NIST Interagency Report
NIST IR 8352sup1

CSAFE Bitemark Thinkshop Report

Center for Statistics and Applications in Forensic Evidence (CSAFE)

This publication is available free of charge from:
https://doi.org/10.6028/ NIST.IR.8352sup1

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Interagency Report
NIST IR 8352sup1

CSAFE Bitemark Thinkshop Report

Center for Statistics and Applications in Forensic Evidence (CSAFE)
Ames, IA

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8352sup1

September 2022

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

NIST IR 8352sup1
September 2022

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose. This work was partially funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreements 70NANB15H176 and 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College and University of Nebraska, Lincoln.

Publication History
Approved by the NIST Editorial Review Board on 2022-09-07

How to Cite this NIST Technical Series Publication
Center for Statistics and Applications in Forensic Evidence (2020) CSAFE Bitemark Thinkshop Report. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency Report (IR) NIST IR 8352sup1. https://doi.org/10.6028/NIST.IR.8352sup1

Contact Information
alicia@iastate.edu
john.butler@nist.gov

Background Information

This report is supplemental information to NISTIR 8352 **Bitemark Analysis:** ***A NIST Scientific Foundation Review***. See also https://www.nist.gov/forensic-science/scientific-foundation-review-bitemark-analysis.

The NIST Special Programs Office requested that the Center for Statistics and Applications in Forensic Evidence (CSAFE), a NIST Forensic Science Center of Excellence, organize a Bitemark Thinkshop to gather input from the bitemark analysis community and its stakeholders. This event was called a thinkshop (rather than a workshop) as the purpose was to hear points of view on scientific questions important to bitemark analysis claims. The thinkshop, which was held October 17-18, 2019, in Arlington, Virginia, included almost 50 participants from across the United States as well as representatives from Australia, Canada, New Zealand, and the UK. Facilitators and notetakers from SNA International (Alexandria, VA) were key to encouraging valuable discussion during the event, capturing what was said by participants, and preparing a summary report.

An 18-member steering committee, organized by Rich Cavanagh and Karen Reczek from NIST, selected the invited participants and introductory speakers, defined the meeting format, and decided on topics for discussion. Following introductory remarks to the entire group from six speakers where background information was provided, participants were divided into three groups of 12 to 15 individuals each to facilitate discussion on specific questions predefined by the steering committee. The composition of each discussion group was shuffled over the two day event to maximize exposure to different perspectives. At the end of each breakout session, the entire group reconvened to hear a summary of what had been discussed in each discussion group.

Those invited to participate in the October 2019 Bitemark Thinkshop were Kenneth W. Aschheim, Charles Michael Bowers, Cynthia Brzozowski, John Butler, Anthony Cardoza, Alicia Carriquiry, Rich Cavanagh, Sarah Chu, Mary Cimrmancic, Alexander Comsudi, Glinda Cooper, Greg Davis, M. Bonner Denton, Robert Dorion, Gregory Dutton, Alex Forrest, Adam Freeman, Lynn Garcia, Will Guthrie, Barbara Hervey, Karen Kafadar, Gerry LaPorte, Julia Leighton, John McDowell, Robin Mejia, Roger Metcalf, John Morgan, Stephen Morgan, Marcela Najarro, Niki Osborne, Chris Plourd, Rich Press, Iain Pretty, Karen Reczek, Jeff Salyards, David Senn, Barbara (Bobbie) Spellman, Hal Stern, Isiah Warner, Frank Wright, Mark Bernstein, Jonathan McGrath, Danielle McLeod-Henning, and Shannan Williams.

The entire summary report received by NIST from CSAFE on January 27, 2020 (and from SNA International to CSAFE on January 21, 2020) is provided here. Points of view presented at this thinkshop and in the provided report are those of the presenters and authors and do not necessarily represent the official position or policies of the National Institute of Standards and Technology.

Iowa State University
195 Durham | Ames, IA 50011
515-294-7278
ForensicStats.org

csafe
Center for Statistics and Applications in Forensic Evidence

Susan Ballou
Acting Director
NIST Special Programs Office

Date: January 27th, 2020
RE: Bitemark Thinkshop Report

Dear Sue,

Enclosed is our report from the Bitemark Thinkshop, which was held October 17-18th, 2019 in Arlington, VA. The report we are submitting includes both a brief CSAFE-authored summary of the Thinkshop and a more detailed report authored by SNA International. The latter includes all of the raw notes produced during the Thinkshop and a detailed summary of the Thinkshop presentations and breakout group discussions.

Submission of this report concludes our official effort on this project. If you have additional questions or need any additional work performed, please contact me.

With warm regards,

Alicia Carriquiry

Alicia Carriquiry
Distinguished Professor
Director of CSAFE
Iowa State University

CSAFE Notes on the Bitemark Thinkshop

Background: The Center for Statistics and Applications in Forensic Evidence (CSAFE), a NIST-funded Center of Excellence, held a Bitemark Thinkshop in Arlington, VA on October 18 and 19, 2019. The Thinkshop brought together bitemark stakeholders from several relevant communities including odontologists, lawyers, statisticians, and researchers. The aim of the Thinkshop was to explore open questions and knowledge gaps in the area of forensic bitemark analysis to inform a NIST Scientific Foundation Review of the discipline. The National Commission on Forensic Science made a recommendation in September 2016 that NIST establish the capacity to perform independent scientific evaluations of the technical merit of the methods and practices used in forensic science disciplines. With funding from Congress, NIST agreed to undertake a series of reviews as an experiment; bitemark analysis is one of those experiments. The Thinkshop is the first part of the NIST evaluation that will culminate with a public report from NIST on their findings regarding the foundations for performing forensic bitemark analyses. A planning group comprised of NIST personnel and representatives of the relevant communities designed the Thinkshop. The Thinkshop included a half-day of presentations to provide participants with some background on the current practice of forensic bitemark analysis followed by 1.5 days of small group discussions around topics relevant to the scientific foundation review of forensic bitemark analysis. CSAFE contracted with SNA International to facilitate the group discussions, record notes from the discussions, and produce a final report summarizing the Thinkshop (provided separately).

The Thinkshop: The Thinkshop began with a series of presentations: a charge to the group from NIST's Dr. Richard Cavanagh; an introduction to bitemarks by Dr. David Senn and Dr. Robert Dorion; a presentation by forensic pathologist Dr. Gregory Davis; a literature review of bitemark analysis by Dr. John Morgan; and a discussion of bitemarks from a standards development perspective by Dr. Kenneth Aschheim. Subsequently, three breakout sessions focused on (1) understanding human dentition, (2) understanding bitemark patterns, and (3) data interpretation strategies. Each breakout session was motivated by a single scientific question (see below). To facilitate small-group discussions, each breakout session was further divided into three focus areas (subtopics within the overall breakout session topic). Small groups (10-15 people with different expertise) discussed each focus area within each breakout group. Every individual thus had a chance to comment on every topic. The composition of the small groups was varied across the three breakout sessions so that individuals were able to hear more varied opinions. SNA has provided a detailed report that includes a detailed description of the agenda and discussion topics as well as a detailed summary of the breakout discussions. The aim of these notes is to provide a summary of key points that CSAFE, as conveners of the Thinkshop, believe are relevant and important to NIST's Scientific Foundation Review of forensic bitemark analysis.

Science Question 1 (Understanding Dentition): Are there measurable characteristics or features in human dentition that vary among individuals and are persistent within an individual?

The first question focused on human dentition as distinct from bitemark impressions. The questions addressed concerned uniqueness of human dentition, best measurement methods, understanding the population, and relevant features and parameters. There was broad consensus that the technology exists to obtain reliable measurements of human dentition. Odontologists agreed that 3-D scanning equipment works well for this purpose. There was some discussion of the challenges associated with taking 3-D representations and presenting them as 2-D overlays as is commonly done in forensic

bitemark examinations. The appropriate 2-D overlay in a particular case depends several aspects of the pattern evidence being reviewed.

There was some discussion on the question of whether each individual's dentition is unique; this discussion was motivated in part by historical cases in which claims of identification based on bitemark evidence suggested the ability to uniquely identify the biter. Thinkshop participants quickly reached consensus however that this question is not relevant for a number of reasons. First, even if dentition is unique it is essentially impossible that a bitemark impression in human skin would record sufficient detail to make an identification to a single source. Second, current ABFO Standard and Guidelines do not sanction identification as a final conclusion in an open population case. Thus, the group did not discuss uniqueness of dentition further.

Given that current guidelines allow final conclusions that exclude a questioned dentition as the source of a bitemark or don't exclude the questioned dentition, the group did feel it important to discuss the need for relevant population information in order to assess the size of the population of potential sources that would not be excluded. There was some discussion of generating databases of 3-D dentition scans or casts through partnerships with orthodontic companies or university clinics. At present a relevant database does not exist.

Science Question 2 (Understanding Bitemarks): Do bitemarks transfer measurable characteristics of the dentition to the substrate?

The second question focused on injury patterns found on human skin and the subgroups discussed a range of data collection strategies as well as the amount variation to be expected in repeated bitemarks from the same dentition. The original Thinkshop plan envisioned a discussion of other substrates but all discussion ended up focusing on injury patterns found on only human skin thought possible to be a human bitemark. There was generally good agreement on the approaches relevant to the collection of relevant information regarding potential bitemark injury patterns. These include photographic documentation (potentially obtained serially over time), possibly including infrared, ultraviolet and alternate light photographs when relevant. Impressions may also be taken if 3-D properties are present, but the consensus was that this is not common. Tissue samples may be taken in the case of a deceased victim.

The first determination required by the ABFO Guidelines and Standards is whether a particular injury was caused by a human bitemark. There was some disagreement about the reliability of such determinations. It was noted that injuries from animal bitemarks may resemble injuries from human bitemarks as might other non-bitemark injuries. This question is discussed further below under Science Question 3. There was agreement that injury pattern evidence could be useful as an investigative tool, especially when there appeared to be evidence of unusual characteristics associated with the cause of the injury (perhaps dentition); there was not agreement over the course of the Thinkshop about using assessments of the injury pattern as evidence at trial.

Assuming for the moment that a pattern injury could be identified as a human bitemark, it was noted that there were many sources of variability that would be expected to impact bitemarks from a given fixed dentition. These include the amount and direction of pressure applied, characteristics of the substrate (area of the body), age and medical condition of the bitemark victim, duration of the bite, motion of the biter and/or victim, and presence/absence of clothing at the bitemark. It is of great interest to learn more about these factors and their relative importance. The study of bitemarks is challenged however by the difficulty in collecting data about variation in bitemarks through human studies. All noted that IRB approval for such studies is a challenge, and concern was expressed that, even if experiments are permitted, the resulting data may not reflect bitemarks produced in the commission of a crime.

Science Question 3 (Data Interpretation Strategies): What interpretation strategies (techniques and practices) produce the most accurate and reliable results?

Having discussed dentition and bitemark patterns, the groups' final discussion focused on strategies for assessing the use of pattern injuries on skin as forensic evidence. The discussion included a review of ABFO Standards and Guidelines (most recently revised in February 2018), a consideration of whether they are followed in practice, and the need for studies of reliability and validity for determinations made about bitemark evidence. The most recent ABFO Guidelines and Standards describe best practice procedures that were supported by the group. This includes blinding (separate odontologists should collect and analyze the evidence), the use of dental lineups (multiple candidate dentitions), removal of irrelevant contextual information, and the need for independent verification. Concerns were expressed that these relatively new guidelines were not always followed. Some called for data to be collected on the frequency with which these procedures were followed.

There was wide agreement that bitemark analysis is a subjective examination. To many this means that there is a critical need for assessments of reliability and validity. Traditional use of the terms in the measurement science world considers reliability as being focused on whether measurements or determinations are consistent and validity as being focused on whether determinations are accurate or correct. When it comes to validity, it was noted at the Thinkshop that one reason for the criticism of forensic bitemark analysis is the number of exoneration cases in which bitemark evidence played a role (more than 30). Some downplayed the significance of these exoneration cases by noting that they reflected poor practice that violated current guidelines. Others argued that without data on the validity of assessments made under current guidelines it is difficult to be confident that errors are not continuing to occur. Additionally, the group discussed the importance of obtaining a better understanding of the number of cases in which bitemark evidence has been used to better assess the number of exonerations that have occurred.

Given the lack of casework data with known ground truth and the challenges associated with collecting experimental data (as described above) to assess validity, the group ended up spending more time on the assessment of reliability. Reliability focuses on consistency of measurements and determinations. Studies can include assessments of intra-individual reliability, sometimes known as repeatability, and inter-individual reliability, sometimes known as reproducibility. For bitemark analysis such determinations would need to be made for the two main questions addressed: determining whether a pattern injury is or is not a human bitemark, and determining whether a questioned dentition could have been the source of the bitemark injury. It was noted that one earlier, unpublished study, found results suggesting limited reliability which naturally implies limited validity. The authors of the unpublished study are now planning to submit the results for publication (with some previous obstacles recently removed).

Moderated Final Discussion: Conclusions and Takeaways

The Thinkshop closed with a whole-group moderated discussion focused on identifying conclusions and next steps. The discussion grew heated with a number of individuals advocating that the group should reach a decision regarding whether to pause the practice of forensic bitemark analysis while research was carried out. As funders of the event NIST determined that such a decision was not part of the scope of the activity.

There emerged a very strong consensus that there is a critical need for research to explore the scientific foundations for bitemark analysis. It was proposed that this research should begin with a study assessing intra-individual and inter-individual variation in determinations of whether a given injury is a human bitemark. This is essential but it should be emphasized that this must be only the first step in a series of relevant studies. If determinations of the type of bitemark can be made by odontologists

with high reliability, then it is critical to assess the accuracy of these determinations using injury marks where the ground truth of their cause is known. We must recall that reliability is a requirement for validity but it is not a guarantee of validity. And if determinations of whether an injury is a human bitemark are found to be reliable and valid, there will be the need for studies on the reliability and validity of determinations made in linking dentition to bitemarks. Additional studies should consider the interpretation by law enforcement and the legal system of a finding that a questioned dentition could not be excluded as the source and whether it is possible to assess the size of the pool of sources that would not be excluded.

NIST CSAFE Bitemark Thinkshop

October 17th – 18th, 2019

Report

January 21, 2020

Table of Contents

1. EXECUTIVE SUMMARY

Following a September 2016 National Commission on Forensic Science recommendation and with funding from the U.S. Congress, the National Institute of Standards and Technology (NIST) is carrying out a Scientific Foundation Review for bitemark analysis. A critical first step in this review was a Bitemark Thinkshop held on October 17-18, 2019. The Thinkshop convened bitemark stakeholders to focus on points of agreement and disagreement, challenges, knowledge gaps, and next steps within the field of forensic bitemark analysis. These stakeholders included dentists, researchers, statisticians, legal experts, and image analysts. The two-day meeting focused on three scientific questions posed to the group and consisted of presentations, breakout sessions, and moderated whole-group discussions. Discussion focused on the next steps for research needed to establish a solid scientific foundation for bitemark analysis. The Thinkshop culminated in a whole-group review to determine next steps for exploring whether bitemark analysis can be placed on a firm scientific foundation. After two days of discussion, the participants concluded there is an urgent need for fundamental research to focus on the reliability and validity of the various determinations made in bitemark analysis: Is an injury a human bitemark? Can a questioned dentition be related to a bitemark? As a starting point, the group recommended a study to focus on the question:

"Can it be determined if a pattern or pattern injury is a human bitemark?"

Prior to conducting this study, the group agreed that a detailed plan should be developed and agreed upon by a wide range of stakeholders to ensure that the results are broadly accepted. The study plan should include sample size, sample types, outcome measures, etc.

It should be noted that there is redundancy in the various parts of this report. This reflects the design of the Thinkshop, which enabled various topics to be revisited in different focus area discussions.

2. INTRODUCTION

The use of bitemark evidence in the courtroom can reportedly be traced back to the Salem Witch Trials of 1692. Since then, one of the most famous criminal cases in the U.S. involving bitemark analysis was that of serial killer Ted Bundy. In 1979, Bundy was convicted of 36 murders, one of which involved two bitemarks. In 1984, the American Board of Forensic Odontology (ABFO) first established guidelines for forensic bitemark analysis, and currently, bitemark evidence is admissible in most courts. However, opinions on the reliability and scientific validity of forensic bitemark analysis continue to vary. In a 2009 National Academy of Science report, *Strengthening Forensic Science in the United States: A Path Forward*, it was noted that bitemark identification testimony has been "introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing." In recent years, several convictions based on bitemark evidence have been overturned using DNA analysis. In addition, in 2016, the Texas Forensic Science Commission issued a decision recommending a moratorium on the use of bitemark evidence in future criminal prosecutions in the state of Texas until forensic bitemark analysis could be scientifically validated. At that time, the Texas Forensic Science Commission also ordered a review of every conviction in Texas that used forensic bitemark analysis.

In September 2016, the National Commission on Forensic Science recommended that the National Institute of Standards and Technology (NIST) perform Scientific Foundation Reviews for a range of forensic science disciplines. Congress provided funding for a number of Scientific Foundation Reviews. NIST selected a range of disciplines for these initial reviews, including forensic bitemark analysis. A Scientific Foundation Review is a study that seeks to document and evaluate the foundations of a scientific discipline. As a first step in the review of forensic bitemark analysis, NIST, in collaboration with CSAFE, organized this Thinkshop to carry out an in-depth discussion of relevant issues. The goal of the Bitemark Thinkshop was to bring together bitemark stakeholders to focus on the current challenges and knowledge gaps within the field. Attendees of the Thinkshop included dentists, researchers in different disciplines, statisticians, legal experts, and image analysts.

2.1. Purpose and Methodology

The purpose of this Thinkshop and the Scientific Foundation Review was to "document and consolidate information supporting (or opposing) the methods used in forensic bitemark analysis and to identify knowledge gaps where they exist." The stakeholders were expected to identify:

- Disagreements
- Agreements
- Knowledge gaps
- Next steps

The two-day Thinkshop consisted of presentations, breakout sessions, moderated discussions, and culminated in a whole-group review to finalize common themes and decide on a way ahead. NIST opened the Thinkshop by outlining the meeting format, process, and objectives. The Thinkshop was collaborative, using the diverse expertise of the participants to identify gaps and challenges to forensic bitemark examinations and comparisons. Three scientific questions were posed to the group to focus discussion and determine whether it is possible to develop a scientific foundation for forensic bitemark analysis.

It should be noted that there is repetition in the raw notes (Appendix C) due to the Thinkshop format where, in an effort to be thorough and transparent, Groups A, B, and C were all tasked with discussing the same topics. All the Thinkshop participants were highly engaged and enthusiastic about expressing their viewpoints, and the Thinkshop notes represent the note recorders' best efforts to reflect the ideas expressed by all Thinkshop participants fairly.

2.2 Speaker Presentations

Following a NIST introduction and overview, seven presentations established a common foundation for the discussions that followed.

Richard Cavanagh PhD, Director of NIST Special Programs Office-Bitemark Thinkshop: *Why Are We Here?*

Dr. Cavanaugh spoke of the efforts made by NIST to address concerns about forensic evidence raised in the 2009 National Research Council (NRC) report. This includes NIST participation in many organizations, including the National Commission on Forensic Science (NCFS) and the Organization of Scientific Area Committees for Forensic Science (OSAC), the NIST Forensic Science Research Program, the formation of the NIST-funded Center of Excellence known as the Center for Statistic and Applications in Forensic Evidence (CSAFE),

and the program of Scientific Foundation Reviews. Dr. Cavanaugh noted that NIST chose to start its program of Scientific Foundation Reviews with a range of disciplines. Bitemark analysis, the subject of the Thinkshop, is included as an example of a field with limited foundational science. NIST has organized this Thinkshop to document and consolidate information supporting (or opposing) methods used in forensic bitemark analysis in a brief, intensive, exploratory program that focuses on open challenges and knowledge gaps. Dr. Cavanaugh also stated that NIST has organized this Thinkshop group to include a mix of perspectives from odontologists, researchers, statisticians, legal experts, as well as image analysts and emphasized that NIST values all these perspectives.

Robert Dorion, D.D.S., Current chair of ABFO Bitemark Evidence & Patterned Injury Committee: *Bitemarks 101*

Dr. Dorion spoke about the number of wrongful convictions identified by critics as being based on bitemark evidence, such as the case involving Ray Krone. Krone was convicted of murder in 1992 using bitemark evidence and exonerated in 2002 after serving ten years in prison. Dr. Dorion noted that there were indeed bitemark errors in many of these cases but that there were other contributing factors as well. Dr. Dorion pointed out that the bitemark analysis processes, including traditional and alternative light photography, bitemark impressions, scanning, and macroscopic and microscopic analysis could benefit from additional research. Dr. Dorion acknowledged that bitemark patterns are impacted by a number of intrinsic and extrinsic factors, and that environmental exposure affects bitemark analysis. These too could benefit from additional study. Dr. Dorion pointed out that, unfortunately, despite these research needs, no bitemark research proposal has been funded since 2009, and the 2016 President's Council of Advisors on Science and Technology (PCAST) report advised against devoting significant resources to bitemark research efforts.

