	2018-VA-CX-0004

Final Research Summary Report

FEDERAL AWARD ID
2018-VA-CX-0004

PROJECT TITLE
Enhancing Foundational Validity of Forensic Findings in Medico-Legal Strangulation Examinations

RECIPIENT ORGANIZATION
University of Virginia
P.O. Box 800826
Charlottesville, VA 22908

PRINCIPAL INVESTIGATOR
Kathryn S. Laughon, PhD, RN, SANE-A, FAAN
Associate Professor, University of Virginia, School of Nursing
klc6e@virginia.edu

CO-INVESTIGATOR
Karen Kafadar, PhD,
Commonwealth Professor, Department of Statistics
University of Virginia
kkafadar@virginia.edu

PROJECT/BUDGET PERIOD: January 1, 2019 to December 31, 2023

FEDERAL FUNDS AWARDED: $726,344.00

This project was supported by Grant ID 2018-VA-CX-0004, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Summary of the Project

Nonlethal strangulation of women—a felony crime in 45 states—poses a significant threat to public safety, resulting in a loss of safety for victims and an increased risk of homicide. As many as 13 million women in the United States (U.S.) are strangled each year. Nearly 80% of strangulations in the U.S. are committed against women by their male intimate partner (IP). Nonlethal strangulation is a key predictor of IP homicide. Evidence also suggests that strangulation perpetrators are more likely to commit police officer homicide. Victims who survive strangulation suffer long-term negative health outcomes. Strangulation victims tend to experience frequent and severe forms of violence by the strangulation perpetrator, suggesting that early criminal justice intervention may reduce homicide risk. Thus, appropriate identification and prosecution of strangulation cases is an important public safety consideration. Nonlethal strangulation is difficult to prosecute because it leaves little external physical evidence when not assessed using careful forensic examination protocols. Some strangulation injuries are non-visible and easily overlooked, such as loss of memory, involuntary urination, and breathing changes. Other injuries, such as petechiae and bruising throughout the face and neck, can be non-specific and may be attributable to other causes, including underlying disease, medication, and other assault-related injuries. Difficulties documenting forensic evidence in strangulation crimes reduce the likelihood that prosecutors are able to file charges that may lead to the conviction of a strangulation perpetrator.

To date, studies typifying strangulation injuries offer only descriptive information regarding types of injuries associated with nonlethal strangulation. Many such studies lack a comparison group, threatening the internal validity of such findings. This may lead to weak evidentiary hearings in strangulation prosecution cases. Therefore, expert testimony provided by forensic nurses and other medical professionals is limited. The current state of strangulation

science confines expert testimony to merely describing injuries attributed to strangulation 'based on the expert's experience and training,' thereby leaving testimony open to critique. Expert testimony that can quantify the likelihood that observed injuries were attributable solely to strangulation (versus other assault-related injuries or disease processes) may drastically improve conviction rates, ultimately reducing homicide risk and increasing public safety. More rigorous research differentiating strangulation injuries from other injuries with probabilistic metrics is needed to hold offenders accountable while ensuring that incorrect interpretation of injury findings will not lead to wrongful convictions.

Major Study Goals and Objectives

Goal 1. To create a large de-identified database of forensic findings from existing medico-legal strangulation and non-strangulation exams of adult women.

Goal 2. Use probabilistic modeling to identify injury findings or clusters of injury findings accompanying reported strangulation of adult women.

Research Question

Are there documented features, or clusters of features, that are associated with cases where strangulation is reported compared to where strangulation was not reported?

Research Design, Methods, Analytical and Data Analysis Techniques

Design, data source, and sample

This study is a retrospective analysis of data from medical record data from patients seeking forensic examinations for strangulation and/or sexual assault who presented to the emergency department of an academic medical center from January 2018 (when this team began systematically performing strangulation examinations) to June 202. Eligible medical records were entered into a research database. Records are those of forensic patients aged 13 years and older who reported sexual assault (SA) by any perpetrator and/or strangulation (ST) by an intimate partner. The final dataset for this study included the following information on all 170 patients: a) type of forensic case: *ST-present* or *ST-absent,* b) numbers of injuries of various types and at multiple body locations, and c) demographic information.

Injury type and location: As part of their routine examinations, trained forensic nurses documented the type and anatomical location of all injuries. Types of injuries documented include bruises, lacerations, abrasions, petechiae, redness, incision, puncture, tenderness, swelling, bites, and scratches. The 37 specific body locations were consolidated into five non-overlapping anatomical categories, detailed in Table 1. Further, injury types (abrasion, petechiae, etc.) tended to be present on a victim (count > 0) or absent (count = 0), so, in some instances, a dichotomized variable from this injury type was created, e.g., "petechiae present" versus "petechiae absent." Finally, while SA examinations included assessment of genital injuries, ST examinations did not, so those injuries and injury sites were omitted from this analysis.

Table 1. Injury Location Categories

	Number of Locations	*Locations Included*
Category 1: Head & Neck	8	Head, Neck, Left/Right Pinna, Left/Right Behind Ears, Left/Right Ear Canals
Category 2: Mouth	3	Cheek, Tongue, Lips
Category 3: Face	7	Face, Left/Right Eyelids, Left/Right Eyeballs, Nose, Under-chin
Category 4: Upper Body	9	Left/Right Arms, Left/Right Hands, Left/Right Shoulders, Chest/Breast, Back, Abdomen
Category 5: Lower Body	10	Left/Right Thigh, Left/Right Knee, Left/Right Foot, Left/Right Lower Extremity, Buttocks, Pelvis

Figure 1. Injury Types

List of Injury Types Documented Among Cases

Injury Type
Abrasion
Bite
Bruise
Incision
Laceration
Petechiae
Puncture
Redness
Scratch
Swelling
Tenderness

Demographic variables (age in years, race, and ethnicity) were recorded, as was the estimated time between the assault and the forensic examination, based on the patient's report in the medical record.

Measures and data analysis

Descriptive statistics were used to report sample characteristics such as age, race, and ethnicity. Chi-squared statistics for numbers of injuries of different types and locations, by ST outcome, guided the selection of variables to be included in a discrimination model. An alpha of 0.02 (see next section) determined the significance of the association. Guided by these single-feature results, a logistic regression model was fitted to distinguish *ST-present* from *ST-absent* cases, using the victim's age and the variables identified in the chi-squared tests. Lastly, a simpler, more intuitive classifier than logistic regression (with similar accuracy) was created to identify aspects of the forensic examination that distinguished cases where ST was and was not attempted. The logistic regression and the simpler classifier were compared in terms of predictive performance to determine the most parsimonious and interpretable algorithm for classifying cases as `*ST-absent*' versus `*ST-present*.'

Missing data

Missing data were imputed or recoded, depending on the type of variable and percentage of missing values. All statistical analyses and imputations were performed using the R statistical software (version 4.2.2 [2022-10-31]).

Expected Applicability of the Research

To our knowledge, this study has produced the first algorithm in medico-legal research to assist with classifying cases of strangulation, given specific forensic examination characteristics. **We emphasize the critical importance of further testing this algorithm to ensure its accuracy and reliability in a court or clinical setting**. Nonetheless, this approach shows promise. This methodology enables forensic examiners and expert witnesses to use data-based techniques to assess the findings from examinations after strangulation events. This would allow an expert witness to quantify their degree of confidence in describing the injuries as associated with strangulation, thus enhancing the scientific validity of their testimony. Notably, the influential report "Strengthening Forensic Science in the United States: A Path Forward" emphasizes the pivotal nature of including an expert's confidence level in their conclusions when presenting forensic evidence in court.

Participants and Other Collaborating Organizations

Dr. Karen Kafadar conducted the statistical analysis. Dr. Sherry Kausch participated in the data cleaning and initial analysis. Ms. Reanna Panagides provided statistical support and writing assistance and helped prepare the datasets for archiving. Dr. Andrea Cimino participated in the design of the project and early analysis. HonorHealth in Maricopa County, Arizona, provided the initial dataset.