David Senn, D.D.S., Clinical Assistant Professor, University of Texas Health Science Center at San Antonio: *Bitemarks 101*

Dr. Senn stated that there have been very few studies conducted with bitemarks on human skin, partly because there have been challenges obtaining Institutional Review Board (IRB) approval of bitemark studies on living human skin. Dr. Senn then presented a case of a 29-month-old child who was bitten multiple times. Four suspects were identified, including the mother, babysitter, babysitter's boyfriend, and the babysitter's 5-year-old son. Based primarily on photographs taken at the hospital, an odontologist concluded that a child made the bitemarks on the victim because of class characteristics identified in the bitemarks. Therefore, Dr. Senn asserted, bitemark analysis can be a probative forensic examination and was useful in this case.

Gregory G. Davis, MD, MSPH
Professor and Director, Forensic Division, Pathology Department, UAB
Chief Coroner/Medical Examiner, Jefferson County, Alabama: *Mechanisms of Skin Injury with Application to Pattern Injuries*

Dr. Davis postulated that bitemarks may be compared to a pattern injury on human skin from a weapon where the weapon is human dentition. Dr. Davis emphasized that human skin is elastic; forensic pathologists may be able to provide information about the mechanism causing a pattern injury on skin, but no information can be ascertained concerning the exact item which caused the pattern injury. Dr. Davis pointed out that different types of injury can occur to the human skin, including abrasions, contusions, and lacerations, which may or may not produce

useful information regarding the weapon used to cause the pattern injury. In addition, tissue inflammation caused by trauma, individual differences in skin types, and variable thickness of the skin can all cause errors in the measurement of pattern injuries on human skin. In his experience, Dr. Davis said it is often difficult to determine the exact instrument which caused a pattern injury on human skin, and forensic examiners should not say more than they can about the exact cause of the pattern injury. Dr. Davis emphasized that any injury shown to a jury must be understandable to its members with some explanation and asserted that if the forensic examiner states, "You'll just have to trust me on this," then the examiner has gone too far in the interpretation of the pattern injury.

John Morgan, PhD, Senior Director, Center for Forensic Science, RTI International: ***Scientific Foundations of Bitemark Examinations***

Dr. Morgan listed nine assumptions bitemark examiners make about their ability to associate a mark on human skin with an individual's dentition:

- Bitemarks remain on skin and are persistent to a sufficient degree to permit examination within a time frame relevant to police investigation.
- Bitemark patterns may be distinguished from other skin patterns not associated with human bites.
- Bitemark pattern distortions are known and can be distinguished from bitemark patterns.
- It is possible to collect and analyze skin patterns in a reproducible and quantitative way.
- The statistical likelihood that a set or subset of dentition patterns is unique to an individual is known.
- Human dentition is persistent to a sufficient degree to permit examination and comparison relevant to police investigation.
- It is possible to collect and analyze human dentition patterns in a reproducible and quantitative way.
- The process practitioners use to extract and match features accurately and reliably relates human dentition patterns to bitemark patterns.
- The probability of a particular feature match relative to all other feature matches has been established.

Dr. Morgan then reported on the interim results of an ongoing literature review regarding the validity of these assumptions.

Kenneth W. Aschheim, D.D.S., Associate Clinical Professor, Oral & Maxillofacial Surgery, The Mount Sinai Hospital: ***Forensic Standards and Bitemarks: A Path Forward***

Dr. Aschheim asked, "How do we make some change?" Dr. Aschheim noted that there had been cases in which bitemark evidence contributed to wrongful convictions but also noted that there is potential information in bitemark patterns in some cases. He asserted that bitemark analysis standards are badly needed and reviewed the standards development process. Dr. Aschheim explained that creating these standards will be a balancing act between two opposing theories (law vs. science). Dr. Aschheim said that the American National Standards Institute (ANSI) defines a standard as a document, established by consensus, that provides rules, guidelines, or characteristics for activities or their results and that the ANSI standards

development process includes consensus, broad-based public review, response to comments, availability of an appeal, and is subject to a neutral third party. Dr. Aschheim listed the organizations currently working on forensic odontology standards: the International Organization for Standardization (ISO), the Auditing Standards Board (ASB), the American Board of Forensic Odontology (ABFO), the Standards Development Organizations (SDO), OSAC, and ANSI-NIST. Dr. Aschheim said that the standards development process can be difficult. It often requires the melding of many different viewpoints; it may involve talking points that will be debated; it needs a non-judgmental environment where good ideas can flourish, and less proven methods can wither.

Amanda Sozer, PhD, Owner, SNA International: ***Breakout Overviews***

Dr. Sozer explained that SNA is tasked with capturing the information from this Thinkshop and putting it into a useful format. She assured attendees that the Thinkshop is designed to promote discussion and indicated that everyone present will have the opportunity to tackle the same topics in turn. Dr. Sozer explained that SNA facilitators and note-takers will capture ideas from the Thinkshop discussions and assemble a report of the results which will be provided to CSAFE and NIST. Finally, she said that the final NIST Thinkshop report will be made publicly available in a National Institute of Standards and Technology Interagency Report (NISTIR).

2.3 Breakout Groups

After the presentations, the attendees divided into three breakout groups. The groups included statisticians, researchers, attorneys/jurists, policy experts, victim advocates, imaging experts, and a NIST representative.

The three scientific questions, defined by the Bitemark Thinkshop Planning/Steering Committee, referred to three major topics: "Understanding Dentition (Teeth)," "Understanding Bitemarks," and "Data Interpretation Strategies." Each of these three major topics was further broken down into three sub-topics to narrow the scope of the breakout discussions.

In breakout groups on day one, attendees discussed scientific topics 1 and 2 and their associated claims, and on day two discussed scientific topic 3. During the breakout sessions, each group convened in a separate room. A facilitator and a recorder were assigned to each group to guide and record the discussion. Each breakout group rotated through the three focus areas for each question (see below):

Discussion Question 1: Dentition			**Thursday, 1:00 – 3:00 PM**
	1:00 – 1:35 pm	**1:40 – 2:15 pm**	**2:20 – 2:55 pm**
Group A	Room 1	Room 2	Room 3
Group B	Room 2	Room 3	Room 1
Group C	Room 3	Room 1	Room 2
Discussion Question 2: Bitemarks			**Thursday, 3:15 – 5:15 PM**
	3:15 – 3:50 pm	**3:55 – 4:30 pm**	**4:35 – 5:10 pm**
Group A	Room 1	Room 2	Room 3
Group B	Room 2	Room 3	Room 1
Group C	Room 3	Room 1	Room 2
Discussion Question 3: Analysis and Interpretation			**Friday, 10:15 – 12:15 PM**
	10:15 – 10:50 am	**10:55 – 11:30 am**	**11:35 – 12:10 pm**
Group A	Room 1	Room 2	Room 3
Group B	Room 2	Room 3	Room 1
Group C	Room 3	Room 1	Room 2

This rotation allowed each participant the opportunity to provide input in all areas. To ensure as much interaction as possible among participants, the composition of the three breakout groups varied for each science question. The facilitators began each session by reviewing the science question and the claim associated with it, then encouraged participants to focus discussion on the specific sub-topic assigned to that break-out room. The recorded discussions about each sub-topic were then combined and synthesized by the facilitators.

The conversations regarding each scientific question culminated in a moderated discussion, with all attendees participating, to review the information gathered during each of the break-out sessions. The moderated discussion summarized the agreements and disagreements expressed during the breakouts and provided an opportunity to solicit additional input or clarification.

The final event on day two of the Thinkshop was a moderated panel discussion with all attendees. This session provided participants with an opportunity to revisit the discussions from the entire Thinkshop and to agree on next steps. While no consensus was reached regarding the current value of bitemark analysis as a forensic tool, the majority of participants agreed on the need for basic foundational research.

2.4 Brief Overview of Current Practices Regarding Bitemark Analysis

Current practices regarding forensic bitemark analysis can be found in the ABFO Standards and Guidelines for Evaluating Bitemarks. The ABFO standards and guidelines include best practices for evidence collection, case information, chain of custody, documentation, photography, swabbing, and capturing impressions from the bitemark. The ABFO standards also include general considerations for best practices in evidence collection from suspects, including photographic records, intraoral examinations, impression lifting, and casting. In addition, the ABFO standards include high-level best practices for bitemark analysis, comparisons across injuries or between an injury and dentition impressions, and reporting of results of the analysis.

3. SCIENCE QUESTION 1: DENTITION

Are there measurable characteristics or features in human dentition that vary among individuals and are persistent within an individual?

Claim: Characteristics of human dentition are unique or can be divided into reliable fractions of the population, provided consideration of any changes with morphometric parameters over time and events.

3.1. Breakout Session

3.1.1. Focus Area A

What measurement method(s) provide the best information for capturing reliable information about the dentition?

The measurements that are the most useful in characterizing dentition are the width and height of the teeth, spacing between teeth, and arch measurements. After taking dental impressions, dental casts are made, and measurements are taken on the casts. Dental casts may vary among casting media, and casts may be subject to wear and damage over time. A ruler and protractor are usually used to measure the size and relative position of teeth in the dental arch. It was noted that dentition measurements vary among practitioners depending on the precise

location of the measurement points and that there is the need for a standard protocol for measurements. Participants mentioned using 3D orthodontic scanning tools to obtain consistently accurate representations of the dentition. They generally agreed that 3D imaging would be best for exemplars. However, participants noted a lack of studies on the conversion of a 3D image to a 2D overlay that can be used in comparison with bitemark pattern injuries.

Measurement variability among practitioners is not related to the ability to measure dental characteristics, but, more importantly, what features are the most informative and characteristic for an individual's dentition? There is a lack of agreement on what should be measured and precisely where to take the measurements. Determining what part of the tooth should be measured, and the resolution of the measurement is challenging. What is measured and how the measurement is made can impact forensic bitemark analysis results. Measurement variability should be studied to determine its impact on characterizing dentition and bitemark analysis.

Participants generally agreed that comparison studies are needed among different models of 3D capture devices. While 3D scanning companies have instrument performance data, the information may be proprietary and difficult to access.

3.1.2. Focus Area B

How do we appropriately collect information to create population databases that can be used for scientific and statistical analysis of human dentition?

Participants agreed that there is no database to demonstrate the uniqueness of human dentition. They also generally agreed that proving "uniqueness" or individualization of human dentition is highly unlikely because the characteristics of dentition that might be used for defining individuality do not exist. Attendees also said that forensic odontologists are looking for a scientifically based, reliable, and reproducible method to determine the rarity of an individual's dentition, and this does not require uniqueness. Forensic odontologists moved away from uniqueness when they went from identification of an individual from a bitemark to classifying a bitemark as "cannot exclude" an individual. Statistics from a database could provide weight to claims of the "cannot exclude" conclusion. In other words, forensic odontologists are looking for scientific studies to provide evidence for statements about the probability of an incidental association.

There are many variables to consider in creating a database of the characteristics of human dentition. These include but are not limited to: populations contributing, physical location of the individuals contributing, socio-economic make-up of individuals, and age of individuals. If a database were developed from data associated with orthodontic treatment, it would exclude people with straight teeth and/or those who cannot afford orthodontic treatment, and this would possibly affect the results of statistical analysis. In addition, dentition changes over time, and orthodontically treated teeth create more similar dentition. Therefore, long term studies are required, and those studies may need to consider orthodontically treated dentition separately.

Before creating the database, it is essential to define what characteristics would be measured for inclusion and how those measurements would be performed. Experts should measure class characteristics, such as arch form and intercanine distance, along with individualizing characteristics such as presence, arrangement, and size of individual teeth. Participants suggested that a survey of experts (both within and outside the community of odontologists) be undertaken to decide what features should be measured before defining which characteristics are to be

studied. For example, facial recognition experts might be consulted to discuss developing facial recognition algorithms. Most participants agreed that currently, existing databases containing large numbers of 3D representations of dentition are likely the best source of information for the creation and study of human dentition. However, access to this data may be difficult because of the proprietary nature of the data and possible privacy issues associated with it. In addition, dental models could be scanned and included in the database study.

Artificial intelligence (AI) tools may be used to assist with data analysis, for example, by examining data to obtain a list of appropriate points of variability or by identifying related populations for statistical analysis. Attendees suggested that forensic odontologists follow what anthropologists have been doing for many years to identify variable locations, populations, etc. They also suggested that to study human dental characteristics comprehensively, the database must contain hundreds of thousands of samples.

3.1.3. Focus Area C

What are the most probative features/parameters to use, and what are the limits associated with each?

Prior to consideration of the most probative features/parameters to use when examining bitemark evidence, the groups noted that the phrasing of the question, "what are the most probative..." is not objective and is open to interpretation. To understand what may be "unique" to an individual's dentition, it is first important to understand "normal" dental anatomy. The groups also noted, as mentioned earlier, that identifying "uniqueness" of dentition is not necessary for bitemark analysis.

Forensic dentists use the following characteristics when examining dental evidence:

- Presence/absence of dentition, including supernumerary teeth
- Gaps/spacing between teeth
- Arch form
- Intercanine distance
- Pathological characteristics, including wear, chips, etc.
- Odontological characteristics, including restorations, orthodontics, etc.
- Alignment and rotation of teeth, including angulation of individual teeth
- 3D orientation of individual teeth and overall dentition
- Width x Length x Height of dentition and individual teeth
- Individual and overall dentition pattern analysis is at the core of forensic dental exams - "It's the pattern that matters"

Participants noted there are many knowledge gaps in bitemark analysis. These knowledge gaps include but are not limited to:

- Lack of frequencies of dental features and long-term studies regarding the stability of the arrangement of teeth/dentition over time
- Lack of studies of individual or larger dental patterns and their stability over time
- No existing data to support which features/patterns may be most probative in bitemark analysis
- No studies conducted about the changes of highly variable traits over time

- Currently, no information to support differences in class vs. individual characteristics of teeth/dentition
- Current results of analysis do not give any level of confidence in the accuracy of the analytical results
- There is also no data to address whether levels of confidence for forensic dental exams varies from orthodontically vs. non-orthodontically treated dentition

When the result of a bitemark exam concludes that a known dentition "cannot be eliminated" as the cause of the injury, there is no level of confidence associated with that result. Participants said there is a need for an extensive database to study specific dental characteristics, including patterns of dentition and changes of these patterns over time to address these knowledge gaps. This study should include the review of the features of all teeth and not a selected few (for example, anterior teeth only). An extensive database study is necessary, and funding would be needed to create such a database. Data may currently be available, which could help with the study. The available data, however, are largely proprietary and unusable because the data belongs to companies such as iTero, CEREC, and E4D. Even if made available, it is unknown if data could be easily compared among the companies retaining them. There also may be ethical and legal issues around obtaining data from these companies without the consent of the patients. In addition, the data may only apply to specific socioeconomic groups whose information is retained by the dental imaging companies. The frequency of characteristics contained in these data may not be representative of the general population. In addition to possibly obtaining dental database information from private companies, this type of information might be available from university dental clinics. Attendees suggested that the University of Colorado's dental clinic might be a good source of data. The creation of a large database of dental characteristics would be an excellent project for a master's or PhD student. Researchers can study the characteristics contained in the database in many ways; their studies should be carefully planned and include examinations of which dental characteristics are independent/dependent of other characteristics. In addition, if more than one group of researchers addresses the creation and study of a dental characteristics database, consistent protocols must be created for all researchers. Funding should be appropriated to operate and maintain the database over time.

3.2. Moderated Discussion

There was consensus in the moderated discussion that 3D scanning is the best measurement tool to capture information about dentition. There is potential for significant variability in capturing dentition features using casting methods because the preparation of casting mixtures can vary among dentists. There was agreement on the features/parameters of teeth to be measured in bitemark analysis, but participants said these features might change over time. They acknowledged that bitemarks are bruising on a flexible piece of skin (this is discussed further in subsequent sections). When performing bitemark analysis, forensic odontologists don't know what position the jaw was in when the bite occurred. Furthermore, when people bite, the mechanism/dynamic can change from bite to bite. Finally, there are privacy concerns in obtaining the necessary measurement data from dentition.

Limitations in this area include:

- 3D scanning's inability to capture tooth height or tooth imperfections (e.g., chips and cracks)

- No research assesses the ability to accurately convert the 3D image to a 2D overlay to make dentition comparison to a bitemark
- The lack of a large, representative 3D database available to forensic odontologists

Moving forward, participants identified the need for the development of databases of 3D human dentitions representative of the population and suggested attention to statistical considerations in the design of the databases. They also noted that current databases (from Invisalign 3D, iTero, etc.) might be leveraged for information in addition to information obtained from existing casts in dentist's offices. There may be a challenge in securing funding and permission for access to data and studies of this nature. The development of a validated protocol for taking dentition measurements is needed, along with understanding the variability associated with those measurements. Documenting the variability/reliability of 3D scanning devices to establish instrument criteria will be necessary to move bitemark analysis forward. Participants agreed that studies need to be conducted that compare information from casting with *in situ* 3D scanning results. Studies to characterize the differences between 3D scanning information and 2D representations used to compare dentition to bitemark pattern injuries are also needed.

Attendees also discussed that, ideally, statistical interpretation should include three age sets of dentition: child, teen, and adult. The study should include an effort to identify class characteristics, natural vs. post-orthodontic dentition, and racial differences in dentition in a longitudinal study over a period of years sufficient to demonstrate common factors and potential rates of change. Additionally, forensic odontologists should follow the lead of forensic anthropologists in identifying and relating variables of bone structure. Participants also discussed the need to engage other experts/fields (e.g., facial recognition experts) to identify probative variables to look for in the database. Leveraging dental schools may be a way to obtain data from individuals who do not regularly seek dental treatment to create a more representative sample of the population. Participants also mentioned that the Center for Disease Control (CDC) and the National Health and Nutrition Examination Survey (NHANES) might be vehicles to obtain dental scans from a nationally representative population.

4. SCIENCE QUESTION 2: BITEMARKS

Do bitemarks transfer measurable characteristics of the dentition to the substrate?

Claim: Bitemarks in human skin and other substrates reliably reflect the features of dentition.

4.1. Breakout Session

4.1.1. Focus Area A

What imaging and measurement method(s) provide the best information for capturing reliable and reproducible information about the bitemark?

Discussions focused on bitemarks on skin and there was little discussion about bitemarks on other substrates (e.g., food). The observations and comments below all discuss marks on skin.

Imaging currently in use is generally limited to capturing bitemark images and bruising using various wavelengths (e.g., visible, infrared (IR), near IR, ultraviolet (UV)). The consensus was that these methods do not capture all characteristics that might be needed for bitemark analysis. There is also no agreement on the resolution required: increasing image resolution may not

provide significant benefits. Information not readily discerned includes the force of the bite, bruising depth, recognizing skin/mark distortion, victims as unconstrained targets (i.e., subject to movement), or other features of the contusion that cannot be readily visualized. All of these variables can impact the appearance of a bitemark on skin. Participants agreed that other tools to capture these images should be explored. Image enhancement and associated software tools may have the ability to "subtract" hemorrhaging, which may be masking relevant information. Confocal Raman was suggested as a possible method to determine the depth of the bruise. Participants discussed the use of 3D imaging to capture bitemarks, but the consensus was that 3D bruising was rarely encountered, and no one was aware of a demonstrated benefit. Participants agreed bites from other animals are distinguishable from human bitemarks, but other items that may cause contusions can appear to be human bitemarks. There was general agreement that bitemark evidence and associated observations could be a viable investigative tool. This evidence could be particularly important when there are unique features, such as clear differences in the mark and a potential suspect dentition (e.g., dentition has only a few teeth, and the bitemark appears to have a complete set of teeth represented).

Several practitioners make dental casts from the subject's teeth to enable comparison with bitemark contusions. While attendees agreed that casts are an indirect method for representing dentition, practitioners present at the Thinkshop stated that casts are "sufficiently accurate" to represent known dentition and are a valuable tool for the analysis of bitemarks. "Sufficiently accurate" was not defined. Participants suggested that the marks made by upper teeth (drag) and the lower teeth (more distinct) in a bitemark should be considered together to determine if an injury is a human bitemark.

Participants identified the following knowledge gaps:

- No current standard exists that defines the characteristics needed to identify a contusion as a human bitemark, including class characteristics, as well as features that may be individualizing.
- There is no agreed-upon set of features for a single set of teeth that could be used to identify the person responsible for causing the injury.
- There is currently no practical way to create "knowns" that account for the variables associated with how bitemarks are made.
- There are no data about the variation observed in real-world bitemark samples from the same source and, as a result, no way to characterize a single variable's (e.g., force) impact on bitemarks.