Changes in Approach from the Original Design and Reason for the Change

Our original plan was to analyze cases from a large dataset from Arizona that contained 12,999 cases: 5,784 cases of sexual assault only (SA), 5,469 cases of strangulation but no evidence of sexual assault (ST), and 846 cases where both ST and SA were evident (ST+SA). This large database included the number of injuries of many types, the number of injuries at many body locations, and the stated outcome (SA, ST, SA+ST). SA/ST cases represented only 7 percent of the 12,099 cases; accordingly, classification algorithms would assign these 846 cases randomly to one of the other classes (SA only or ST only). Classification algorithms do not perform well with highly disparate proportions of cases in the classes. The seriousness of ST in such ST+SA cases led us to combine them with ST-only cases rather than SA-only cases.

Despite the advantage of its size, many months of data cleaning and multiple analyses revealed not a single variable (#injuries of 25 types noted at 20 locations, for a total of 45 variables) associated with the outcome. Comparing the number of injuries in the two classes (SA; ST/ST+SA) at each of the 45 variables at a significance level of 0.10, one would have expected, with probability 0.948, at least one variable to have been significant by chance alone. The numbers of injuries of each type and at each location were nearly identical for the two classes.

Several inquiries led us to suspect inconsistencies in both the recording of the data and the translation of the data from hardcopy forms to the computer. We suspect that the "Exam Type" variable (ST, SA, ST+SA) may have been disassociated with the columns denoting the number of injuries in a random way. After much investigation into this database, we concluded that a high-quality database would be less likely to suffer from recording inconsistencies and difficulties in correctly transferring the data into the computer. We then hoped to include data from additional sites, but administrative difficulties rendered that plan unfeasible.

Outcomes

Activities/Accomplishments

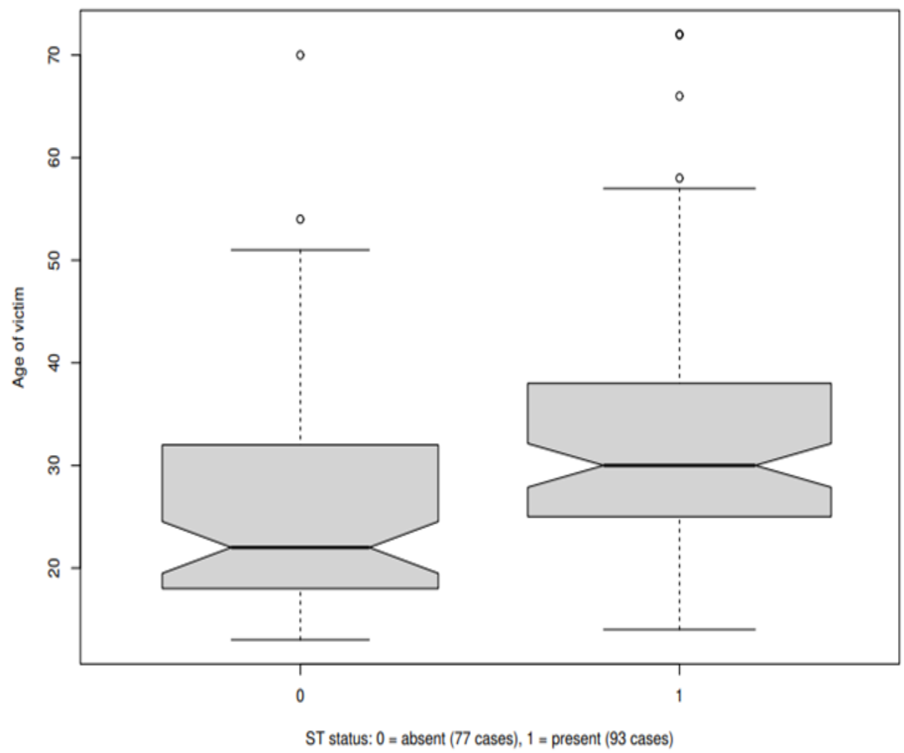
- We have built and refined a research database for medical record data entry. This database will be maintained moving forward to allow for future, larger analyses.
- Two nurses, in addition to the principal investigator (PI), were trained to enter data.
- Development of a potential algorithm to assist with classifying cases of strangulation given specific forensic examination characteristics.