Participants agreed that an important task is to decide on a standard set of features to characterize human bitemarks. To do so, studies should include a large number of injuries on skin, some made by humans, some made by animals, and some made by other mechanisms. In addition, each type of injury should be made on different parts of the body, with varying degrees of force, and considering a wide range of positions of the victim relative to the instrument that made the mark on the skin. Even in the most carefully planned study, it would be difficult to account for the element of surprise; a victim's instinctive reaction when attacked is to move away from the source of harm, thereby distorting the injury measurements of the width, depth, and angle of the bitemark impressions. In addition, studies should be performed to identify which imaging and measurement methods are appropriate for capturing reliable and reproducible information about a bitemark.

There was general agreement on the importance of conducting initial studies to explore the foundational validity of forensic bitemark analysis. Is it possible, even under controlled conditions, to determine whether a pattern injury was caused by a human biting down on the skin? Currently, there is no agreed-upon set of features that can be used to accurately determine if a pattern injury is a human bitemark or another type of trauma with characteristics similar to bitemarks. All studies should include inter/intra-rater reliability for a wide range of measurements. There was also discussion about the potential use of AI tools to "learn" those features that accurately characterize bitemarks and potentially facilitate bitemark comparison with dentition. To enable such learning, AI tools would require a significant number of ground-truth knowns (i.e., marks made by known dentition).

4.1.2. Focus Area B

What contributes to the variability in bitemarks from dentition, and how can the variability be determined?

The identified sources of variability in bitemarks from dentition include:

- Amount and direction of the pressure applied to create the bitemark
- Substrate (different areas of the body)
- Age and medical condition of the bitemark victim
- Duration of the bite
- Distortion caused by movement of the body while the bite is inflicted
- Posture or position of the body when bitten
- Medications taken by the victim
- Motion of the biter
- Whether the bite occurred on bare skin or through clothing
- Body temperature of the victim when the bite occurred

There is currently no information regarding the variability in bitemarks from dentition associated with any of the factors mentioned above, such as the amount and direction of the pressure applied to create the bitemark, distortion caused by movement of the body while the bite is inflicted, presence of hair or lotion on the area bitten, duration of the bite, and the posture or position of the body when bitten. There was some agreement that it may be impossible to find ways to measure all these variables under controlled conditions. To further complicate matters, the impact of these factors may vary because of variation in the human victim healing response. Given clear risks to participants, it was unclear whether it would be possible to get such studies approved by IRB. Even if approved, there was a consensus that it would be challenging to create a real-world scenario because of the variability in response of the simulated victim to the bitemark. The person being bitten would know the bite is coming. They also may be anesthetized. This is different from real-world bitemark scenarios where bitemark victims move, react, do not expect the bite, or are in a fight or flight mindset. The reasons listed here may account for the lack of research in this area.

The path forward to understanding variability in bitemarks and how that variability can be determined is difficult to ascertain. Certainly, if possible, studies should be performed to duplicate human-biting dynamics and collect data on live victims.

4.1.3. Focus Area C

What data collection techniques are sufficient to collect evidence of pattern injuries on human skin?

In addition to the collection of visual evidence from the apparent bitemark on human skin, other available appropriate evidence should also be collected. This may include trace evidence, such as hairs, fibers, etc., which should be carefully collected first. In addition, the area of the bitemark should be swabbed for possible DNA analysis. Trained and qualified personnel in each jurisdiction should collect this evidence.

As for the bitemark itself, participants discussed various photographic and lighting methods used to capture as much information as possible. At a minimum, the bitemark should be photographed in color and then with alternate light imaging methods, including UV and IR light. Distant and close-up photos should be taken, with scale. A trained evidence photographer should take the photos, but, practically, photos can be taken with a phone camera if necessary. Photos should be taken as soon as possible after the bitemark was inflicted, and the time elapsed between injury and recording should be noted. When photos are taken of a live victim, analysts should be aware that the bitemark can change over time. Therefore, additional photos should be taken of the bitemark approximately seven and twenty-one days after the injury was inflicted and continue until the images are no longer providing additional information. Participants noted that UV photos may show more distinct characteristics at later photographic intervals. However, scientific articles on the use of UV in bitemark analysis are not common. In addition, participants mentioned that hyperspectral imaging and Raman spectrometry may be useful tools in bitemark imaging. In addition to photography, experts should take a tissue impression or a 3D scan of the bitemark. On deceased individuals with pattern injuries, more invasive techniques have been employed. Techniques include tissue excision, microscopy, trans-illumination, and histological examination of the tissue. Participants emphasized that tissue excision of pattern injuries on skin have been proven to impart distortion on the bitemark, and experts do not generally use tissue excision. They also noted that forensic dentists can attempt to take measurements of the injury on the skin to compare with the putative dentition, but it is unknown whether the skin reliably captures the actual attributes of the teeth.

Forensic dentists have not established whether what they do works. The subjective nature of bitemark analysis has been an issue since 1974 and remains so. Subjectivity does not in itself invalidate the methods, but it does mean that careful studies on measurements and bitemark decisions are required. Obtaining more information is necessary before it can be established that forensic bitemark analysis is valid and reliable.

The knowledge gaps identified regarding data collection techniques include:

- A lack of reliability of measurements on skin that exhibit a pattern injury
- A lack of criteria for determining the sources of pattern injuries
- A lack of data to understand the distortion of bitemarks inflicted on skin
- A lack of models of injuries on skin that consider bite dynamics
- The methods currently employed for bitemark analysis have not been systematically tested
- There are no universal training standards, detailed protocols, or workflows for bitemark analysis

- Current standards for collection of bitemark evidence are not specific (e.g., need distance, light, etc.)
- A lack of large ground-truth studies of tooth marks on skin with victims in motion, animal bites, etc.

The next steps include those associated with measurement of injury patterns and interpretation of those measurements. Steps should be taken to help determine if researchers can establish reliable and reproducible metrics for differentiating between a bitemark and another pattern injury. Participants suggested that movement forward should be minimal in scope so that the basic issues mentioned above can be thoroughly addressed: whether it is possible to distinguish between different types of pattern injury (bitemark vs. non-bitemark). The first question to be addressed should be, "Is it a bitemark or another pattern injury to skin?" Participants mentioned that current facial recognition technology can identify a specific face in a large number of faces. Is it possible that this technology can be somehow applied to assist with bitemarks? In addition, they suggested that AI may be able to assist in recognition of a bitemark vs. non-bitemark. However, machine learning or artificial intelligence techniques can be used only if large data sets are available. Another key interpretation issue is where bitemark analysis results in the "non-exclusion" of a suspect. There are currently no means to quantify the rarity of that bitemark association. Participants agreed that standards must be developed for data collection and interpretation.

4.2. Moderated Discussion

This moderated session included discussion about the lack of information about types of bitemarks and their causes. Participants agreed there is variability between bitemarks made by the same set of teeth. Variables affecting bitemarks include the location of the bite on the body, the angle of bite, and the force of the bite. There were disagreements regarding the type of conclusions that can be reached about the age of the person who inflicted a bitemark (very different in living vs. dead victims). There was, however, general agreement that the sheer number of variables that affect the attributes of a pattern injury on skin may be too large for a truly controlled experiment. Participants agreed that it may not matter how well injuries are imaged because the injury may be distorted, obscuring the relationship to the actual teeth that may have caused the injury. If something *can* distort, that doesn't mean it *always* will distort, and experts should not assume that a picture is 100% accurate.

It was generally agreed that the some of the best studies that have been conducted on human bitemark identification have been conducted by Mary A. Bush, DDS. The Bush studies[1][2][3] revealed the distortion that can be created by biting. However, it is unknown if primary distortion is distortion at all, or it may be the most accurate representation of the bite. Attendees concluded that there is no perfect model but that, even though limited, the Bush model is the most reliable.

Participants raised concerns that if forensic dentists cannot reliably reproduce a human bitemark, reaching conclusions about the person who produced the injury is problematic. The

[1] M.A. Bush, et al., "Biomechanical Factors in Human Dermal Bitemarks in a Cadaver Model", Journal of Forensic Science, January2009
[2] M.A. Bush, et al., "Inquiry into the Scientific Basis for Bitemark Profiling and Arbitrary Distortion Compensation", Journal of Forensic Science, July 2010
[3] M.A. Bush, et al., "A Study of Multiple Bitemarks Inflicted in Human Skin by a Single Dentition Using Geometric Morphometric Analysis", Forensic Science International, September 2011

objective, repeatable, and scientific identification of a bitemark is not substantiated, and even known bites can produce different results. If we cannot answer these questions, how can we exclude or conclusively determine if an individual inflicted a bitemark? The Texas Forensic Science Commission has issued a moratorium on judges admitting bitemark analysis, and, without empirical data, forensic odontologists cannot continue to testify about bitemark evidence in Texas cases. Participants discussed that forensic odontologists should be courageous enough to say, "we don't know." They should also do the work first to gain confidence in performing these examinations. Failure to gain such confidence in forensic odontology examinations risk putting innocent people in prison or letting guilty people walk free.

Participants agreed that there are significant knowledge gaps associated with bitemarks and the transfer of measurable characteristics of dentition onto human skin. The limitations discussed include:

- Not understanding how bitemarks/contusions change over time and the associated impact on making bitemark comparisons
- Not having useful/relevant information for comparing bitemarks to dentition
- Not knowing what the scope of the damage done by the "impact tool" (i.e., dentition) is on the bitemark and how it influences the ability to characterize the bitemark and compare it to possible dentition sources
- Not understanding how the extent of the contusion on the skin reflects the dentition characteristics
- Lack of understanding what features, including depth of injury, in a bitemark are characteristic of a human bitemark
- Lack of understanding of what characteristics/observations of a bitemark may make that mark unique
- Not understanding the extent to which skin conditions may impact bitemark contusions
- Not knowing how to recognize distortion and understanding its impact on the ability to characterize and potentially compare bitemarks to dentition
- Realizing that it may not be possible to experimentally study actual bitemark dynamics due to IRB approvals and difficulties in finding volunteers
- The problem that bitemark studies, even if allowed, are not necessarily probative because bitemark study participants are aware they are being bitten and therefore will react differently to being bitten than individuals who are not aware of it
- Accounting for the fact that skin is an unreliable impression medium, and the consequently low reliability of pattern injury transfer to skin
- Lack of data for studies due to the limited availability of advanced equipment (i.e., scanners, camera)
- Lack of standardized training and protocols
- Lack of data on known sources of pattern injuries

Challenges to overcome in this area include funding for studies and equipment, examination of current protocols compared to current scientific literature, standardized education and training for forensic odontologists, standardized workflows for data collection, identification of metrics to interpret pattern injuries (e.g., human bitemarks), threshold requirements for technology to collect pattern injury data, ground truth studies needed to establish reliability and repeatability, and possible use of Raman to examine and characterize subsurface injuries.

Participants identified steps to move forward. These include experimentation (identify nine to twelve variables to limit effort), a historical case review study, determination of what variables interact to help design experiments, and the design of better biting devices for experiments on living people. Future studies should include the following topics:

- Mechanisms of biting by the human jaw
- Determination of variables in bitemark examinations and the relative contribution of each variable
- Development of a ground truth database to assess contusion variability
- Determination of what class and individualizing characteristics determine a human bitemark
- Determination of what may be individualizing features in a bitemark and how they relate to dentition

A study should be done to assess the viability of using upper teeth drag marks and lower teeth features displayed in a bitemark to determine if the mark is indicative of a human bitemark. Also, to be explored is the potential use of image processing tools (e.g., filtering and other image processing techniques) to find additional meaningful information from bitemarks. Studies should include an inter-rater and intra-rater reliability of measurements and pattern injuries inflicted by objects other than bitemarks. The focus is reliability, and to understand measurement consistency over time, between items, and among different practitioners. Pathologists and emergency medical personnel are experienced in the examination of pattern injuries and could aid in differentiating between accidental injuries and injuries caused by a violent act. Forensic odontologists should be open to embracing new ideas and different perspectives when designing and conducting these studies. Ultimately, the science needs to move forward before forensic odontologists can make decisions based upon reliable and rational data.

5. SCIENCE QUESTION 3: ANALYSIS AND INTERPRETATION

What interpretation strategies (techniques and practices) produce the most accurate and reliable results?

Claim: Selected data interpretation strategies produce more reliable/defensible results.

5.1. Breakout Session

5.1.1. Focus Area A

What defines sufficiency to establish reliability in the association of bitemarks to dentition?

Participants generally agreed that there is a need for more data to assess the degree to which one can reliably associate bitemarks to dentition. There is no established minimum number of features defined to 1) identify a bitemark or 2) exclude/include a suspect dentition to a bitemark. They also noted that bitemarks are highly variable and are often limited in quantity and quality of features.

Attendees discussed many gaps. The topics included:

- An understanding of class and individual characteristics of dentition

- A minimum threshold for the quality and quantity of information to identify a pattern injury as a human bitemark
- Lack of knowledge regarding the impact of distortion occurring in the bitemark process
- Limited ground truth data in this area, including rarity of dental features in a population
- Variation in bitemarks inflicted by the same individual
- How to correlate a bitemark to dentition

To address the gaps to establish reliability in the association of bitemarks to dentition, there should be studies to establish definitions of class vs. individual characteristics of dentition, establish error rates for bitemark exams (all aspects of ABFO comparison), establish quality and quantity of characteristics needed to define a bitemark, and evaluate pattern matching software for possible use in bitemark comparisons. In addition, participants suggested that bitemark testimonies be reviewed and parameters for association testimony prescribed. They also noted that a study of jurors involved in bitemark cases be used to determine the juror perception of a "cannot exclude" bitemark analysis result.

5.1.2. Focus Area B

What other data are relevant to bitemark examination and analysis?

To minimize bias when performing examination, forensic odontologists should not have any information except what is essential to perform their analysis. Of course, separate and apart from swabs of the bitemark taken for DNA analysis, the most relevant data needed to perform bitemark examinations includes high quality, well-documented photos of the bitemark, documentation of the position on the body and where on the body the bitemark occurred, 3D imaging of the bitemark, impressions of the bitemark, approximate time of death (if appropriate), and if available, information about the time between when the bitemark was inflicted and the information was obtained. It could be useful to investigate using non-traditional spectroscopic technology such as Raman or hyperspectral imaging to obtain additional information during bitemark analysis.

Gaps in relevant data for bitemark analysis include:

- A lack of studies comparing techniques to capture relevant bitemark information
- A lack of ground truth studies to validate the current approach
- A lack of definition of minimum standards to identify a bitemark as such
- Ground truth of meta-data concerning the uniqueness of dentition

The path toward identifying data relevant to bitemark analysis includes funding and resources for studies of Raman and other non-traditional spectroscopic techniques applied to bitemarks, studies to determine minimal characteristics required to identify a human bitemark, studies of 3D scanning of bitemarks, and studies to validate the "dental line-up" approach. In addition, studies should be conducted to establish the error rates for bitemark analysis and juror studies to shed light on the significance of bitemark analysis conclusions.

5.1.3. Focus Area C

What are the key approaches to take in bitemark analysis that will ensure the comparison is objective and, if the dentition is not excluded, the significance of an association is accurately reported?

Initially, attendees agreed that examiners should determine the quantity and quality of the data in the Questioned sample (Q) before looking at the Known (K) sample to decide if there are sufficient data present in the Q for comparison to the K sample. Baseline considerations of the quality of the Q data should include scale and adequate photo documentation of the bitemark, but there is no standardization for the baseline considerations. Often "dental lineups" are incorporated into the comparison process: experts compare five total dentitions (including the suspect's dentition) to apparent bitemark injuries. However, participants noted that there is no standardization for selection of the dentition included in the "line-up." There is no scientific basis for including a total of five samples in the "dental line-up," and it is often difficult to find dentition similar to the Q to use in the line-up. Current ABFO guidelines suggest that bitemark exams include an independent reviewer in the examination process. However, attendees noted that the presence of the independent reviewer has not been demonstrated to improve the accuracy of bitemark examinations. In addition, participants agreed that the independent review should be more of a technical review for quality assurance purposes. Forensic odontologists must disclose what they can and cannot determine in a bitemark exam in their reports and testimonies. Some attendees proposed that if forensic odontologists cannot say that the suspect's dentition and the bitemark are unique, no comparison should be made. They also noted that criminal prosecutors may try to overemphasize the "cannot exclude" conclusion. Participants emphasized that bitemark excision in deceased victims distorts the tissue, and forensic dentists should not use this method for bitemark pattern analysis purposes. When experts use overlays in the bitemark comparison process, the overlays must be consistent in size. Participants discussed that bitemark analysis is a qualitative comparison that is subjective by nature, and the forensic dentists should not use the term "match" in reports or testimony.

The many knowledge gaps in establishing the key approaches for bitemark analysis include:

- A need for determining and validating the data needed to determine if an injury is a bitemark
- A lack of a minimum threshold of data to compare an apparent bitemark to a known sample
- A need for determining the accuracy of bitemark analysis
- A lack of gold-standard data in the bitemark examination process
- The introduction of bias when exposed to contextual information surrounding the criminal case
- Lack of information on the errors made in bitemark cases

Participants identified studies regarding the subjective decisions that constitute a bitemark analysis as the next step needed to move the field forward. There is a lack of data on the similarity of dentition and no error rate established for bitemark comparisons. Scientific studies must be performed to establish error rate data for bitemark examinations, and funding for bitemark research is very limited. Research should also include studies with judges who have presided over bitemark cases to assess their understanding of the practice. Bitemark cases that first resulted in convictions followed by exonerations should be examined. We have not done studies to determine where the examination process failed. We have not determined if errors were related to the process or if human factors such as bias contributed. A transparent study of bitemark examination case errors is needed. The ABFO has recommendations/guidelines/best practices but no standards for the bitemark examination process. Well defined bitemark examination standards, including detailed process maps and complex decision trees, are needed.

Root cause analysis for cases with errors has been requested but has not been carried out. Challenges include obtaining funding and resources to establish, at minimum, a threshold for data to make a bitemark comparison.

5.2. Moderated Discussion

Participants agreed that bitemark analysis is a subjective examination. They also agreed that bitemark photos should include different lighting techniques and scales. To minimize bias, forensic dentists have employed the "dental lineup" in the comparison process. However, there has been no validation of the dental lineup process. Attendees discussed that contextual information surrounding the crime may introduce bias into the comparison process. They noted that peer review/independent verification is used to try to produce more accurate and reliable results but agreed there is a lack of data available about the use of peer review and its impacts. In addition, independent analysis of the questioned evidence for quality, quantity, and significance of data helps produce the most accurate and reliable results.

Limitations in this area include:

- A lack of standard protocols for bitemark examinations
- A lack of sufficient experimental data to establish "best" practices
- No universal agreement on class and individual characteristics
- A lack of the minimum number of criteria/features (quality and quantity of data) needed to identify a pattern as a bitemark
- A lack of data to determine the minimum quantity and quality of data used in the bitemark analysis process
- A lack of knowledge regarding the range of variability in human dentition and how orthodontic procedures affect this range
- A lack of research on best practices for the "dental lineup" (no other forensic discipline utilizes a lineup of this nature)
- A lack of knowledge of spectral interpretation from data created by non-traditional spectral techniques (Raman and hyper-spectral techniques)
- A lack of studies concerning 3D scanning on bitemarks
- A lack of knowledge on the effects of body position and movement of the bitemark victim
- A lack of data concerning time of death vs. first observation of wound pattern

Participants also observed that there is a lack of studies to establish data consistency, a lack of error rates (bitemark vs. non-bitemark, human vs. non-human, etc.), and a lack of root cause analysis on cases of error. It was also noted that there were no judge and juror studies regarding bitemark testimony (what does "cannot exclude" mean to them).

Steps to move forward in this area include the need for significant funding and resources for studies to address all gaps listed above. It is well understood that generating ground-truth bitemark samples to evaluate the ability to match dentition to bitemarks accurately is very problematic, if not currently impossible. One suggested approach was to generate bitemark samples in silicone and vary features to define the limits of comparison. If sufficient ground truth samples are available, AI techniques could be used to learn what is needed to make effective comparisons. Pattern matching software and data analysis tools (e.g., PCA) should be evaluated

to determine their applicability to bitemarks and comparison to dentition. In addition, novel imaging techniques should be studied and compared.

Studies should be conducted to establish the minimum acceptable standards associated with bitemark analysis. Studies should also include assessing the ABFO guidelines to determine how widely they are being used and whether they are effective at making bitemark comparisons more reliable. Work is also needed to establish limits on report statements and expert testimony about bitemarks. There was no agreement about the effectiveness of the use of the "dental line-up." Additionally, in the absence of error rates, bitemark reports with a "cannot exclude" conclusion should state that it is unknown if the bite could have come from another individual. Error rate studies on measurements and comparisons should be conducted. During the discussion, it was agreed that experts should not derive racial or gender information from a bitemark. Participants also discussed that the age of the biter and the victim may be important. Bites from a child are not a criminal act. Bitemarks found on a child may heal faster than those on an adult. There was also discussion about the qualifications of forensic odontologists and a need for minimum standards of training.

6. MODERATED PANEL DISCUSSION – CONCLUSIONS AND TAKEAWAYS

The final discussion focused on the pitfalls of bitemark analysis and how we can come together as a community to assess the scientific foundation for the field and, if justified, increase confidence in it. Participants noted that the Thinkshop, bringing both sides of the aisle together, was an accomplishment in itself, and one that needs to continue. Participants emphasized the urgency of the situation: the lives of innocent people may be on the line. They posed the following questions regarding next steps:

- Should different thresholds be established for investigative lead value vs. information that could be provided in court?
- Is there a scientific basis for such thresholds to be established?
- Can there be a minimum threshold for determining
    - Injury to skin?
    - Pattern injury?
    - Human bitemark vs. non-bitemark?
- If so, what are these thresholds?
- Can we develop
    - Common terminology?
    - Minimum characteristics for bitemark analysis to be performed?
    - Minimum standards for interpreting findings?
    - Transparent and consistent reporting conclusions?
- Is research feasible/possible?
- What are the research priorities?

There was much discussion among the participants about the logical next steps for NIST and others to take to accomplish these tasks. Participants concluded it is necessary to prioritize fundamental research studies. Throughout the two days, the need for reliability studies was consistently brought up as a way to assess consistency and to ensure we can accurately measure

bitemarks. The group ultimately agreed the first fundamental research study for NIST to conduct to define the scientific foundations of bitemark analysis should focus on the question:

"Can it be determined if a pattern or pattern injury is a human bitemark?"

Prior to conducting this study, a detailed plan should be developed and agreed upon. It should include information on what is needed to perform this study, including sample size, sample types, outcome measures, etc.

Appendix A
BITEMARK THINKSHOP AGENDA

Bitemark

Thinkshop

October 17-18,

2019

UVA Darden Sands Family Grounds – UVA Darden
DC Metro

1100 Wilson Blvd, Arlington, VA 22209, on the 30th and 31st floors of the Monday Properties

(Please use elevators near the security desk in the building lobby to access)

Contact Information: Harlie Jud, harliej@iastate.edu (O: 515 294 7278, C: 515 351 9407)

	Thursday, October 17	
7:00 AM	Breakfast Buffet	Flat Classroom 313- 31st Floor
8:00 AM	"Charge to this Thinkshop Group – Why are we here?" - Dr. Richard Cavanagh	Large Tiered Classroom – 31st Floor
8:15 AM	"Bitemarks 101" – Dr. David Senn and Dr. Robert Dorion	Large Tiered Classroom – 31st Floor
9:00 AM	"Mechanisms of Injury with Application to Pattern Injuries" – Greg Davis, Medical Examiner	Large Tiered Classroom – 31st Floor
9:45 AM	Break	
Refreshments available outside of room from 7:30 – 11:30 AM		
10:00 AM	"Literature Review Output" – Dr. John Morgan	Large Tiered Classroom – 31st Floor
10:45 AM	"Bitemarks from a Standard's Perspective" – Dr. Ken Aschheim	Large Tiered Classroom – 31st Floor
11:30 AM	Breakout Topics Overview	Large Tiered Classroom – 31st Floor
12:00 PM	Lunch Buffet	Flat Classroom 313 – 31st Floor
1:00 PM	Breakout Sessions – Topic 1, Focus areas A, B & C	Refer to individual schedule for room assignment
3:00 PM	Break	

Refreshments available outside of room from 12:30 to 4:30 PM		
3:15 – 5:15 PM	Breakout Sessions – Topic 2, Focus areas A, B & C	Refer to individual schedule for room assignment

	Friday, October 18	
7:00 AM	Breakfast Buffet	Flat Classroom 313 – 31st Floor
8:00 AM	Moderated Discussion of Topic 1	Large Tiered Classroom – 31st Floor
9:00 AM	Moderated Discussion of Topic 2	Large Tiered Classroom – 31st Floor
10:00 AM	Break	
Refreshments available outside of room from 7:30 – 11:30 AM		
10:15 AM	Breakout Session – Topic 3, Focus areas A, B & C	Refer to individual schedule for room assignment
12:15 PM	Lunch Buffet	Flat Classroom 313 – 31st Floor
1:45 PM	Moderated Discussion of Topic 3	Large Tiered Classroom – 31st Floor
Refreshments available outside of room from 12:30 – 4:00 PM		
2:45 – 4:00 PM	Moderated Panel Discussion – Conclusions and Takeaways	Large Tiered Classroom – 31st Floor

Appendix B
HANDOUT ON SCIENCE QUESTIONS

Bitemark Thinkshop Breakout Discussions

October 17-18, 2019

Breakout for Science Question 1

Understanding Dentition (Teeth)

Claim

Characteristics of human dentition are unique or can be divided into reliable fractions of the population provided consideration of any changes with morphometric parameters over time and events.

Science Question

Are there measurable characteristics or features in human dentition that vary among individuals and are persistent within an individual?

- ☐ Do we fully understand the discriminating power of the features of interest within human dentition?
- ☐ Are these features of interest well characterized?
  - Do we know what we are measuring?
  - Are these features meaningful to measure?
- ☐ How unique are human dental characteristics in an absolute sense (dentition itself)?
- ☐ How much do the morphometric parameters change over time due to various events?

Focus Areas Facets

Focus Area A

Concentration: issues with imaging (casts vs. 3D), accuracy, number of replicas, resolution of image

What measurement method(s) provide the best information for capturing reliable information about the detention?

- Imaging technology and requirements
  - Casts
    - Hard
    - Soft
  - 3-D digital scanning
  - Direct impressions on test material/media

- Determination of theoretical limitation of measuring biometric parameters of the dentition (e.g. image resolution, etc.)

Focus Area B

Concentration: databases and collection analyses

How do we appropriately collect information to create population databases that can be used for scientific and statistical analysis of human dentition?

- Development of 3D databases of human dentitions (i.e., cadaver dental impressions to avoid Institutional Review Board (IRB); other?)
- Large population studies including statistical consideration of the minimum number of previously defined biometric qualitative and quantitative parameters or discrimination of the dentition.
  Will the database be reflective of the population over time?

Focus Area C

Concentration: biometric and temporal variables, what are relevant features, teeth in relation to each other, is there consistency over time

What are the most probative features/parameters to use and what are the limits associated with each?

- Biometric qualitative and quantitative parameters for describing and measuring human dentition
- Determination of the uniqueness of dentition based on these biometric parameters
- Biometric parameters of tooth and arch form for determination of primary/mixed/adult bite dentitions
- How do we take into account changes in morphometric parameters over time due to various events?

Bitemark Thinkshop Breakout Discussions

October 17-18, 2019

Breakout for Science Question 2

Understanding Bitemarks

Claim

Bitemarks in human skin and other substrates reliably reflect the features of dentition.

Science Question #2

Do bitemarks transfer measurable characteristics of the dentition to the substrate?

- Do we understand the stability, transferability, and/or persistence of the discriminating features used in bitemark interpretation? (Do we understand the meaning and reliability our measurements?)
- How consistent/variable are bitemarks made by the same dentition?
- What resolution is required for meaningful analysis?
- What issues impact the above questions for human skin?
    - Need to address specifics around temporal variability (a body's response to an injury, temporal issues – biological, physiological aspects.)
- What issues impact the above questions for other substrates?

Focus Area Facets

Focus Area A

Concentration: imaging analysis (various substrates)

What imaging and measurement methods(s) provide the best information for capturing reliable and reproducible information about the bitemark?

- Fidelity of 2D imaging of bitemarks (various substrates)
- Technologies for 3D imaging of bitemarks (various substrates)
- Are there other techniques to measure and characterize bitemarks that may be informative?
- Determination of theoretical limitations of measuring biometric parameters of the bitemark (e.g., image resolution, etc.)
- Biometric qualitative and quantitative parameters for describing and measuring a bitemark
- Determination of the uniqueness of a bitemark based on these biometric parameters

Focus Area B

Concentration: variations and different substrates

What contributes to the variability in bitemarks from the same dentition and how can the variability be determined?

- Variability of multiple bitemarks from the same known source
- Effects of biting mechanics variables such as force, duration, angle, movement, etc.
- Study of temporal variation of bitemarks over time

Focus Area C

Concentration: looking at skin as a substrate

What data collection techniques are sufficient to collect evidence of pattern injuries on human skin?
Specific issues associated with imaging of patterned injuries on human skin (2D and 3D)

- Development of reliable skin model for imaging/feature studies: range of Young's modulus; linear imaging material
- Biometric qualitative and quantitative parameters for describing and measuring a bitemark on human skin
- Consideration of skin variables such as topography, underlying structures, skin type (old, young, fat, slender)
- How to measure characteristics to identify the pattern injury/injury source, (i.e., is it a human bitemark?) and capture sufficient information in detail that would be representative of the injury
- How much information is sufficient to draw a conclusion about the dentition?

Bitemark Thinkshop Breakout Discussions

October 17-18, 2019

Breakout for Science Question 3

Data Interpretation Strategies

Claim

Selected data interpretation strategies produce more reliable/defensible results.

Science Question #3

What interpretation strategies (techniques and practices) produce the most accurate and reliable results?

- Can we establish the reliability of the methods involved in identifying and characterizing the discriminating features used for associating an evidentiary bitemark with a questioned individual's dentition?
- Can we demonstrate the combined reliability of our measurements and interpretation strategies?
- What type of conclusions can be drawn based on the reliability of this type of evidence?
- What is necessary to establish accuracy, reliability, and defensibility?
- What are the best features to consider?
- What are the best techniques and practices?
- What are keys to proper data interpretation?

Focus Area Facets

Focus Area A

Concentration: number of samples and distinct comparisons required

What defines sufficiency to establish reliability in the association of bitemarks to dentition?

- Minimum number of dentition samples required for comparison/association with a bitemark
- Minimum number and type of distinct features of dentition and of the bitemark required for association
- How do we assess the role of distortion and observed non-matching features?
- What is the role, if any, for digital/state-of-the-art imaging plus pattern recognition software?

Focus Area B

Concentration: metadata, subpopulation questions, making comparisons within certain classes

What other data is relevant to bitemark examination and analyses?

- Inclusion of metadata such as DNA, other biomarkers, age/ethnicity/gender from the subject and victim / evidence
- Comparison of association strategies

Focus Area C

Concentration: statistical analysis, strategies to overcome human factors

What are the key approaches to take in bitemark analysis that will ensure the comparison is objective and if the dentition is not excluded, the significance of a match is accurately reported?

- Relative roles of statistical analysis
- Are there lessons learned from other disciplines that can be applied in bitemark analysis (e.g., two-stage methods, white box studies, likelihood ratios, machine learning)?
- Human factors study
- What are the range of conclusions that should be issued in reports?

Appendix C
RAW NOTES

NIST CSAFE Bitemark Thinkshop

October 17th – 18th, 2019

Presentation, Breakout Sessions, and Moderation Discussion Notes

It should be noted that there is repetition in the notes due to the Thinkshop format where, in an effort to be thorough and transparent, Groups A, B, and C were all tasked with discussing the same topics. All the Thinkshop participants were highly engaged and enthusiastic about expressing their viewpoints and the Thinkshop notes represent the note recorders best efforts to fairly reflect the ideas expressed by all Thinkshop participants.

Table of Contents

1. Opening Talks

1.1 Richard Cavanaugh -Bitemark Thinkshop: Why Are We Here?

Thursday, October 17th, 8:00 am

NIST Efforts since 2009

- NRC (National Research Council) 2009 report
- Since 2009: NCFS (National Commission on Forensic Science), OSAC (Organization of Scientific Area Committees for Forensic Science), NIST Forensic Science Research Program, CSAFE (Center for Statistic and Applications in Forensic Evidence), Scientific Foundation Reviews
- SFR (Scientific Foundation Reviews) = technical merit evaluation
    - NIST agreed to undertake a series of experiments:
        - DNA mixtures: extensive foundation science in DNA
        - Bitemarks: limited foundation science in bitemarks
        - Toolmarks and firearms – a lot is happening in objective comparisons
        - Digital – a fast moving field
    - Purpose: To document and consolidate information supporting (or opposing) the methods used in forensic analysis and to identify gaps

Science

- The science of reliability – a quality system (control chart)
- Science as a research enterprise (discovery) – ask a question, do background research
    - https://www.sciencebuddies.org/science-fair-projects/science-fair/steps-of-the-scientific-method

Steering Committee

- Robert Barsley – passed away last month; Sept 28, 2019

Thinkshop

- Why a Thinkshop instead of a workshop?
    - Thinkshop: usually a brief intensive exploratory program that focuses on open challenges and knowledge gaps
- Asking the right questions is as important as answering them (Benoit Mandelbrot)
- Bitemark analysis rests on two additional postulates:
    - First, that the anatomical configuration of every mouth is unique
    - Second, that this asserted uniqueness is graphically displayed in the bite mark
- Questions
    - Do we fully understand the discrimination power of the features of interest within biting teeth?
    - Do we understand the stability, transferability, and/or persistence of the discriminating features used in bite mark interpretation?
    - Can we establish the reliability of the methods involved in identifying and characterizing the discriminating features used for associating an evidentiary bitemark with a questioned individual's dentition?

Mix of perspectives

- Take a diversity of perspectives: object looks like square, circle, or triangle
- Very familiar with forensic odontology
    - Odontologists
- Not so familiar with forensic odontology
    - Researchers
    - Statisticians
    - Legal experts
    - Image analysts

- All perspectives are insightful, creative, and open thinkers

1.2 Robert Dorion – Bitemarks 101

Thursday, October 17th, 8:15 am

Wrongful Convictions - Bitemark

- 800 cases in CA alone over 13 years
    - 14 systemic issues
- Since 1996, less wrongful convictions overall
- 2007 - 15 institutions/universities with schooling
- Human bitemark/not human bitemark/inconclusive
    - Excluded/not excluded/inconclusive
- Similar to FR, can we create an algorithm to identify bitemarks from the same individual? American Board of Forensic Odontology (ABFO, started in 1976)
- CASE: Ray Krone, 2002
- AAFS 2008 – four talks given
    - Mary Bush – biomechanical properties of skin, distortions in bitemarks
    - Peter Bush
    - Fabric impression on skin from clothing
- JFI 2009, 59(3)
- NRC 2009 concerns
    - Uniqueness of the bitemarks
    - Transfer
    - Persistence
- NIJ bitemark proposals were not funded
    - No bitemark proposal has been funded since 2009
- Bitemark photography – protocol not well defined
- PCAST 2016 – we advise against devoting significant resources to such efforts
    - 406 bitemark references
- 26 publications, 12 AAFS talks were submitted by Robert Dorion to PCAST for their addendum but were not included
- Bruce Budowle – comments on PCAST were cited
- Uniqueness of the dentition: Sognnaes et al. 1982 … Franco et al. 2017
- Factors affecting bitemarks: intrinsic factors, extrinsic factors, environmental exposure
- Why postmortem bitemarks have limited experimental value – 11 reasons given
    - Shape retention - Rebounded tissue examined
    - Is pattern analysis a science at all? (Jim Lewis talk)

W.R. Oliver 2016 J Forensic Sci. 61(4), 62(3), …

- Freeman et al. 2005 JFS
- How many teeth to call it a human bitemark?
- What if victim claims to have been bitten? Will this bias the examiner?
- Methods
    - DNA collection
    - Hair removal
    - Color photography
    - UV photography
    - Alternative light imaging photography
    - IR photography
    - Impression
    - Scanning
    - Macroscopy
    - Microscopy
- 3D tissue excision on the dead – histology of the tissue can reveal timing of the death relative to the bite however, tissue excision will distort tissue/bitemark
- Improper use of ABFO standards
- Media has never reported a non-U.S. exoneration of bitemarks
- Prosecutorial misconduct
- Am J Forensic Med Pathol 2018 – 14 factors involving wrongful convictions
- 311 bitemarks (274 antemortem, 37 postmortem) in 2019 largest bitemark research database

1.3 David Senn – Bitemarks 101

Thursday, October 17th, 8:15 am

IRB study - 15 participants bit with a cast

- Challenge of getting IRB approval on living human skin
- Showed video

Case

- 29-month-old bit multiple times
- Primary photographs at hospital
    - Dental taken later by forensic dentists
- Arrange photo anatomically
- 4 suspects
    - Mom
    - Babysitter's boyfriend
    - Babysitter
    - Babysitter's 5-year-old
- Photos of teeth from suspects taken
- Universal tooth numbering system used
- Persons of interest in case: mother, babysitter, babysitter's boyfriend, babysitter's 5-year-old son
- Evidence: dental photographs of each POI

- Digital overlays created
- If the distortion is great, then a comparison is not made
- It was determined that the bite mark was made by a child due to class characteristics
- It is not a crime for a child to bite another child – but it is a crime for an adult to bite a child

1.4 Gregory Davis – Mechanisms of Skin Injury with Application to Pattern Injuries

Thursday, October 17th, 9:00 am

- Weapon=teeth
- Skin is elastic
- Can tell a lot of the mechanism but not the exact item per se
    - Abrasion
        - Injury to the surface of the skin
        - Damages epidermis at the least
            - Tangential force with scraping
            - Perpendicular force with crushing
    - Contusion
        - Injury to blood vessels deep to skin
    - Laceration
        - Rupture of skin that cannot withstand force
        - Occurs along weak line, like tearing paper
        - Does not produce useful pattern of weapon
    - Fracture
        - Broken bone
    - Incision
        - Caused by sharp edge (incisors)
    - Chop wound
        - Special category of wound by sharp edge
- From looking at the wound itself: This is a knife wound
    - But need other tests to determine exact
- Patterns
    - Example of a zipper wound on a car accident victim coming from her clothing
    - Other people ought to be able to see the patterns and how they might fit
- *Resist the temptation to say more than you can say*
    - Attorneys have a job to do and that is to advocate for their client and client's position
    - Experts have a job to do, hold the line for appropriate interpretation of injury
    - Any injury you show to a jury should be understandable to the jury with some explanation. If you have to say "You'll just have to trust me on this" to every juror, then you have gone too far in your interpretation.
- Bitemark analysis must be measurable and reproducible
- We want to stick with science and not voodoo – we should be able to show it to someone and them to understand

Questions

- Skin rebounding of tissue – your body tries to heal an injury, which leads to inflammation; measurements on the order of a millimeter is pushing it due to skin changing over time
- Differences in skin – types of skin, diabetic, thickness of skin, differences in individuals?

1.5 John Morgan – Scientific Foundations of Bitemark Examinations

Thursday, October 17th, 10:00 am

Scientific foundations of bitemark examinations

- Scientific Foundations Overview – methodology and online assessment database
  - Definition of the assumptions underlying a technique
  - Collection of all primary research related to the assumptions
  - Double reviews of all research papers
  - Determination of the consensus of reviews and research basis for the assumptions
- We are not establishing the acceptability of the technique as evidence in legal proceedings

Bitemark Examination

- Bitemark analysts make nine assumptions concerning their ability to associate a mark on human skin to an individual's dentition
  - Bitemarks on other media were not examined
  - Other bruising mechanisms were not examined
  - Non-primary research was excluded, including reviews of bitemark examination in the legal and scientific literature
  - Case studies were included if the paper related to one of the assumptions of bitemark examination
    - In general, case studies provide weak support for any supposition

1. Bitemarks remain on skin and are persistent to a sufficient degree to permit examination within a time relevant to police investigation.
2. Bitemark patterns may be distinguished from other skin patterns not associated with human bites.
3. Bitemark pattern distortions are known and can be deconvolved from bitemark patterns. Distortions may be due to the limitations of skin as a transfer medium, healing, postmortem changes, variations in bite force and angle, other skin marks, variations related to placement on the body/head/extremities, and other variables.
4. It is possible to collect and analyze skin patterns in a reproducible and quantitative process.
5. The statistical likelihood that a set or subset of dentition patterns is unique to an individual is known. Subsets may include single teeth or a set of teeth. The most relevant subset is the front six teeth on the top and bottom of the jaw.
6. Human dentition is persistent to a sufficient degree to permit examination and comparison relevant to police investigation.
7. It is possible to collect and analyze human dentition patterns in a reproducible and quantitative process. Human dentition patterns may be compared to bitemark patterns in a format that is compatible from a practical and data perspective.