Results and Findings

Descriptive Statistics

All 170 cases from the University of Virginia dataset were included in the sample for analysis. Among the 170 cases, 77 were documented as SA-only (recoded as *ST-absent*), 74 were documented as ST-only (recoded as *ST-present*), and 19 were documented as ST+SA (recoded as *ST-present*). In this study, the median age among the 77 *ST-absent* and 93 *ST-present* cases were 22 (quartiles 18, 32) and 31 (quartiles 25, 40), respectively, indicating slightly older victims involving strangulation than those experiencing sexual assault only. Boxplots of age by ST status (Fig. 2) demonstrate that ST-present victims tended to be older in this dataset: 70 percent (54/77) of *ST-absent* victims were under age 30, versus 47 percent (44/93) of *ST-present* victims. A large majority of the sample was documented as white (78%, n=132) and non-Hispanic (90%, n=154).

Figure 2. Age by Outcome Group



Note. Box width is scaled, so datasets with fewer observations (here, *ST-absent*: n=77) are narrower than those with more observations (here, *ST-present*: n=93). Non-overlapping notches indicate that the medians of the two populations may differ at the 5 percent level of significance.[11]

Table 2. Demographic Characteristics by Outcome Group (observed counts and percentages)

	Strangulation-Absent	*Strangulation-Present*
Race		
White	62 (81%)	70 (75%)
Black	10 (13%)	19 (20%)
Other	4 (5%)	2 (2%)
Unknown	1 (1%)	1 (1%)
Missing	0 (0%)	1 (1%)
Ethnicity		
Hispanic	68 (88%)	86 (94%)
Non-Hispanic	9 (12%)	5 (5%)
Missing	0 (0%)	2 (2%)

Feature Screening Using Chi-Squared Tests

The results of multiple independent chi-squared tests show that certain body locations and injury types impact ST case classification substantially. First, a chi-squared statistic was calculated to assess the difference in the proportions of petechial injuries present by ST status. Of the 77 ST-absent cases, only one (1.3%) showed a single petechial injury. Conversely, among the 93 *ST-present* cases, 11 cases (11.8%) showed a single petechial injury, four cases showed two petechial injuries, four cases showed three petechial injuries, two cases showed four petechial injuries, and one case showed six petechial injuries. Dichotomizing this variable as "petechiae present" versus "petechiae absent," the results of a chi-squared test statistic on one degree of freedom (17.2, p-value = 3.4e-05; cf. 95%-point 3.84) shows strong evidence of an association between ST status and presence/absence of petechiae (Table 3), suggesting that the presence of petechial injuries is highly indicative of *ST-present* versus *ST-absent*.

Table 3. Presence of Petechial Injury by Outcome Group

	Petechial Injury		
	Absent (n)	Present (n)	Total Cases (n)
Strangulation Outcome			
Absent	76	1	77
Present	70	23	93

We followed this strategy for other injuries to identify those that these data suggest might be related to ST outcome. For some injuries, we dichotomized them as we did for petechiae ("present" or "absent"); a chi-square statistic that exceeded 3.84 (corresponding to the 5% point of a chi-square distribution on 1 degree of freedom) suggested that the variable might be a useful discriminator. For other injuries (e.g., bruises), we first collapsed some values of k, the number of injuries, due to the wide range in the injury counts. As an illustration, Table 4 shows the number of cases that experienced k bruises, where k = 0, …, 28. Because most of the cells are zeroes, the validity of the "p-value" from Pearson's chi-squared statistic on 28 degrees of freedom (df) [= (#rows – 1) x (#columns – 1)] is doubtful, so we collapsed categories before computing the chi-square statistic; see Table 5.

Table 4: Number of cases with k bruise injuries (actual counts)

	k Number of Injuries (n)																		
	0	1	2	3	4	5	6	7	8	9	10	11	12	15	17	18	22	27	28
ST Absent	43	8	8	3	5	3	1	1	2	0	1	0	0	0	0	0	1	1	0
ST Present	30	11	14	3	11	4	2	2	2	2	0	4	1	2	2	2	0	0	1

Table 5: Number/Percentage of cases with k bruise injuries (collapsed counts)

	k Number of Injuries (%)		
	0	1-4	5-28
ST Absent	43 (56%)	24 (31%)	10 (13%)
ST Present	30 (32%)	39 (42%)	24 (24%)

Note. Chi-squared statistic 10.2 (df = 2), p-value = 0.0006.

We conducted 17 chi-squared tests on all injuries (total injuries and the 11 injury types listed in Table 1 and Figure 1) and on the five anatomical location codes, resulting in 17 p-values. Using a false discovery rate of 0.05,[12] variables having p-values less than 0.02 were retained for consideration in the logistic regression model. All but five variables (bite, laceration, incision, puncture wound, combined lower body) met this criterion. To include all anatomical locations (as variable combinations) in the model, we decided to include "combined lower body" in the model. Conversely, although "scratch" and "tenderness" did have p-values less than 0.02 (0.020 and 0.008, respectively), neither variable improved the model (in terms of either accuracy or predictive power). Finally, because the square roots of Poisson-distributed variables (as "counts" often are) tend to be more normally distributed,[13] and because Figure 2 suggested the value of age as a discriminator, we included the injury variables in the model as "sqrt(count)" as well as victim's age.

Logistic Regression

The chi-squared statistics discussed above guided the selection of variables to be included in a classification model. Logistic regression was used to classify the 77 *ST-absent* cases from the 93 *ST-present* cases. The most parsimonious model uses the victim's age, the five injury-location combinations (mouth, head-neck injuries, face, upper body injuries, lower body injuries), and the number of injuries due to scratches, redness, petechiae, abrasions, or bruises (scratch, redness, petechiae, abrasion, bruise). As indicated above, the "count" variables were entered into the model as "sqrt(counts)."

Table 6. Most parsimonious logistic regression model for classifying ST cases

Effect	Estimate	*SE*	z-score	*p*
Intercept	-2.167	0.640	-3.385	0.0007 *
Age	0.034	0.018	1.932	0.053
sqrt(Mouth)	-1.515	0.936	-1.619	0.105
sqrt(Face)	0.701	0.506	1.385	0.166
sqrt(Head & Neck)	1.363	0.452	3.018	0.003 *
sqrt(Upper Body)	0.383	0.481	0.795	0.427
sqrt(Lower Body)	-1.354	0.507	-2.671	0.008 *
sqrt(Swelling)	1.600	0.863	1.855	0.064
sqrt(Redness)	0.883	0.618	1.429	0.153
sqrt(Petechiae)	3.399	1.554	2.188	0.029*
sqrt(Abrasion)	-0.545	0.469	-1.161	0.246
sqrt(Bruising)	0.565	0.521	1.085	0.278

Note. * indicate p-values < 0.02.