8. The process practitioners use to extract, and match features accurately and reliably relates human dentition pattern to bitemark patterns.
    - Feature extraction is reliable for bitemarks.
    - Feature extraction is reliable for human dentition.
    - Bitemark features may be related accurately and reliably to morphological variation in human dentition.
9. The probability of a particular feature match relative to all other feature matches has been established.
    - This assumption is necessary for the use of bitemark evidence as an identification method, but it is also relevant to the establishment of the evidential value of consistency between evidence and reference.

- Have completed reviews of a large section of the papers – this will be published in a journal article in the future and in an FTCOE report – have reviewed >60 papers (out of hundreds that were considered); will not publish the bibliography until publication

Assumption 1

- Minimal research has been completed to look at how long a bitemark will be forensically useful
    - Studies indicate that measures of forensic utility are reliably applied by examiners, but these metrics do not reflect research concerning the persistence of bitemarks on skin
    - One study determined that the usefulness of bitemarks on cadavers is limited to 30 minutes post-application (Bush et al. 2009; 23 bites on un-embalmed cadaver skin)

Assumption 2:

- Our review found no studies that addressed the differences between bitemarks and other skin patterns, including other injuries
- NIJ has funded research on bruising that links bruise appearance to skin color, the environment, the type of injury, the lighting the observer is using, and how much time has passed since the impact first occurred (Karen S. research at George Mason University)
    - This work suggests caution in the interpretation of bitemarks on skin, especially if interpretation relies only on bruising.

Assumption 3

- Cadaver studies have established a limited basis for identifying distortions associated with biomechanical factors including relative change in position after bitemark production. Bitemark distortions have greater variability than dentition patterns.
    - The "arbitrary distortion of a bitemark photograph to 'match' a dental overlay in an attempt to compensate for tissue distortion is not an appropriate technique."
        - Bush et al. 2009: examined 23 bites from a single dentition on cadaver skin
        - Bush 2010: 46 bites on 6 cadavers using a single dentition with anterior teeth removed; 38% of bites showed significant distortion based primarily on the stiffness/anisotropy of the tissue

        - Bush 2010: 66 bites on 6 cadavers; 38% of bites produced misleading distortions
        - Bush 2011: 89 bitemarks (one arch only, impressed not "pinched") from a single dentition on cadaver skin; cadaver model; bitemarks compared to 411 dentitions using principal component analysis and canonical variates analysis
        - Sheets & Bush 2011 FSI: 36 bites on cadavers with single dentition; affine methods cannot be used to account for distortions related to skin deformation because deformation is non-uniform and anisotropic
- Other studies have examined distortions on pig skin or associated with the removal and preservation of bitemarks, which produces distortions that prevent reliable analysis.

Retrospective studies of cases

- Bowers & Pretty JFS 2009: 49 cases in retrospective study for the forensic significance of the bitemark used for examination
    - Bitemarks with lower forensic significance are more subject to disagreement among examiners
    - DNA exoneration cases had similar significance to overall convictions
    - Child abuse cases were more likely to involve lower forensic significance
- There were 11 additional papers that were part of our review that were primarily case reports. In each case, reviewers determined that isolated case reports did not constitute scientific research
    - Issues include selection bias, lack of ground truth, no hypothesis testing

Assumptions 4 and 7

- It is difficult to establish an objective method for the quantitative characterization and comparison of bitemark images. Radford 2009 used Procrustes analysis of landmarks in bitemark images to argue that 3D bites should not be approached as 2D, plan-view images, but this only looks at relative deformation and may not represent the stresses associated with the physical act of biting into skin.
- The review included several technical notes published in forensic science journals. Although the technical notes demonstrate that the field has attempted to standardize its procedures, these papers did not include research to determine the reliability of the methods described.
- Several research studies have demonstrated significant inter-examiner consistency concerning the process of bitemark examination.
    - Suitability/quality metrics are applied consistently
    - ABFO guidelines were used by examiners and applied consistently
    - Overlay techniques are applied consistently. Several studies examined the use of software tools such as Adobe Photoshop to analyze bitemarks to demonstrate that practitioners could perform overlay techniques in a consistent manner.

Assumption 5

- Although the uniqueness of dentition is often assumed by bitemark examiners and forensic odontologists doing human identifications, most research studies involving large data sets and rigorous methodologies suggest that there are limits to the measurable uniqueness of an individual's dentition.

- Bitemark examination relies on anterior teeth, i.e., a more limited data set than is used in odontology for identification.
- In 1984, Rawson proposed an approach that includes 150 different positions of teeth as an approach to statistical characterization. Variables include presence, displacement, rotation, fracture variables. The Rawson study itself was found to have poor statistical analysis, but the approach was adopted by later studies.
- Studies of young people in developed countries show that differentiation is still possible but may be getting more difficult due to improved dental health generally.

Assumption 5

- Bush/Bush/Sheets 2011 IJLM and 2011 JFS:
  - 100 mandibular models scanned in 2D and 3D, 500 mandibular and maxillary models scanned
    - Discriminatory power is improvised using 3D images
    - Discriminatory power higher for mandibular than maxillary models
    - Two unrelated individuals in the set had matching mandibular and maxillary dentition in 3D. There were 197 maxillary-only matches and 51 mandibular only matches.
  - Two sets of dental models (n=172 and n=344)
    - 7/172 and 16/344 matches seen in anterior teeth in the two data sets
    - Did not reproduce Rawson results using his approach morphology
  - Dentition is not unique in the populations studied

Assumption 6

- In the bitemark literature, there was limited research concerning the persistence of dentition
- Trained examiners and dental student were better than untrained individuals at distinguishing bitemarks by children vs. adults
  - Whittaker, 1998: 50 bitemarks from casework, 108 observers of various training in bitemark examination and dentistry

Assumption 7

- McNamee 2005 JFS: examined use of Adobe Photoshop and Scion Image to produce and apply overlays for bitemark examination. Showed that 30 examiners produced overlays reliably

Assumption 8

- Studies indicate that interexaminer reliability is about 80% in bitemark examination using a variety of research methodologies with other outlying studies
  - 2001 study: 30 bitemarks with 2 biters in pigskin: 80% interexaminer reliability
  - 1975 study: 84 bitemarks in pigskin and wax; 78% accuracy for matches on pigskin, 98% on wax
  - 2011 study: 20 bitemarks in pigskin and 20 dentition casts; 95% true match case
  - 2001 ABFO workshop study: 4 bitemarks with a close set of 7 possible dentitions. ROC accuracy score of 0.86.

    - 2010 black box study: 30 examiners of varying experience, 18 porcine bitemarks with three pairs sets of dentition; error rates of 40% to 100%
    - 2009 study of 100 lower anterior dental impressions on cadavers. Between 11% and 86% of potential biters could not be excluded by human examiners of various expertise. Limited use of cadavers, depth of wounds only to incisal edge.
- Reviewers rejected several studies for methodological or interpretation issues, including one with bitemarks on 50 human volunteers. Other studies were included but were considered relevant to examiner consistency, not reliability
    - Rawson 1986: tested the 1981 ABFO scoring system; not "black box"
        - 3 bitemarks, 21 examiners; this phase validated the scoring system approach as a way to produce consistent findings among examiners. Teeth outlines only, so very idealized
        - 4 bitemark cases, 100 examiners; examiners told which bitemarks matched which dentition; so, this experiment only validated the scoring system
        - Showed interexaminer repeatability of 1981 scoring system, i.e. it works if all other assumptions hold true
    - Talabini JFS 2006: 50 bitemarks in human volunteers
        - Unique study of two methods of feature extraction (2D polyline and painting) without adequate reporting of statistical results

Assumption 9

- The field is heavily dependent on examiner judgement, so studies could be interpreted to apply to assumption 8 only
    - The fundamental uniqueness of bitemark features is not understood independent of a human examiner
    - Studies of distortion did attempt to perform an analysis based on various computational models

Q&A

- Looking at materials and methods very closely?
- What were your inclusion criteria for papers that were reliable?
- The field owes an enormous thanks to the Bushes and to Iain Pretty
- Work in this area need further funded research – more is better
- Reviewers
    - Generous to papers included
    - Approach was based on the Campbell approach

1.6 Ken Ascheim – Forensic Standards and Bitemarks A path Forward

Thursday, October 17th, 10:45 am

- How do we make some change?
    - Sometimes not even speaking for myself
    - He chairs several committees – ASB, OSAC Forensic Odontology, ADA
- Goal is to remain impartial, without representing the interests of a particular country or organization
- Bitemarks are not great evidentiary value
- Writing standards is a balancing act between two opposing theories (law vs. science)

    - Law: quantum theory – answers need to be binary (it is or it is not)
    - Science: wave theory – answers are part of a continuum
- Scientific certainty vs. reasonable doubt
- Garden of Eden:
    - First time that the Apple terms and conditions were violated
- Biting in wrestling – referees do not have to see the bite in order to call it biting, one should not assume intentional biting
- Where standards fit in
    - Part of our everyday life
    - Gives little or no thought to everyday products and services and how they work
    - Make modern conveniences possible
        - Industrial revolution – first standard was the railroad tracks; during Civil War the U.S. recognized the military and economic advantages of standardized rail tracks
- ANSI – what is a standard? A standard is a document, established by consensus, that provides rules, guidelines or characteristics for activities or their results
- Definitions adapted from ASB:
    - Standards: document that provides rules, guidelines, or characteristics for activities or their results
    - Technical report: explanatory or informational document that describes the scientific and operational underpinnings of a standard, but does not set requirements
    - Best practice recommendations: document that describes a method or technique that has been accepted as superior to alternative methods
- Core Principles of Voluntary Standards Development:
    - Openness
        - Participation open to all persons who are directly and materially affected by the activity (stakeholders)
        - No undue financial barriers to participation
        - Voting membership not conditional on membership in any organization
        - No unreasonable restrictions based on technical qualifications or other such requirements (participation is open to all persons who are directly and materially affected by the activity, no undue financial barriers to participation…)
    - Consensus
        - Substantial agreement has been reached by directly and materially affected interest categories
        - Signifies concurrence of more than a simple majority
        - Not necessarily unanimity
        - Consensus requires that all views and objections be considered, and that an effort be made toward their resolution
    - Balance of Interest
        - The standards development process should have a balance of interests
        - Participants from diverse interest categories shall be sought with the objective of achieving this balance
    - Due Process

        - Any individual, organization, company, government agency, or other entity with a direct and material interest has a right to participate by
            - expressing a position and its basis
            - having that position considered
            - appealing if adversely affected
        - Due process allows for equity and fair play
    - Dominance
        - The standards development process shall not be dominated by any single interest category or individual
        - Dominance is a position or exercise of dominant authority, leadership, or influence by reason of superior leverage, strength, or representation to the exclusion of fair and equitable consideration of other viewpoints

American National Standards Development Process

- Consensus,
- Broad-bases public review,
- Response to comments,
- Availability of an appeal,
- Subject to neutral third party

Current organizations working on forensic odontology standards

- ISO – new work item proposal, consensus building within technical committee (can take 1-10 years), draft international standard, final draft international standard
    - Just because a standard is written does not mean it has to be adopted
- OSAC does not make the final documentary standard
- SDOs are ADA (only standards related to standard clinical dentistry) or ASB (forensic dentistry)
- ASB – has 138 proposed standards
    - Terminology documents are crucial
- ABFO.org – standards and guidelines for evaluating bitemarks (revised 2-19-2018)
- ANSI-NIST-ITL 2011 – contacted by Brad Wing in summary 2011 – deals with transfer of information between agencies
    - Dental identification data set accepted (for use in Police Activity)
- NSF

Class characteristics of the human bitemark

- Features that allow identification of a generic human bitemark
    - Incisors are located toward the center of each arch
- Canines produce circular, triangular, or diamond shaped marks toward the edges of the arches
- Canines size determined by their incisal surface area and depth of penetration of their conical cusps
- If a premolar marks, it produces a cuspid like marking aligned with the arch and possibly a second mark a few millimeters within the arch, representing the lingual cusp

Individual characteristics of the human bitemark

- Size

- Shape
- Wear
- Restorations
- Rotations
- Diastemata
- Trauma/fractures

DNA is not always the solution

- DNA positive for an individual only – does this prove guilt? Does this prove innocence?

FAQs

- Why does the SDO process take so long?
  - This is not the task group's members primary job, takes time to get consensus
- Who decides what goes into a standard?
  - The content of a standards is guided by its scope
- Why are task groups draft documents not open to the public?
  - The standards development process can be difficult and often requires the melding of many different viewpoints; may involve talking points that will be debated; need a non-judgmental environment where good ideas can flourish, and less proven methods can wither
- How will my voice be heard?
  - There is always public review
- Why do you need an SDO if OSAC is writing the drafts?
  - OSAC is not an SDO – OSAC focuses on the science of it
- Is the OSAC task group's word final?
  - OSAC has multiple layers of review in place before a document is sent to an SDO
  - Once a document is at the SDO the are additional check and balances in place to make sure all points of view are presented
- Is there an appeals process?
  - All ANSI certified SDOs must follow ANSI protocols concerning the appeals process
- What happens once it is in the OSAC Registry?
  - The OSAC Registry is a listing of high quality, consensus-based science-based standards and guidelines that the OSAC community feels has technical merit
  - To be posted to the OSAC Registry, consensus-based document must pass a review of technical merit by forensic practitioners, academic researchers, statisticians, legal professionals, quality managers, and measurement scientists.
  - OSAC does not have authority to enforce standards.
- Can a standard or best practice document be "aspirational"?
  - It is not clear; we want to promote the road to the better, but implementation must be clearly delineated and obtainable for the standard to be implemented

Hurdles to overcome:

- Should bitemarks be an industry standard or a consortia standard?
- How do you report quantitative data on a qualitative assessment?
- How many bitemark cases were CONFIRMED by DNA?
  - The denominator is as important as the numerator on the Innocence Project

- What is an acceptable error rate for those individuals wrongfully incarcerated?
- What is an acceptable error rate for the public for those individuals wrongfully exonerated?
- How do we differentiate between bad science and bad jury decisions?
- Where does reasonable doubt fall?

Standards are not a field of dreams – consensus building is required

OSAC documents:

- Forensic dental data set
- Human identification by comparative dental analysis
- Human age assessment by dental analysis (proposed ADA)
- Bitemark analysis… (sitting in the drafting committee)
- Seed document created at onset of OSAC – current task group chair is James Lewis

ISO TC 272 Forensic Sciences Technical Committee

BM changes to a 5-part document:

- Part 1 Terms/definitions
- Part 2 Recognition, collection, and recording of evidence
- Part 3 Evidence analysis
- Part 4 Evidence interpretation & comparison
- Part 5 Reporting

What standards will not do!

- Sometimes we will be wrong, our goal is to define the scientific capabilities and processes – Voltaire, Blackstone, Ben Franklin (100 guilty vs. 1 innocent)

With open minds and hard work, we can produce forensic standards that will not only prevent guilty persons from escaping justice, but will prevent innocent persons from suffering injustice

1.7 Amanda Sozer – Breakout Overviews

Thursday, October 17th, 11:30 am

SNA: Amanda Sozer, Ron Fazio, Rick Tontarski, Joe DiZinno, Aislinn Berge, Brieanne Knight, Timothy Graham

SNA is tasked with capturing the information and putting it into a format that is useful

- Thinkshop is designed to promote discussion
- Everyone will discuss the same topics in turn
- Facilitators and notetakers will capture ideas from the group

Breakout sessions will take place in groups.

Group reconvenes – report outs, further discussions

Expected outcomes

- Report outs from breakouts will result in publicly available NISTIR which will identify needs, gaps, standards, issues, etc., priorities will be ranked and grouped

- Identify areas for further research – want everyone's voice to be heard. Concern with opinion-driven rather than evidence-based discussions

2.0 Breakout Sessions

2.1 Science Question 1 - Are there measurable characteristics or features in human dentition that vary among individuals and are persistent within an individual?

2.1.1 Focus Area A - What measurement method(s) provide the best information for capturing reliable information about the detention?

What measurement method(s) provide the best information for capturing reliable information about the dentition?

Group A

Discussion:

- Class and Individual Characteristics
  - There are definable features - class characteristics and individual characteristics - to dentition that have been defined historically and there is a lot being done (reverse engineering – what is unique).
  - ANSI-NIST has definitions of class and individual characteristics.
  - Bush papers discuss measurements.
  - Features are designed in a way that can be measured, but not sure if it defines uniqueness.
- How is dentition measured?
  - Impression evidence (casts) as an exemplar
    - Dental impressions are not a measurement, but a replication of the dentition
  - Line-ups
  - Ruler and protractor to measure the width and height
  - 2D and 3D measurements. There are standards. Circumference, which changes when you move up and down the tooth.
- What measures are most probative? Best way to measure the individual characteristics specifically on dentition?
  - Well-developed field now (white light, laser scanning). Orthodontic scanning gives you an STL file
  - Methods have always been there whether you use a ruler, string, or high-tech ruler. We are measuring physical characteristics of the border of something
  - Measuring the physical characteristics of the external surface of tooth. Comparing to impression, it makes a difference.
- What you are measuring? – might take different measurements based on what you are measuring
  - Method may vary based on what you are trying to record
  - Can do it with dental model
  - Modern version of handheld scale (orthodontic scanning) is used to scan a model (dental stone) or of the actual dentition
  - Could start using a mold of the dentition
  - The scanner takes a metric measurement, whether handheld scale or 3D laser scanner. It is a scaled device.

- What is an example of models? What type?
    - Dental stone models, epoxy, etc.
- What are the limits of these techniques?
    - Most overlays are 2D representations of 3D dentition
    - Limit to resolution - fundamental limitation to any measurement. How accurate can you measure a person's weight – it is based on the ability of the scale.
    - 3D laser scanning is very accurate and has much less of an error rate
    - Challenge of 2D representation of 3D features
    - An individual using the same ruler and model may get differing results; therefore, you must report uncertainty associated with the method/tool.
    - Accuracy can be more consistent by deciding on what to scan and what points to measure when measuring tooth length
    - Black box algorithm for scanner to measure the teeth. The 2D representation (digital pictures) are limited physically by the operator of where you determine two points are for measuring. There is a resolution limit. Is the difference in what you are measuring from a 2D image significantly different from what you are measuring in 3D?
    - Tooth height is not a good comparison.
    - There is an objective element to measuring - not a bias.
- How to address challenges (whether uniformity or tool use) related to 2D vs. 3D?
    - Replicas are important to discuss for 2D and 3D measurements. We need to know how many replicas needed - take the average of measurements for accuracy. Knowing measurement variability could help drive us towards the number of replicates needed.
    - Depending on depth, one assumption about bruises that darker areas there is more force on the skin. 2D provides length of the information. One method to develop, because of volume difference due to circumference of teeth – people will take sequential outlines and immerse model into wax at different depths. Each wax impression would show the different layers representing the 3d.
    - Most evidence is just bruising
    - It is advantageous to have a 3D bitemark.
    - There is a desire to have a 3D bitemark converted (translated) to 2D overlay, particularly for exemplars. However, overlays will still be 2D unless sequential.
    - 3D is a pyramidal solid in the x, y, z direction, where Z = thickness.
- Ability to determine force and impression impact
    - The more you push teeth down, the more the width of the tooth transfers on the skin. Skin is elastic so there is deformity and it might not be replicated even a minute later.
    - If you have 3D info (cast) and a puncture in the skin, can you associate it?
    - Teeth are not all the same length; therefore, each tooth would hit at different times upon coming into contact with a substrate. How far are the teeth off of the plane?
    - Do fillings in the teeth play a role? - No, the density of the filling doesn't typically matter in the elasticity in the skin.
- Dentition measurements are well defined
    - Dentition measurements are not an issue. Measurements would be reproduced in the 99th percentile. Setting aside pattern injuries and distortion, just

measurements of dentition, there is good reproducibility among general odontology population.
    - Issue is what we are measuring, not the ability to measure
    - Need to address the variability
- Is there some gap that is associated with the force and pressure and impact when looking at that in regards to dentition?
    - Force of the dentition. Issue deals with the skin as the recording medium.
    - Method of measuring and where to measure it. What part of the tooth are we measuring? Not a method issue.
    - We can make the measurements accurate. If you are asking how wide the tooth is, where on the tooth do you want us to measure. Where to measure – the place measurement is taken is the issue.
- Final Considerations
    - What is the standard way of doing measurements in the discipline? Do we just measure once?
    - These measurements are only to show uniqueness. This has to deal with the claim and the uniqueness of the dentition. Is the error rate greater than the differences? It depends on how many different types of measurements you take. Method is well defined but for determining uniqueness we need to determine what to measure.
    - No standard yet in resolution.
    - A protocol is needed for the place the measurement is taken.
    - Variability – need to understand it to determine the number of replicas.
- Summary of Group A
    - Agreed that 3D works well
    - Measurements are not the issue, what to do with the measurements is
    - Where are the measurements taken and what are significant?
    - Measurement variability may define how many replicas are necessary