The classification model was then used to determine the appropriate threshold for determining which cases are classified as *ST-present* or *ST-absent*. The logistic regression model expresses $y = logit(p) = log(p/(1-p))$, where p = probability of *ST*-present (vs. *ST-absent*), as a linear function of the variables in the model [here, age and sqrt(counts)]. A predicted logit of y = 0 (p = 0.50) indicates that the two possible outcomes (here, *ST-absent* and *ST-present*) are equally likely; hence, values above a threshold of 0 [respectively, below 0] indicated cases predicted to be *ST-present* [respectively, *ST-absent*]. In situations where the two classes have different sizes (here, 77 and 93) and/or one misclassification rate (e.g., false positive) is deemed more serious than another, often a different threshold may be more suitable. For these data, misclassification rates were calculated using threshold 0.5: cases with predicted values of y that exceeded 0.5 (p = 0.62) were classified as *ST-present,* while those for which y < 0.5 were classified as *ST-absent*. This choice of threshold took context into account: misclassifying an *ST-absent* case as *ST-present* (false positive) was deemed the more serious error. The false positive rate for the model in Figure 5 is 6.5% (5 of the 77 *ST-absent* cases were misclassified as *ST-present*) and a false negative rate of 22.6% (21 of the 93 *ST-present* cases were misclassified as *ST-absent*).

Simpler Classifier

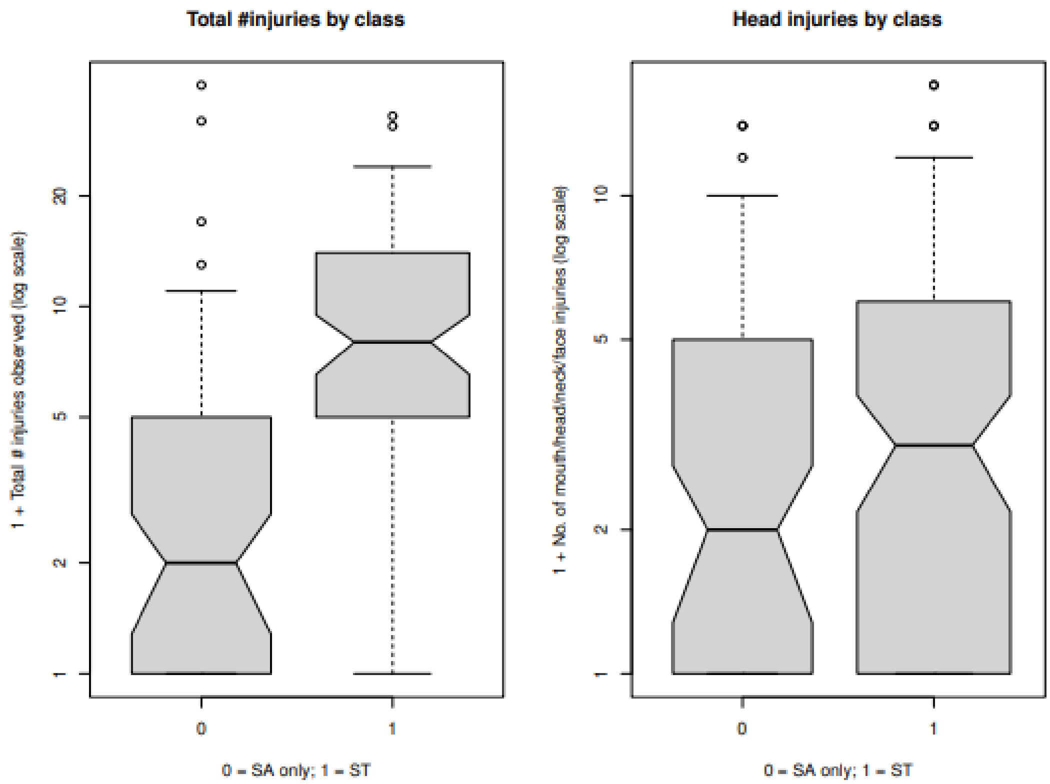
After noticing three significant characteristics of these 170 cases, a simpler classification model was developed and found to be almost as effective as the logistic regression model. First, the presence of petechiae is highly suggestive of ST cases. Among the 77 *ST-absent* cases, only one had petechiae (1.3%). Among the 93 *ST-present* cases, 23 had petechiae (24.7%). A rough 'likelihood ratio' derived from the ratio of these two proportions (*ST-present* versus *ST-absent*) is 19.0. The uncertainty in this ratio is large due to the relatively small counts in the numerator

(93 cases) and denominator (77 cases). Via simulation, assuming binomial proportions of 0.247 and 0.013 with 93 and 77 cases, respectively, the standard deviation is 12.8; 95 percent of the simulated likelihood ratios fell between 5 and 45.