Captured on flip charts:

- Agreements - Knowledge of Gaps
    - Features: measurements, best mechanism to capture
        - Measurements are not an issue - pretty good at taking measurements
        - 2D applied to 3D dentition is a gap
    - Resolution
        - Protocol needed for taking measurements (where on the tooth do you measure?)
    - Ability to determine force/pressure impact on 3D vs. 2D
    - Steps to Move Forward
        - Wax impressions
        - Orthodontic scanning (handheld)
        - Dental stone/epoxy
        - Know measurement variability defines requirement of replicas

Group B

Discussion:

- 3D dataset is the best and most reliable for methods
  - 3D not a challenge, 3D scanning is the measurement to use
  - 3D scan is done on original teeth (best mechanism, most accurate), should not be done on a cast but can be done. 3D scan on original teeth is timely.
  - A 3D model or cast does not provide same info - secondary source
  - Models are less accurate over time. Wear, chip, humidity can affect them. A lot of variability with a stone model. The models wear out. Amaterial, plaster material. Imbibe moisture and rub against each other, they can lose accuracy
  - Measure of uncertainty if you don't take a 3D scan of the mouth. More uncertainty of dental model scanning
  - 3D scans are more transferrable
  - Direct 3D scan provides the degree of accuracy and stability over time and transferability
  - You can 3D scan the model to keep forever and 3D print it at any time with the same level of quality as the day you took it.
  - Same level of fidelity of the cost as the 3D? On one hand, cast is just as good if it's all you got but there are limitations. May not be as accurate or precise?
    - Depends on user. If you mix materials properly that are not expired and poured correctly. Ratios for the mixture, this is not the way it is typical in real practice. They eyeball the ingredients. Where it is stored. If you scan an item, it is accurate.
    - Chemical reactions occur, changes can produce minor inaccuracies like air bubbles. Scans don't change accuracy over time.
- Quality of 3D Scanning
  - Does it matter who takes the scan?
    - You need to have a degree of familiarity with scanning
    - If you are missing an area in the scan, the computer will tell you it is missing. In the end, the scan will be complete and the same as an expert. It might take more time.
  - X-rays aren't always perfect the first time you take it. Is that the same with 3D?
    - No, you look at shadows with x-rays. Scanning device goes in the mouth and you see the scan come together right there.
  - Cheap or expensive scanners?
    - There are high- and low-quality scanners, but quality and resolution increase all the time
  - Group accepted that scanners work at a level that is clinically acceptable
    - Material in the mouth and scan has to be accurate within one micron
    - Existing scanners have tremendous resolution to characterize dentition that exceeds models.
  - Is it necessary to define resolution?
    - Not provided by manufacturers. Clinically accurate enough.
    - Far more accurate than plaster models and better every year.
    - Create minimum guidelines for the different data sets to later study or compare the identifying characteristics, you can compare it

        - ADA has specifications, that info might be there.
    - Do we know what reproducibility is for 3D scans?
        - There have been studies on this, and it is reproducible.
        - Need comparison study between different models and 3D capture devices.
        - Studies have to exist as the FDA will make sure this is necessary to be used on humans.
        - Getting info from manufacturers is difficult.
- Dentition Measurements
    - Arch shapes, distance between teeth, inter-canine distance.
    - Identified what was needed in the data from dentition measurements.
    - Inter-canine distance – measure from center point? Where is the point A to point B measurement? What is uncertainty of measurement? Not always reproducible between dentists?
    - Are the measurements between dentists significantly different?
        - There is literature on the significance of the reproducibility
    - Will there be variations in measurements and is it critical?
        - Not critical. Don't know if it really impacts or not.
    - Take models and get an accurate 3D model. Coming up with measurements is a different question.
    - Metric analysis of models will help to develop to determine if the dentition is identifiable. Comparisons to injury – a small degree of difference between dentists is negligible.
    - Can you define these measurements? - These can be created. Always some level of uncertainty.
    - There is a degree of individuality of teeth and should be discussed through and distribution of the individuality and how often is this encountered in and across populations
- Gaps associated with 2D and overlays?
    - Yes, there are gaps. Using pseudo 2D techniques.
    - Overlays are only looking at incisal edges, just the biting surface. Things like height, incisal height are not in a 2D overlay. Chips and sharp edges are not appreciated. Lots of weakness. Can use photographic overlay.
    - Scanning dentition in 3D and slicing the 2D
    - Taking 3D data and horizontal relationship of teeth (incisal height) – lower teeth and higher teeth. Horizontal = incisal height
    - Using orthopix, it is in the software that you can slice the tooth at any height. If you have 6 teeth at different heights, with the software, you can decide which tooth touches the horizontal plane first. Should be getting different impression if it stops at the first tooth vs. if the tooth goes all the way through the substrate. Trauma in the skin would be expected to be equal for the teeth of similar height
    - Summary - using 3D data to assess incisal height impact
    - In addition to other factors, the arc is a factor. You cannot see this alignment in the 2D alignment.
    - Still 2D in overlay, you can't see that the teeth were higher than the other. Even with the software on orthopix, you only see one plane.
    - Summary - a challenge is the ability to factor in the various planes of the teeth

    - Acquiring the 3D data because it gives you all of this info, 2D does not provide the data. 3D data can do scans to determine the height of each tooth to be used later to see what impact it has on penetration - like making a topographical map
    - Group consensus - if you capture 3D info of teeth appropriately, you have sufficient data to do a variety of things to characterize dentition
- Summary from Group B
    - 3D is good
    - Companies have info on resolution
    - Standards defined for measuring dentition
    - Casts are not sustainably accurate.

Information from Flip Charts

- Agreements - Knowledge of Gaps
    - Limitations of 3D scanning (resolution)
    - How individual characterization of dentition may be
    - Guidelines for data sets (ADA may have):
        - Meta data
        - Resolution
        - Precision of measures
    - Where are measurements taken/precision
    - Using overlay to compare to dentition
        - Height
        - Chips
        - Incisal height (horizontal height) - above/below plane
    - Ability to access slices of tooth
- Steps to Move Forward
    - 3D scanning
        - Primary source - mouth
        - High accuracy
        - Sharable immediately
    - Casts are less accurate
        - Indirect
        - More uncertainty
        - Subject to damage (humidity)
        - A lot of variability depends on user
    - Currently sufficiently accurate for needed accuracy
    - Study to compare devices
    - Use 3D data to assess incisal height impact

Group C

Discussion:

- 3D Scans are the best methods to measure dentition
    - Group C general consensus that there are no gaps associated with 3D scanning
    - The 3D scan is of the entire mouth that can be taken of the mold or in vitro. Both the molds and in vitro are accurate to microns. 3D scan is the same whether you take it of the impression or directly in the mouth.

        - There was consensus that there is no difference in scans of models and the scans in vitro
    - 3D scanning can be done without getting an impression taken. The digital scan is then emailed to the lab and then can be made. Crowns can be made from these scans. Very accurate.
    - Group consensus that casts and 3D imaging are equally as good.
        - For research, you ultimately still need a 3d scan for measurement.
    - Non-color scans are false color - tell you nothing. Color has no relevance to bitemark analysis. Color calibrated scanners to choose tooth color.
- 3D Concerns and Considerations
    - 3D is not accessible and expensive - go from hand scanners and tiny scanners ($70-100K).
        - 3D scans can't be used in private practice due to cost
    - Are they calibrated?
        - Hundreds of studies on scanners on different types of crowns.
    - Regardless of 3D scanner used, you get reliable measurements
        - Crowns always work, always accurate. Standard impressions also give you detailed information.
    - Impressions
        - Need to scan impressions to take measurements
        - In private practice, you take physical measurements of impressions.
        - Density or porosity concern
        - Dentists/Odontologists don't look at density/porosity for bitemarks in forensics
        - This will not tell you anything about predictability as there is nothing to measure the age associated with teeth
        - Porosity is added to the parking lot
- Variability in measurements
    - Variability is irrelevant.
    - When you do repeat scans from a trained scanner it is not variable. Reliability of good quality scan done well.
    - If you know what the measurement uncertainty is of the systems, we can report that
    - Need to agree on threshold level to accept vs. not accept with scans.
- 2D Overlays
    - 2D overlays have limitations. You are taking 3D objects and reducing them to 2D objects. 2D overlays do not consider 3D. Dentists alter overlays to take into consideration 3D.
    - Impressing casts into wax and seeing what levels the teeth press into
    - Technique most recommended is using flatbed scanner to produce 2d scan of 3d object
    - Length of teeth is lost with 2D scan. Limitation of 2D vs. 3D
    - Gap is resolved by only using 3D scanning
    - There is a study that took models and made overlays to show 3D in 2D (like CT scan).
    - 3D is not always the best for demonstrations, use 2D for demonstration

    - Can outline edges of 3D scan to measure and make a 2D CT scan that you don't get from an overlay
    - Challenges with false matches with multiple layers?
    - If you find congruence at 1mm down but ground truth is tooth went 3mm down. Need to create that way of testing. No way to test skin to find out. How much value do you get out of 3D because we don't know how deep the bitemark went?
        - For a tooth that is shorter, you still get some degree of 3D data
        - Need to test the deepness of the bite
- What data needs to be harnessed out of 3D scan?
    - An efficient set of features vs. high degree of features – not sure how to measure this.
    - Image processing – forensic app needs to be written
    - Need a plug in of some sort and software for dentistry
    - There is no automation in this process - PCAST criticized lack of automation
- Summary of Group C
    - Agreement that 3D is accurate whether cast or scan
    - Limitations of 2D and 3D is a gap - can be approached by taking a "CT Scan" of the tooth length as a 2D image

Information from Flip Charts

- Agreements - Knowledge of Gaps
    - How to harness 3D scan data and what to extract
        - Efficient set of features/discriminate
    - Define appropriate thresholds
    - Limitations of 2D compared to 3D
- Steps to Move Forward
    - 3D scanning
        - Mouth
        - Or cast
        - Reliable measures with range of scanners
    - Series of 2D exam
    - molars to show length variation
        - CT scan-like
        - Use 3D scan data to show variation that might be in 2D

2.1.2 Focus Area B - How do we appropriately collect information to create population databases that can be used for scientific and statistical analysis of human dentition?

Group A

Discussion:

- Population databases:
    - Present databases can be leveraged, and information gathered from dental schools/dentist
        - Dental schools have standardized practices
        - Would need to pay for high res scans

        - Not representative of populations
    - Odontosearch
        - Over 100,000 participants/records
        - Rarity/uniqueness of pattern fillings
    - Need to agree on what characteristics to collect
        - What variables are worth collecting?
            - Define characteristics to collect
                - Need database first
                    - Data determines metrics
        - Does cosmetic dental work impact this?
            - Goal to make dentition similar from person to person
    - Need to secure funding
    - Can diet affect teeth?
- Knowledge Gaps:
    - Money for data collection
        - High resolution scans are expensive and time consuming
    - Understand database design
        - Need a database to define the metrics we will be looking for
    - Representative population issues
        - Socio economic, medical history, geography
            - Need to have inclusive studies
    - HIPPA constraints
        - Dental records are protected
            - Are teeth and dentition PII?
- Challenges/Steps moving forward:
    - Follow anthropologists lead on identifying and relating variables of bone structure
        - Applied to dentition
    - Design and implement 3D scan and or model databases
    - Engage other experts/fields to identify process to identify variables to look for in database
        - Facial recognition
    - Identify factors on variability
    - Representative population inclusion

Group B

Discussion:

- Population databases:
    - Present databases can be leveraged
        - Invisalign 3D, OrthoPics, existing casts dentist's offices
        - Not representative of populations
        - Use of AI learning to look for areas of statistical variation
    - Need to agree on what characteristics to collect
        - What variables are worth collecting?
            - Define characteristics to collect
    - Need to secure funding

- Knowledge Gaps:
    - Statistical interpretation has to vary between 3 sets of dentition
        - Child, teen, and adult
    - What is rate of variability?
    - Identify and classify class characteristics
    - Identify and definition of characteristics
    - Natural vs. post orthodontic dentition
        - Discussion on the science of orthodontics is to normalize teeth
    - Understand racial differences in dentition
        - If you want to know uniqueness, then you need to collect lots of samples
    - Representative population issues
        - Differences in importance of cosmetic dentistry
        - Regional differences
        - Socio-economic
    - Challenges/Steps moving forward:
    - Leverage existing databases for forensic purposes
        - Invisalign, OrthoPics, molds
    - Find a way to get information across the board
    - Survey other experts to decide on variable features (e.g., facial recognition algorithm – not just color of eyes; dependent on the data used to train the algorithms)
    - Interview dental practitioners for a list of variable features to identify in database
        - Determine what characteristics to collect
            - Look at biting surface vs. entire 3D surface of tooth?
            - Arch characteristics
            - Leverage AI eventually
                - Generate list of characteristics from dentists
    - No good current database on uniqueness of human dentition

Group C

Discussion:

- Population databases:
    - Need large 3D imaging compatible databases
        - Needs to be longitudinal study over 20 years
        - Images every 6-12 months would be ideal
    - Present databases can be leveraged, and information gathered from dental schools/dentist
        - Dental schools have standardized practices
        - Not representative of populations
        - Need representative samples
            - Ongoing surveys going on in US and Canada (NHANES)
                - Over sampled for Hispanics and Blacks
                    - Selection is rare
    - Need to agree on what characteristics to collect
        - What variables are worth collecting?

        - Define characteristics to collect
    - Need to secure funding
    - Challenge of IRB approval for collecting information
- Knowledge Gaps:
    - Lack of cohort database (longitudinal study)
        - Time consuming and expensive
    - Representative population issues
        - Dental practices are not inclusive
            - Need to get records or scans from across several socio-economic backgrounds and regional backgrounds
    - Which to do first database or decision of variables that are important to measure
        - Chicken vs. egg
        - Would database be useful?
- Challenges/Steps moving forward:
    - Leverage dental schools for information on those that do not go to formal dentist
    - CDC and NHANES to get dental scans or create NHANES like program for dental scans
        - Due to mission of NHANES this is considered unlikely
    - Critically discuss if database would prove probative
    - Determine variables to collect as part of database
        - May be after database is made
    - Need to study and determine error rates

2.1.3 Focus Area C - What are the most probative features/parameters to use and what are the limits associated with each?

Group A

Discussion:

- Probative Parameters/Features
    - What makes these features probative?
        - Usually statistics
    - Class characteristics and shapes
        - Variations
    - Vertical height
    - Arch width/arch form
        - Cuspid to cuspid
    - Tooth condition
    - Restorations
    - Tooth condition (pathological)
    - Angulation
    - Occlusion/dental alignment
    - Tooth shape
        - What metrics are we using?
        - Shape itself is hard to define consistently
    - Metrics of tooth height are pretty well defined

    - In terms of giving a statistical number to it - has been well defined in populations
        - Goal is to make those normative
    - How do we deal with normal biological deviation vs. outlier?
    - Do orthodontists make people non-unique?
    - Cold cases aren't done BECAUSE there are changes that occur over time
    - Some features change quickly, some fast, and some at a normal rate
- Knowledge Gaps
    - Can a 60yo make the same bitemark as they did at 18yo?
    - Do orthodontists make people non-unique?
    - Collecting a database of 3D images
    - Can a database of the relevant population be created?
        - May discover trends were not what were thought of before
    - Would like to know the frequencies of all the features?
    - Open population studies – what images would be needed?
    - Some features could change over time (persistence issues) – rescan people over time
    - Bitemarks are compared close to the time period of the crime (Q-K comparison)
    - Cold cases not done because of mouths are likely to change over time? (a rare occurrence)
- Challenges/Steps to move forward
    - Acquisition of data
        - We don't want to exclude
    - Invisalign - getting the data doesn't fit in with their business model - clients may not know they are collecting said information
    - A protocol that is consistent across laboratories – precision is better than our eyes can see
    - 3D database of images you can recreate whatever you wanted
        - Want a list of specific things to measure
    - Orthodontists have a wealth of info, but often young people with crooked teeth
    - Studies over time
    - Open population database
        - Operation and maintenance
        - Outliers – always going to be there, how do we deal with them?
    - Teeth considered PII?
        - An issue?
        - HIPPA release?
    - Funding

Group B

Discussion:

- Probative Parameters/Features
    - Number of teeth should include supernumerary teeth
    - Alignment
    - Normal dentition vs. restoration or orthodontic
    - Vertical height

  - Shape of the arch
  - Intercanine distance
  - Presence/absence of teeth
  - Diastema - Spacing of anterior teeth
  - Incisor ledges
  - Normal dentition vs. restoration or orthodontic
  - Chips/breaks
  - We want features that are highly distinctive
  - Intrinsic parameters
    - Age - wear/change
    - Ethnicities - shape of teeth and arch form may vary
- Knowledge Gaps
  - Population specific traits - and how far back does it go?
    - Ethnicity: self-reported vs. actual
  - Statistical evidence, but would have to be advanced
    - Digitally
    - 3D
  - Ethical and legal issues in obtaining data
    - If the patient dies, then HIPAA doesn't apply
      - Can we use decedent data to make a longitudinal study?
  - If you're getting all of the info from Invisalign - is your data applicable to only a certain socioeconomic group?
    - Is there a difference in the quality of data? Can those be compared?
    - Invisalign data - could it be used in DVI
  - Look at factors anew due to admixture in US population
  - Trait frequency may not be the same
  - The goal would be to obtain a database that is representative of the general population
  - Statistics analysis
  - Need for a Database
- Challenges/Steps Forward
  - Can university dental clinics be used for data collection?
    - Expedited exempt for IRB/human subject if we go thru universities
    - Masters or PhD study in this
      - University of Colorado could be a start
  - Studies
    - Which features are independent vs. relation to one another?
  - Arch shape to be reduced mathematically
    - Or other features
  - Proprietary information
  - Getting the data IS the priority
    - How you look at it comes second
    - Once you have the data you can do it multiple ways

Group C

Discussion:

- Probative features/parameters
  - How do we know what is probative?
    - Open to interpretation
  - First, we need to understand 'normal' dental anatomy
    - Then get into the things that make people 'unique'
      - Uniqueness is a great concept, but it's not NECESSARY for bitemark analysis
- Forensic dentists use the following characteristics to determine which features to characterize an individual's dentition:
  - Arch form
  - Intercanine Distance
  - Gaps/spacing between teeth
  - Natural vs. pulled/missing
  - Pathological changes vs. odontological
    - Dental work adds to individual characteristics
    - Orthodontic teeth are more uniform compared to non-ortho teeth
  - Alignment and rotation of teeth
  - 3d orientation
    - WxLxH
    - Individual vs. 6 anterior vs. all
  - Grinding, chipping, etc.
  - When a tooth is extracted, other teeth may start to move
  - There is no data to support that these features are examined because they have a lot of variability or because they are observed?
  - Dentition is fairly stable over time
  - Small patterns and big patterns - It's the pattern that matters
    - Small patterns are less informative - only using a portion of the teeth, and not the entire pattern of the teeth
    - What are the patterns, how stable are they?
    - We need to refine what is actually studied, be broader and less narrow
- Knowledge Gaps
  - Need a large database for studying patterns of dentition
  - How long is the data good for?
    - Need a longitudinal study as dental patterns may change over time
  - Study should examine all teeth instead of selecting few
  - 'Cannot exclude' bite mark analysis result does not give a level of confidence
  - Is there data on the dental patterns of orthodontically treated teeth?
  - Data needed on features of the dentition -> both class and individual
  - Have we defined class/individual characteristics?
- Challenges/Steps to move forward
  - Funding needed for studies
  - Studies
    - Likelihood ratio vs. 3 categories of inclusion/exclusion
    - Multivariate, need a LOT of samples - 10,000 - might be undoable

        - Studies using points for statistics
    - Design research protocol to collect 3D data and store it in the same way to be analyzed across the world
    - Proprietary data - itero/serac/e4d currently have databases that contain the info, but it is locked down and unusable
    - NHANES

2.2 Science Question 2 - Do bitemarks transfer measurable characteristics of the dentition to the substrate?

2.2.1 Focus Area A - What imaging and measurement method(s) provide the best information for capturing reliable and reproducible information about the bitemark?

Group A

Discussion:

- Are the reliable and reproducible features of the bitemark being captured?
  - What variations are there? One would like to get the full image of the tooth on the bitemark. But we do not get that info, we get a small portion of that info at best.
  - Maybe what is being used now with photography cannot be examined.
- Different Imaging Methods to View Skin Bruising
  - What is the impact of forces on the tissue below? What if we can get more info about what the bruise looks like.
  - The word "bruise" should not be used - better to use "tissue damage".
  - You can see under pigmented skin to see bruising below using IR at different wavelengths. Using IR, you can see if a small unseen bruise is there, especially in individuals with darkly pigmented skin
    - IR equipment is readily available and widely purchased
    - Certain IR wavelengths are better than others
    - Is there a study that shows different wavelengths and filters and spectral contrast? - Yes, Frank wrote an article that address it. If there is an abrasion, the skin can heal with the same pigment, it is the pigment that the UV sees.
    - Focus the laser on a spot and focus the spot deeper. Photons are focused that gives you a depth profile, can be used for imaging skin, blood, and capillaries. This is what is done in spectrometry, not sure about odontology and bitemark analysis.
    - Using IR, you can find a specific blood vessel in the middle of a hematoma and remove the images of the hematoma. You can just look at the tissue. Nondestructive method, can do it on a living person
    - Since the tooth is going into a solid structure, you need to know the plasticity. 3D imaging does not answer the issue with plasticity
  - Bruising may mask that there is some dentition in the skin that you can't see through the bruise, other imaging technology you may be able to see this. Has there been research?
    - Yes, research on ideal imaging technologies. Conclusions of these are that you can see under the skin using IR, penetrates 3mm below the skin. Near IR that does the trick. Deeper wavelength like thermography you can't.
    - With IR, you have the potential of seeing the micro veins that were cut by the tooth on the way into the skin. You can still image what the damage was to the tissue itself. That can then give you an insight of the geometry of the bite infliction.

        - Imaging technology could show blood vessels and ruptured cells and not where the blood flows. This is used at 785 UV (near IR).
    - Info on bruising is the least helpful in bitemark analysis, follows the path of least resistance. Doesn't help you in making a comparison such as something that occurs on the surface of the skin.
        - Humans all bruise differently due to the variability of the tissue
        - The bruise will not show morphology of the tooth that made the bruise
        - Depth of the impact doesn't tell you about the power of the bite, the shape of the tooth, etc. Doesn't tell you about how you were hurt.
    - Other imaging options include confocal Raman spectroscopy
    - 3D Imaging for Bitemark Analysis
        - You can look at the depth of the impact, use as a measuring tool for depth
        - How deep did the damage go? Can subtract things out to see different damages.
        - This gives you insight of the depth, breadth, angle of the tooth in the bitemark
    - Skin Distortion
        - IR and other imaging methods capture unreliable evidence due to the unknown amount of distortion in the skin
        - You can look at damage, you know very little about the inter human variation
        - There may be a possibility with imaging techniques, such as IR, that we can see through distortion. Requested to look more into IR and other imaging techniques.
        - The value in this imaging technique is that you can distinguish between things that are not human bitemarks. More information matters but maybe not enough if you don't answer what you don't know about the human body.
- Effectiveness of Bitemark Analysis
    - DNA testing has disassociated these problems from actual court cases. Bitemark analysis is essentially ineffective.
    - If you look at the picture of the dead kids, how could you not want to make this science better?
- Reliability and Accuracy
    - What do we consider reliable?
    - You can take an image three times, but it doesn't represent what is shown, but if it is wrong, it is not accurate.
    - Looking at the biting edge, introducing depth to info we have. Maybe that can tell us about activity or violence. If we are talking about using bitemark to connect to a person, we are limited to biting edge. This is conditioned on what we get from the biting edge.
    - Does the skin accurately produce what the biting edge transfers?
    - You might be able to improve BM with imaging techniques. Collecting more info is a good thing.
    - Summary - explore imaging possibilities, improve spectroscopic imaging and all other imaging techniques

    - You don't get to examine teeth first. Examine pattern injury first. Record it using photographic image correctly. Need to start here. Once we have the image, then do the comparison using dentition. If there is distortion, need to throw it out.
        - How to quantify distortion?
        - Distortion may be a problem or may not. If there is distortion it will make it more unlike the actual biter's teeth.
        - No way of knowing if distortion includes or excludes
        - You may wrongly exclude but may include something innocent
    - Characteristics that are there are individualistic, the bitemark and teeth, have similarities.
    - Scales have color references on them. Correct images to the most correct color we can. Not based on reference.
- Group A Summary
    - Not capturing as much discriminating info as we would like
    - Does bitemark inform dentition?
    - Depth of damage
    - Learning about what we can't see beneath skin
    - Exploring imaging possibilities, subtracting info
    - Inform angle, width, depth
    - Agreement that looking at a variety of imaging techniques is of value

Information from Flip Charts

- Agreements - Knowledge of Gaps
    - Wax impressions
    - Orthodontic scanning (handheld)
    - Dental stone/epoxy
    - Know measurement variability defines requirement of replicas
- Steps to Move Forward
    - Explore imaging possibilities
    - Filter/subtract to see tissue damage, may inform on tooth geometry (Raman?)
    - Angle
    - Width
    - Depth

Group B

Discussion:

- Distortion of a Bitemark on Skin
    - There are distortions on bitemark that we can't account for: movement of body after the bite is inflicted, tension markers, longer lines, pulling back or movement of live person after receiving the bite (physiological rotation of the jaw). No matter how good measurements are, those are things that we can account for, or on occasion, recognize.
    - Need to also account for the movement of the teeth, jaw is closing, and person being bitten moves to cause distortion.

  - Rotation of the jaw upon a bite introduces distortion that we cannot measure and sometimes it can't be recognized. For example, we wouldn't know if a person is bitten when their shoulder is extended.
  - Does the pattern of contusion reflect the dentition (the teeth) that created the bitemark in the skin? How reliable is the pattern?
  - Time increases the issues of distortion; however, sometimes the dentition is prolonged post-mortem rather than antemortem due to the remains being placed in a freezer.
  - Skin is too variable a material.
- Analysis Techniques
  - Imaging modalities can show bruising better than others.
  - With artificial learning you need to have 1000s of knowns
  - If you have a single bitemark and take multiple photos and wavelengths of light over a course of a few days, does this strengthen our ability to say it's a bitemark and can we profile the evidentiary value of it?
  - Really bad teeth (crooked, gaps, etc.) are good for identification. Perfect teeth are hard to uniquely distinguish.
  - Some skin conditions might mimic a bitemark. A time series would be helpful in order to differentiate a bitemark vs. a skin condition (5 skin conditions that can mimic a bitemark).
  - Do measurements really mean anything if you are measuring distorted skin? How to calculate or take into account distortion?
    - Comes down to pattern recognition instead of specific measurements
    - 3 stage process - look at class and individual characteristics. Does the pattern relate to a set of teeth?
    - Relative measurements are more important than absolute measurements
  - Can't always be certain that pattern in skin is reflective of teeth or specific teeth. Need to study and determine how teeth impact skin. How are patterns formed and how they are assessed?
    - Need to study how patterns are formed on an awful substrate that has a degree of variability and distortion in comparison of teeth to a bitemark in a subjective nature of interpretation
    - Limitation is the ability to create and understand the variation that can exist with real-world samples
- Ideal Information Gathered from Contusions
  - Ideally, you want to know who made the contusion.
  - You need data that establishes ground truth. We don't know this until we have a large data set. Limitation identified as the lack of data - no ground truth data to assess comparison variability. Need to know ground truth.
    - Need a lot of people that suffer bites (knowns), bitten by dogs, horses, pigs, humans. Some are sitting, standing, bitten in leg, some aren't bitten, some have prior injuries. They don't know they are going to be bitten so they act like a victim. Need to capture the biting and then measure the bite and teeth

        - Ex. ABFO odontologists received a picture via email which ~80% of the recipients believed that it showed a perfect bitemark. The mark turned out to be skin where an EKG pad was removed.
        - If a case goes through court, how do you know it reflects truth?
    - Biting dynamics are what we don't understand – the tool hitting the surface. We need to learn this.
        - Lower teeth hold the object, the top teeth move. Leave a distinctive pattern, drawing a bar pattern. These seem to be distinctive to any one tooth. Able to be correlated to account for the marks in 3D.
        - Might be sufficient data there within certain parameters that the bitemark is human, which can lead to an exclusion, not inclusion. Need to substantiate this. Individual barcodes (when teeth drag along skin).
    - If we are not measuring the right thing, then it's pointless
    - We don't know if the information being collected is relevant

Information from Flip Charts

- Agreements
    - 3D scanning
    - Primary source - mouth
    - High accuracy
    - Sharable immediately
- Casts are less accurate
    - Indirect
    - More uncertainty
    - Subject to damage (humidity)
    - A lot of variability depends on user
- Currently sufficiently accurate for needed accuracy
- Study to compare devices
- Use 3D data to assess incisal height impact

Group C

Discussion:

- What are limitations with understanding bitemarks?
    - 2D imaging of skin is good, 3D is not currently as available and easy to work with
    - Preferable to document bitemark as both 2D and 3D
    - There are no good or cheap 3D imaging systems currently
    - We know we can reliably image models and dentition. Do we know we can image a bitemark reliably using a surface that is not reflective?
    - Working on trials and research on the substitution of impression of the skin. Nothing currently usable, even in dermatology, to scan 3D images of the skin. Tried with equipment used to record teeth but skin is different to teeth.
    - There is no imaging technique looking at bruising, even looking at a 3D bitemark.
        - Not many 3D bitemarks are seen in this field, most are contusions. 3D is not the limitation, impression is the limitation.
    - There is no ability yet to use 3D to examine the skin's ability to heal, particularly looking at natural skin differences, such as age and ethnicity

        - African Americans are more difficult to photograph due to the dark pigmentation of their skin
    - Force is a variable, but the skin is also a variable.
        - The measurement of force is what is used in real life research. We are creating an artificial simulation to reproduce a real situation. As a simulation – not as a variable in an analysis.
        - Everyone's skin is different, and skin differs on different areas of the body.
        - The protocol for capturing info from a skin lesion or pattern is the same whether you are alive or dead, but we also have invasive techniques, this is the same on all parts of the body.
- Imaging of the Contusion
    - Several pictures over time of the living are useful to look at contusion. There is no requirement for how old you would want the contusion to be to take images.
    - If the victim is a child, take pictures over a case of three days x 2/day.
        - Others believe documenting the contusion at the time the victim comes in and taking additional photographs weeks and months out is sufficient.
    - Limitation with doing this - give you info on appearance on injury, but what you do with that info, how can you possibly compare dentition to injury over days or weeks?
    - What is the basis/foundation of the protocol? Is it a bite? Human? A certain individual?
- What information should we ideally extract from the bitemark?
    - We want to see class and individual characteristics. You are arriving at a decision if it is a pattern caused by teeth.
    - It doesn't matter what data you collect, you can't reliably interpret data
    - What is sufficient with class characteristics to say its human dentition?
        - How has this been studied? What else could be out there that produces the same pattern?
        - We have no definite numbers of class and individual characteristics to be able to say yes, it is and no it is not a human bitemark. Depending on circumstances, do we take the victim's response that it is a BM into consideration or is this a bias?
        - These are qualitative issues.
    - Are there studies that show the examiners review and produce reliable and accurate results?
        - Are we giving the impression that data has more relevance than need be? We can't discount collection of data at this point.
- Imaging and Measurements
    - Gap identified - skin is a terrible substrate. Need to better understand skin.
    - Need to get into 3D scanning and do measurements digitally. Some subjectivity and inaccuracy in those measurements.
    - Should use some measurement tool that allows metrics that can be reproduced. There is a need for reliability studies. When you measure a contusion in skin, someone might measure a different point on the tooth depending on what point of the tooth you think it is. Need a study to show we can agree.

    - Need studies of variation across examiners.
    - This qualifies as methods that we need inter-rater reliability and intra-rater reliability tests are needed. Identified as a step forward.
- Group C Summary
    - Cannot define how to determine human bitemark
    - Inter and intra rater reliability testing needed
    - Using other techniques for more accuracy for precise measurement of contusion
    - Issue of time and how it affects the bitemark
    - Can only analyze what is present, info we draw from the trauma is how good or present the trauma is or the insufficiency of the tool being used to capture it. Or the knowledge of variability between humans and measurements. Lack of knowledge not just lack of quality of image. Lack of understanding of the trauma inducing agent.

Information from Flip Charts:

- Agreements - Knowledge of Gaps
    - 3D scanning
        - Mouth
        - Or cast
        - Reliable measures with range of scanners
    - Series of 2D exemplars to show length variation
        - CT scan-like
        - Use 3D scan data to show variation that might be in 2D
- Steps to Moving Forward
    - CAI
    - Class and Individual characteristics determine a human bitemark
    - What is sufficient?
    - What else produces similar marks?
    - What is relevant to determine human bitemark?
    - Inter-rate and intra-rater reliability of measurements

2.2.2 Focus Area B - What contributes to the variability in bitemarks from dentition and how can the variability be determined?

Group A

Discussion:

- Variables Affecting Bitemarks
    - SUBSTRATE
    - Location of bite on body
    - Angle of bite
    - Force of bite
    - Skin of the victim
    - Movement of victim/biter
    - Health of victim/medication
    - Time since bite
    - Time bitten down on

    - Clothing vs. no clothing
    - Ante- vs. postmortem
- Knowledge Gaps
    - Location of bite on body
        - We don't understand how area of body impacts bitemarks
    - Angle of bite
        - Need research on how angle affects appearance of bite
    - Force of bite
        - Soft constant pressure vs. sharp sudden pressure
    - Skin of the victim
        - Health of victim/medication affecting skin
    - Movement of victim/biter
        - Reflexes to pain and fear not well accounted for in current research
    - Time since bite
        - Cannot determine age of bitemark without witnessing the bite
    - Time bitten down on
        - Sudden with release vs. biting and "locking jaw"
    - Clothing vs. no clothing
        - Biting through an object
- Challenges/Steps moving forward
    - Need experimentation
        - Identify 9 to 12 variables to limit effort
        - Using matrices to determine how to best design experiments
    - Historical cases review
        - Identify general characteristics and improve experimental design
        - Gives best view of real-world scenarios and impact of variables
        - Not all variables may be recorded in the case notes
        - A lot of potential unknowns
    - Prioritize variables that can be identified in crime scenes
        - "Forensic rules of thumb"

Group B

Discussion:

- Variables Affecting Bitemarks
    - SUBSTRATE
    - Location of bite on body
    - Angle of bite
    - Force of bite
    - Skin of the victim
    - Movement of victim/biter
    - Health of victim/medication
    - Time since bite
    - Time bitten down on
    - Clothing vs. no clothing
    - Ante- vs. postmortem
- Knowledge Gaps:

    - Location of bite on body
        - We don't understand how area of body impacts bitemarks
    - Angle of bite
        - Need research on how angle affects appearance of bite
    - Force of bite
        - Soft constant pressure vs. sharp sudden pressure
    - Skin of the victim
        - Health of victim/medication affecting skin
    - Movement of victim/biter
        - Reflexes to pain and fear not well accounted for in current research
    - Time since bite
    - Time bitten down on
        - Sudden with release vs. biting and "locking jaw"
    - Clothing vs. no clothing
        - Biting through an object
- Challenges/Steps moving forward:
    - Determine which variables interact to help design experiments
        - Will be difficult or impossible to measure all at once
        - Some aspects may not be possible for ethical reasons
    - Fear response to a human biter and sudden unexpected attack
        - Measuring a few at a time that may influence each other is better than none
    - Design better human biting dynamics apparatus for experimentation on substrates
        - Human jaw moves on more than 2 axes
        - Top teeth slide across surfaces when biting
    - Perform more experiments on live people
        - Ethical issues
        - Issues getting willing participants
            - Willing participants will not be wholly representative

Group C

Discussion:

- Variables Affecting Bitemarks
    - SUBSTRATE
    - Location of bite on body
    - Angle of bite
    - Force of bite
    - Skin of the victim
        - Age and medical condition
    - Movement of victim/biter
    - Health of victim/medication
    - Time since bite
    - Time bitten down on
    - Clothing vs. no clothing
    - Ante- vs. postmortem
    - Medication, drugs, alcohol effects

- Knowledge Gaps:
  - Location of bite on body
    - We don't understand how area of body impacts bitemarks
  - Angle of bite
    - Need research on how angle affects appearance of bite
  - Force of bite
    - Soft constant pressure vs. sharp sudden pressure
  - Skin of the victim
    - Health of victim/medication affecting skin
  - Movement of victim/biter
    - Reflexes to pain and fear not well accounted for in current research
    - Posture/position
  - Time since bite
    - Cannot determine age of bitemark without witnessing the bite
  - Time bitten down on
    - Sudden with release vs. biting and "locking jaw"
  - Clothing vs. no clothing
    - Biting through an object
  - Angulation of the bite
  - Curvature of the substrate
  - Duration of bites is not known (e.g., a kid biting another kid)
  - Anatomical location of the bite (e.g., breast, back, arm)
- Challenges/Steps moving forward:
  - Determine what variables are important
    - Including conditions for when variables are significant
    - Case wise significance and statistical significance
  - Use approximations in experimental design
    - Won't be able to recreate real-world scenarios
  - Work to determine what is the relative contribution of each variable
    - Some variables may not be useful
  - Bitemarks are like snowflakes
    - No two are the same even when coming from the same individual
  - Variability of pressure and location of bite alone can make it impossible to reproduce a bitemark perfectly
  - Can skin accurately record bitemarks?
  - Define how we determine that something is a "human bitemark"
  - May not be possible to experimentally recreate actual bitemark dynamics (IRB approval issues with human subjects)
    - We cannot answer these questions – which is why bitemarks should not be done
  - This is why there have not been much research in this area – except for Bushes studies on cadavers

2.2.3 Focus Area C - What data collection techniques are sufficient to collect evidence of pattern injuries on human skin?

Group A

Discussion:

- Data Collection Techniques (We added to the existing list instead of restarting the list to leave time for the gaps and challenges)
  - Everything that has been discussed, but not all are SUFFICIENT
    - Applies to injuries on skin- NOT necessarily for bitemarks
  - Hyper spectral imaging and RAMAN
  - Many parameters for taking a scientific picture
  - Why are we excising (tissue distortion associated with excision)?
    - Bruce Rockwell's paper - why are we ignoring his work?
    - Need more evidence to bring it back in
  - UV literature has NOT been seen in a sequence of studies
    - Is there a rationale? If so, let's write it down
    - Who's collecting the evidence?
    - In the UK, UV photographs cannot be taken on living persons
    - Injury patterns, using the UV allows us to see the features that reveal more information than the normal light
    - You don't know until you see it - Using UV or IR is a way to visualize in a way we can't see normally
  - Let the bruises develop - keep on taking photographs until they are fading and no longer useful
- Knowledge Gaps
  - Employing measures to avoid bias in data collection (suspect vs. victim)
  - Lack of known sources of pattern injuries and ground-truth data (to separate bitemarks from other pattern injuries; lack of data on does the data reflect the teeth that created the injury; lack of skin model to bite dynamics)
  - Haven't systematically tested the methods
    - We don't know the efficacy depending on the method
    - We don't know empirically - seems worthwhile to at least TRY it
  - Image too soon (first few minutes) - primary distortion
  - The moments immediately after the bitemark might be the *best* representation
  - Evidence based practice relies on clinical, expertise, experience
    - But should establish WHY what we do works
- Challenges/Steps to Move Forward
  - Science needs to be reproducible, reliable
  - Many standards are not very specific (e.g., need distance, light, etc.) – collection issues
    - SOP must be based on empirical evidence
  - Basic science to applied science - if data tells us basic science doesn't apply, then don't use it
    - Applied vs. observational
  - What would a reasonable and prudent forensic odontologist practicing safely in their specialty do under similar bitemark case circumstances?