Additionally, among the 146 cases that showed no evidence of petechiae, only three of the 76 ST-absent cases (96%) showed more than two injuries in the mouth, head, neck, and face combined; these three people showed five, seven, nine injuries across these four locations. Conversely, among the 70 *ST-present* cases that showed no evidence of petechiae, the majority (40, or 57%) showed more than two injuries across these four locations, and half of those 40 cases showed five or more injuries. The two classes also differed greatly in the number of total injuries: more than ten injuries were rare among ST-absent cases (3/76 = 4%) and more common among ST-present cases (14/70 = 20%).

The difference in the distributions of the number of mouth/head/neck/face injuries and the total number of injuries between the *ST-absent* and *ST-present* (those who did not experience petechiae) is illustrated with boxplots of the distributions in Figure 3. These differences, shown visually in Figure 3, can also be seen by the differences in the fitted distributions to the observed proportions of ST-absent/ST-present cases having exactly k total injuries, k=0, 1, 2, ... The former proportions are well approximated by a geometric distribution with a probability of 0.2567, while a geometric distribution with a probability of 0.1357 well approximates the latter. The predicted numbers of total injuries estimated from these fitted distributions are very close to the observed counts.

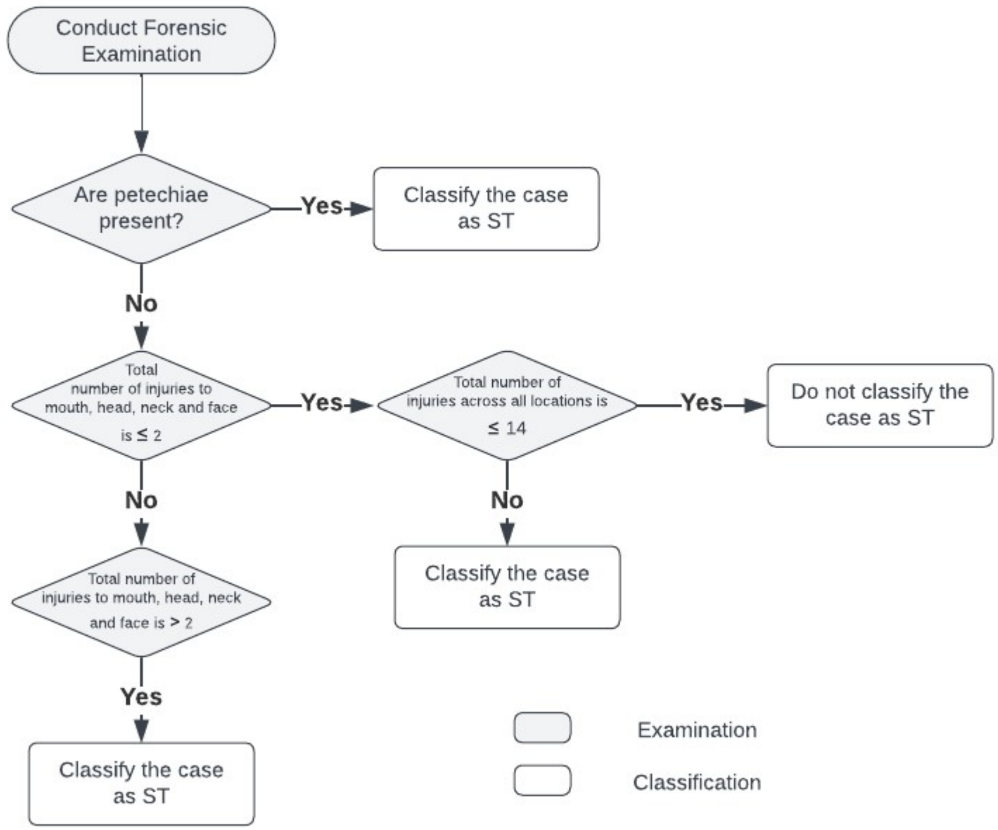
Figure 3. **Comparing Total Injuries and Head Injuries by Strangulation Outcome Among Participants without Petechial Injuries**



Note. "SA only" = ST-absent; "ST" = *ST-present*.

A "likelihood ratio" of the probability of observing k total injuries in an *ST-present* patient (without petechiae) relative to the probability of observing k total injuries in an *ST-absent* patient (without petechiae) was also calculated. At k = 15, this ratio is 4.0, suggesting that 15 or more injuries are four times more likely to have been observed in an *ST-present* patient than in an *ST-absent* patient. These observations led us to produce the following algorithm, shown in Figure 4.

Figure 4. Simple Strangulation Classification Algorithm



The final algorithm produced using this simple classifier assesses a case as "*ST-present*" if a case meets the following criteria: 1) petechiae is present or 2) petechiae is absent, but injuries to the mouth, face, head, neck total L or more, or the number is less than L, and the number of all injuries exceeds 14. As stated previously with the logistic regression model, classifying an *ST-absent* case as "*ST-present*" was deemed the more serious error. A value of L = 2 was selected as the false positive rate increases rapidly when L=0 or L=1. Using this algorithm with L = 2, four of the *ST-absent* cases were misclassified as "*ST-present*" (observed false positive rate 5.2%), and 30 of the 93 ST-present cases were misclassified as "ST-absent" (observed false negative rate 32.2%). Multiple runs of 200 bootstrap simulations on the original dataset confirmed these error rates with standard errors of 0.2 percent and 0.4 percent, respectively.

Limitations

There are several limitations to this work. They include:

- Examiner bias. When strangulation is suspected, the examiner looks more closely for signs known to be associated with strangulation. We attempted to mitigate this issue by using forensic sexual assault cases as the control (versus other forms of assault or injury), as those patients also receive a detailed examination for injuries.
- Representativeness: It is known that skin color may make injuries more or less visible. This algorithm was developed with a small sample of predominantly white, non-Hispanic patients and may not accurately reflect the findings for other populations. Future work is needed using larger datasets that are more representative of the population.

Before applying this algorithm in real-world settings, additional work is needed. Future studies should repeat this analysis with a larger and more diverse sample to refine these results further.

Study Artifacts

List of products (e.g., publications, conference papers, technologies, websites, databases), including locations of these products on the Internet or in other archives or databases

- Laughon, K, Kafadar, K. (2023). Probabilistic Modeling of Strangulation Injuries: Interim Analyses. Academy on Violence and Abuse Global Health Summit, Salt Lake City, Utah. (Virtual)
- Laughon, K, Kafadar, K. (2021) Probabilistic Model of Strangulation Injuries: Interim Results. International Association of Forensic Nurses. (Virtual)
- Laughon, K.; Cimino, A.; Kausch, S. (2020). Enhancing the Foundational Validity of Forensic Findings in Strangulation Examinations. Presentation at the Nurses Network on Violence Against Women, International, Malmo, Sweden. (Conference canceled due to COVID).

- Enhancing foundational validity of forensic findings in medico-legal strangulation examinations. Manuscript in submission to the *Journal of Forensic and Legal Medicine*.
- A research database of forensic patients' findings is maintained at the University of Virginia.

Data Sets Generated

A dataset of 170 patients (strangled and sexually assaulted) that includes demographic variables, exam characteristics, and injury findings (type and location) is archived at the National Archive of Criminal Justice Data (See https://www.icpsr.umich.edu/web/pages/NACJD/index.html).

Dissemination Activities

Interim findings have been disseminated at two major injury and violence conferences. This has allowed the principal investigator to discuss the importance of strengthening the scientific validity of expert testimony by examiners. There is a manuscript in submission to the *Journal of Forensic and Legal Medicine* and an additional manuscript of descriptive findings in preparation.