  - Cannot address black box studies because ground-truth data are unable to be collected
    - Mention of 3 papers from Bill Oliver (based on only receiving photographs, MEs did not do as well - then more information helped; contextual bias concerns)
  - Independent review
  - Specialized photo and scanner equipment to take photos of bitemark evidence
  - Universal training with standard protocols and workflows (training on how to take a UV photo)
  - Limited by lack of ground-truth studies (need lots of people in the study, tool marks on skin, motion, animal bites)
  - Machine learning with artificial intelligence could only be used if data sets are available

Group B

Discussion:

  - Data Collection Techniques (We added to the existing list instead of restarting the list to leave time for the gaps and challenges)
  - With photography, serial photograph overtime (AM and PM)
    - Case: asap, 1 week, 3 weeks
    - After 7 days do an UV photo - shows much more distinct characteristics
    - 28 days+ UV invisible
    - Exception to close-up because it distorts- term for magnification, take from a distance and zoom in
  - SUBJECTIVE - has been the issue since 1974 and will continue to be
    - Have tried systems to reduce
    - Subjective is not intrinsically bad
  - The more information we get, the better we are at identifying a bitemark
    - A doctor takes in all the information they can before making a diagnosis
  - Things become clearer at times as times progresses
  - High level of technical expertise required
- Knowledge Gaps
  - We need some kind of skin modeling
    - Simulation of dynamics
  - Does the pattern reflect the teeth? We don't have a way of creating
  - Study: dentists vs. non dentist - are these bitemarks?
    - Students did better
  - Don't have ground truth to do black box experiment
  - We have the measurements
    - First question - can we identify human bitemark from OTHER injuries?
  - Face recognition works on hundreds of thousands of faces
  - How specific are the standards?
- Challenges/Steps to Move Forward
- Access to specialized equipment such as?
- AI Friend: do you have a collection of known bitemarks?
  - Can be scanned into a machine to be run by AI

- How to quantify rarity of a bitemark association?
- Need standards for data collection

Group C

Discussion:

- Data Collection Techniques
  - What kind of features do we want to capture?
  - If you see trace material in the skin, you will want to collect that
  - DNA
    - Depending on jurisdiction, could be crime scene tech, pathologist or anthropologist
  - Photographic Evidence - happens whether or not the victim is alive; immediately is better
    - Of differing types: Color, Alternate light imaging, UV, IR
    - Distance and up-close (keeping anatomical context)
    - Use a scale
    - As a practical challenge, can I use an iPhone?
  - On a live person, what is the time since incident?
    - Many factors affect change
  - Tissue impression or 3D scan of bitemark
    - Usually lose the 3rd dimension fairly quickly on living victims
  - On deceased, can use invasive techniques
    - Tissue excision (ISSUE with distortion upon excision)
    - Microscopy (analysis of tissue thru microscope - trace evidence)
    - Transillumination
    - Histology
  - You can perform metric analysis on the skin, but you don't know if the skin reliably records the actual measurements of the teeth
- Knowledge Gaps
  - Are we able to associate bite marks in the absence of DNA?
  - How reliable are these measurements on skin?
  - There are issues with skin as an impression medium
  - Reliability
    - Measurements
    - Of transfer
  - We don’t understand the pattern transfer
  - What is the level of reliability?
    - You can perform metric analysis on the skin, but you don't know if the skin reliably records the actual measurements of the teeth
  - Looking at the same factors regarding the physical and interdimensional measurements
    - Class characteristics - anything within the pattern that can distinguish between weapon vs. dentition, human vs. non-human, adult vs. juvenile, individual vs. another individual
- Challenges/Steps to Move Forward
  - Differentiate between physically taking the metrics vs. interpreting

- Determine if we can get metrics to determine between pattern injury (bitemark vs. non-bitemark)
- Need to define **minimal** scope
- Minimum technological requirements for data collection

2.3 Science Question 3 - What interpretation strategies (techniques and practices) produce the most accurate and reliable results?

2.3.1 Focus Area A - What defines sufficiency to establish reliability in the association of bitemarks to dentition?

Group A

Discussion:

- What defines sufficiency to establish reliability in the association of bitemarks to dentition?
    - Data defines sufficiency. Each question requires experimentation and data to establish. The limitation is sufficient experimental data.
    - It is sufficient if you can accurately and reliably make an association with any number.
    - We do not have a minimum number to call a pattern injury a bitemark. If you can establish that, you can establish sufficiency.
- Determining a Human Bitemark
    - Other individuals have said that we don't have the science to determine a bitemark yet.
    - Assuming you can determine a bitemark, you need data to show an error rate for accurately reaching the truth
    - If we have the criteria to determine if pattern injury is a bitemark through the number/type of distinct features, then we can answer the Focus Area A questions.
    - If you see a vague pattern injury (smoke rings), and you know that this pattern can be produced by teeth, we don't call it teeth. It is still valuable in looking at something if it may have been used by teeth even if it won't be used for anything
    - Data allows you to know reliability of assessing a bitemark from a pattern. There is a probability that this injury was caused by a bitemark, then we would know something about what that probability is. Need data to back up saying something is probable
- Training and Experience
    - Some say training and experience does not count as data to validate methods/techniques.
    - In radiology, this is called pattern recognition. You can be trained to recognize patterns. If you don't know what you are looking for you do not know how to interpret it. Are we going to be able to say it's a bitemark or not? Should the "no's" and in between be thrown out? Do we have data to support that? External validity and all data get funneled in to treat people in radiology. External validity – what can we determine from these injury patterns. Can we apply it in medicolegal system and have some result that will help society?
    - Summary - we are not optimizing use of pattern recognition, so we do not know the futility of the information we have
    - Pattern recognition is so important in everything we do that it can be applied to the same model of recognizing a skin injury. And then being able to decide if we are close to the diagnosis. If not close, then you don't move forward. People can

be trained to recognize patterns. Not everyone who goes through training programs is confident.
    - Point made that histology is very different from skin or dental x-rays. The skin is not the same as a pathological slide to capture that disease. Pattern recognition is much more difficult.
    - As soon as we start talking about patterns, we run into the study about whether objects translate into patterns when applied to skin. In radiology, you are working with ground truths.
  - There might not always be an absolute diagnosis. There may be thresholds we can and cannot accept. Limitation identified - we may not recognize and leverage pattern data.
  - External validity is something we don't have. Need to have data to establish the validity of what we do is the problem.
    - Can we at least look at factors, attributes, components of what the injury may or may not be? Can we provide a minimum of what was seen before?
  - Bounding what possibility is for the mark you are observing?
  - Some say you can provide a range of possibilities based on your training and knowledge. Others say you can't rely on people's knowledge to make a determination. In a world where we can measure and experiment to see if our experience is valid. Disagreement. We can test and measure. We can't just rely on experience. We can't stay in the dark ages if you only rely on experience.
  - The only way to know if something is a bitemark or not is if you were there to watch it. It happened in the past, what you see is a pattern. Analysis is examining evidence. Comparison is completely different – comparing to cases. Important to understand that there are class and individual characteristics, when you look at injury. Do we meet those characteristics? No point looking at individual characteristics if you can never prove if something is a bitemark. Can say it has class characteristics and enough individual characteristics based on the information available. It is not arbitrary, and it does involve pattern recognition. How to quantify – can we do it? Pragmatic question. Simply not feasible to go out and bite people.
- Mechanisms for Filling the Gap
  - You need to determine class and individual characteristics universally. Applying numerical code to relevance.
  - An example of individuals is if someone is missing a tooth, you would not expect to see a mark with a tooth in that location where it is missing. If you make a model of teeth and drag it against skin, then it leaves marks. If there are drags, you should be able to marry those. Are there sufficient features to make the comparison? Only comparing it to a small group of individuals (suspects)
  - What if you don't know or have any suspects who did it. Ideal for working with small population, but what if we are in the blind?
  - Population dependent. How common characteristics are – we don't have that data and it will change over time.
  - Individual is going away. Departure from class characteristics. "Randomly acquired characteristics" instead of "individual characteristics".

- Another dataset could be is to test all the performance of the dentists and radiologists and odontologists
- Radiologists have a recertification every so many years
- Does the pattern accurately reproduce in skin? If it reproduces in skin, then how good is the dentist and the method? What is the error rate? Review Pretty and Freeman study on error rates.
- Example of error rate study – had one sample that was not a human BM, from that one data point, you know the % of people who incorrectly identified the BM.

Information from Flip Charts:

- Agreements - Knowledge of Gaps
  - Sufficient experimental data to establish (provides external validity)
  - No minimum number of features to include/exclude
  - Minimum number to identify as a bitemark
  - Data to collect trace from a mark (maybe a bitemark)
  - Not optimizing use of pattern recognition software
    - Utility of software
    - Can it be applied
  - Don't recognize/leverage pattern data
    - Helps determine how to move forward and direction
  - Bound the source for marks observed
- Steps to Move Forward
  - Work toward universal agreement on class and individual characteristics
    - To say randomly acquired characteristics
    - Then apply to comparison
      - Population dependent
      - Different person from ideal dentition
  - Test radiologists
  - Published error rate studies

Group B

Discussion:

- What defines sufficiency to establish reliability in the association of bitemarks to dentition?
  - Need population data to look at frequencies before answering these questions. How many distinct features do you need – depends on rarity of the feature in the population. Need to understand rarity of features in the population.
  - Still don't know exactly how teeth and bitemarks matches. Can get tons of measurements of dentition and can get some measurements of bitemarks. We don't know how to correlate or if there is a correlation of the bitemark to dentition.
  - Of all the issues we identified, dentition is the easy part. Hardest is skin and bitemark and all things that affect that. These studies to determine what we don't know; this is the problem until it is done. Just need RESEARCH.
    - Injury bitemark is the weak point. Identify whether or not statistically you can even do the research. Too many variables to do real research – but

identify that. Is it doable, can we get this research done? Needs to be a will to do it.
    - Can look at dentition across a lot of people, can get inter-biter variability with that. Don't know about intra-person variability within the bitemark. Can you determine if one-person bit something 6 times and it was that person each time? Study across different people and how the bite changes for one person.
    - Pretty and Freeman conducted a study and asked board-certified people from their own cases whether it was a bitemark, this gives reliability – can we distinguish what a bitemark is. This study needs to be done and redone multiple times and ways.
        - This study was considered a "construct validity testing", or the group's preferred name "Reliability Study". This means that you are constructing a test, validity is determined by an agreement of people, the experts.
    - There should always be statistical findings to support data in court
- Defining Criteria for Bitemarks
    - Want to know when looking at injury: overall shape, similar to two opposing arches, can you identify where the midline of the arches is, can you see multiple individual marks that were suspected to be made by teeth, and if those elements are not there, I can't say for sure if it is a bitemark or not.
        - Criteria makes sense and could be empirically studied.
    - ABFO could prescribe a requirement, if you call this a bitemark it must be XXXX. ABFO published it and withdrew this in the 80s. Scoring system can be developed to apply to each pattern to use as a guide to get to the next step
        - Define, rerun testing, and then relax the definition based on data. Why hasn't this been redone?
- Mechanisms to Fill Gaps
    - Keeping up with new technology would identify things to know we can't do or haven't looked at before. Stay up to date.
    - Need ground truth studies and statistics.
    - Need data for the jury to know how likely it is that they should accept it just beyond trusting an expert opinion.
    - The issue is moving a discipline that it qualitative to quantitative. Maybe need to start at 0 and start all over.
    - We try to collect statistics when there is something regular enough to study. As PCAST did, can human beings do this task with a degree of accuracy and reliability. Bitemark comes up often enough that we need to do this study. Suggest a black box study. If we can get gold standard data, we would like to see the results of black box study. Need to know ground truth to the cases being answered.
        - This will show you reliability, not accuracy.
    - There are two questions: reliability vs. accuracy. Can always ask reliability and do that study. If you can't establish reliability, you can't establish accuracy.
    - Group consensus of where to start - reliability study.
        - In the courtroom looking at relevancy and reliability. Accuracy is great, but not the immediate focus in court.

Information from Flip Charts

- Agreements - Knowledge of Gaps
    - Don't have delta data for population
        - Don't know rarity of features in population
        - Don't know how to correlate bitemark to dentition
        - Don't know bitemark variation for individual
        - Ground truth not readily available to gather data
- Steps to Move Forward
    - Do reliability then accuracy studies
    - Study dentition as a starting point
    - Generate simulated bitemark in silicon – vary; distorted to begin to evaluate bitemark
    - ABFO prescribe features to apply to pattern as a guide
        - Supported by study to evaluate prescribed criteria
    - Construct validity study / reliability study to dentist to evaluate agreement on findings
        - Black box study
    - Need to limit/ qualify testimony / statements

Group C

Discussion:

- What defines sufficiency to establish reliability in the association of bitemarks to dentition?
    - Depends on the quantity and quality of info you have in the pattern itself (the pattern has limited quantifiable data - the more information you have, you have a better chance of doing a comparison). If you don't have the quality and quantity in the pattern, not much you can do with it. The more info you have from the pattern, you have a better chance of doing a good comparison. You can't define what is sufficient. It is either a good or bad pattern or no pattern. Must have the class and individual characteristics that is consistent with the dentition.
    - You can categorize patterns as good or bad. You are not in a position to determine how many points you need and level of quality you have. You could begin to say that these patterns can meet. Thresholds need to be defined by an error rate.
        - Lack of agreement (have not begun to scope it yet) on what is the minimum threshold on quantitative
    - Fingerprints have class characteristics that you can use to exclude a person. You can do the same thing with a bitemark using class characteristics.
        - That is the biological plausibility argument that has been made. Someone with only 3 teeth cannot make a mark with 9 teeth. All agree on this. How much to bring this in to make that determination.
        - Don't have enough data to get to that point.
        - Yes, it is possible to exclude people, but no data to prove it. At what point do you need an expert? Do not need an expert to match the broken tip of the knife to the knife, no science there.
    - Error rates can help to define the minimum threshold. At what point do you get error rate vs. information is useful and is optimized. When do you need an expert vs. when can you use common sense? Need to establish the error rates - categories

need to be made based on error rates. Need statisticians to help define experiments first: sample size, data, how is data interpreted?
    - Summary - Need statistically relevant structure to studies
- Exclusionary Threshold in Court
    - When we are talking about assessing a bitemark for exclusion, the intent is to create a one tailed test, but a jury hears it as a two tailed test. If not excluded, then jury would think inclusion. Don't know where the right answer is.
    - Need to give number to the jury to weigh the statement. Gives information on what it means to not be excluded, attribute a number to it. Need to give context or just leave evidence out.
    - Probative value is outweighed by effect. Things can be excluded from a legal case.
    - Legal rule uncertainty of probative vs. prejudicial
- Current Use of Bitemarks and Future Tools
    - Bitemarks are being brought in as a part of the police report and sometimes not sent to expert. Tremendous increase in bitemarks of children at daycare centers and juveniles.
    - There is currently no pattern recognition software. Identified the evaluation of pattern matching software and its applicability as a step forward.
    - This would it be an extra tool in toolbox as it would be objective.
        - It was mentioned that this is not likely to happen due to the problem surrounding the ground truth. If you had ground truth, the machine and software would be useful.
    - Pattern recognition software could be used in a cursory incident. Could use an app to use for evidence.
    - Some techniques in machine learning. Can cluster data into groups and measure similarity easily. Machine learning techniques that performs very well. Good and free software out there. Lasso. R – package of statistical routines, machine learning routines. Has a steep learning curves but introductory manual provided. Widely used. Principle component analysis.
    - Need to plug in limitations in the software. Before being able to use the automated software that we still have underlying work we need to do. What aspects of variability are important and how much variation is too much? If we don't have a sense of a flattened bitemark is the most reflective image of the arrangement of teeth you can use to associate a person to a bite then you can start imaging all of these flat skinned bites, then you are not developing a tool that gets us closer to ground truth. Automating it isn't going to help. Need to assess underlying variables to plug into model.
    - Need to define parameters features, or you only need a certain amount of ground truths samples and let the machine determine it. If you get ground truth and have dentists determine characteristics and compare it to data, eventually it can get to a point where the computer could say that, and the data can be validated. Weak point is that good data must go in.
    - An enormous data set is not required
    - There is probably data out there with ground truth. Collect and document that and do a review process. Need to have a criteria during process. Then put into

computer. Using real-life cases. Everyone agrees with features seen. You see arch size, etc. Computer then has range based on the data you give it. Just like automated fingerprint system.
    - There is danger in using actual case work and using that as ground truth data.
    - Some cases in literature use golden standard as ground truth. Gold standard can become ground truth
    - You don't get the variety of real-world conditions from experimental bites. You are really looking for what happens in the real world.

Information from Flip Charts:

- Agreements - Knowledge of Gaps
    - Pattern has limited quantity and quality
        - Class and individual characteristics of dentition
    - Don't know bounds/threshold for agreement on minimum quantity and quality
    - Don't know distortion impact
        - Variability aspects
        - Range – before applying tools studies
        - Automating doesn't help
    - Very limited ground truth data
    - Lack threshold to make decisions
    - Class characteristics suggest a bitemark
- Steps to Move Forward
    - R+D to define agreement on quant and qual to call bitemark
        - Error rate study would be essential
        - Statistically relevant structure study
    - Evaluate pattern matching software and applicability
        - Not machine learning (lack ground truth)
        - Characterize and compare marks
    - Application of data tools to evaluate data, e.g., PCA cluster info
    - Collect and analyze substantive bitemark / ground truth
        - Agreement a bitemark and features
        - Need to ensure a bitemark

2.3.2 Focus Area B - What other data is relevant to bitemark examination and analysis?

Group A

Discussion:

- Other Data for analysis
    - What items would we collect?
        - 3D scans
        - Bite impressions
        - DNA Swabs
    - May want to know what other evidence has been collected around area of bitemark
        - Or other things that may impact the bitemark

            - Swabbed vs. kissed vs. touched
        - Results of other tests not necessary
        - Are there other items in area that may have made the mark
    - Exemplars for a dental line-up with known non-matches
- Knowledge Gaps:
    - Getting information (DNA results) too soon may impart bias on examination
    - Difficult to get effective dental line-ups
        - Database would help
    - Lack of discussion of the human element in dental examinations/analyses
    - Lack of definition on what constitutes a human bitemark
        - Are characteristics consistent with a bitemark?
            - Should be first major question in any analysis/exam
- Challenges/Steps moving forward:
    - Need to outline and determine best way to create a dental line-up
        - Address how to create best known non-matching line-up
            - Allows us to assess how accurate an analysis on bitemarks is
        - Establish common protocol that person performing the analysis does not collect the evidence and is only given the information they need for analysis
            - What are the specifics they need?
                - May vary by case
        - Establish minimum standards to identify a bitemark as a bitemark
    - Maybe compare characteristics instead of labeling something as a bitemark
        - "Descriptive diagnosis" vs. labeling
        - It has the characteristics 'consistent' (not a loved word) with a bitemark impression vs. it is a bitemark
        - "There is sufficient data to move forward with examination"

Group B

Discussion:

- Other Data for analysis
    - Want to limit bias in analysis and comparisons
        - Only collect the pertinent information
        - The analyst making comparisons should not know anything about suspect(s)
    - Blind/randomized comparisons
        - Dental line-ups
    - DNA and other forensic evidence collected when applicable
        - Not of concern to Forensic Odontologist
    - Information on clothing work by victim can be useful
        - Won't expect anywhere near a perfect bitemark if through a jacket
    - Photos should include:
        - Different lighting techniques
        - Scales
- Knowledge Gaps
    - Difficult to get blind submission of evidence

        - Needs multiple people who interact but only pass off evidence to each other and minimal information
    - Limited or no standardization on what information gets passed and how
    - Selection for dental line-ups are rarely fully blind and variable
        - Many use the same set of dental impressions for line-ups or is selected by police (not best fit)
    - No minimum agreed standard to define what makes a bitemark a bitemark
- Challenges/Steps Moving Forward:
    - Standardize blind submissions of evidence and exemplars
        - Need to outline positions needed in evidence handling
            - Need to outline and standardize what information is passed along and to who
                - Minimize bias in exams
        - Need to standardize outline for picking best exemplars for dental line-ups
            - Best known non-matches
    - General need to validate and create a standard approach to association strategies
        - Identification of a bitemark
        - Comparisons of bitemark evidence

Group C

Discussion:

- Other Data for analysis:
    - Want to get task relevant information for the analysis
        - High quality photos with appropriate scales
    - What other data are relevant to bitemark examination and analysis?
    - 2 issues
        - Victim deceased
            - Want to see the position of victim and other information about the body at the scene
                - Was it moved?
            - Time of death
                - Especially vs. when photo was taken/pattern observed
            - Getting into bias and contextual information
                - Skirting the edge
        - Victim Alive
            - Won't see injury on scene
                - Effects when you see it which can create variability
    - Always try to collect DNA
    - Virtual autopsy data (would be beneficial)
    - Meta-data, sub-population questions, mining comparisons within certain classes
- Knowledge Gaps:
    - No way to determine when a bitemark occurred without seeing it
        - Has led to cases with competing motives stating different timelines
    - Lack of research and funding
    - Lack of leveraging of new technologies in other fields of science
        - Raman

    - Lack of standardization for dental line-ups
        - Selection issues
        - How many samples?
            - Why?
    - Some "Black Box" studies are being mislabeled as actual black box studies when they are not
    - There are no "magic number" of appropriate class characteristics – there need to be some minimum criteria
- Challenges/Steps moving forward:
    - Training of CSU to identify bitemarks to enable better collection and preservation of evidence
    - Performing real Black Box studies
        - Generate and report error rates
    - Leverage technologies not currently used in field for cases
        - Joint research and funding opportunities
        - Virtual autopsy and Raman among others may help analysis
    - Need general agreement on general requirements to identify a mark as a bitemark

2.3.3 Focus Area C - What are the key approaches to take in bitemark analysis that will ensure the comparison is objective and if the dentition is not excluded, the significance of an association is accurately reported?

Group A

Discussion:

- Key Approaches
    - Significance of the data
        - Need to mark what I see as artifact and as impression
    - Excision of tissue further distorts.
    - Anatomical distortion
    - Matter of opinion on what order you look at evidence
    - Need to take photos 'in situ' regarding position
    - Before and beyond the dental lineup: making sure overlays are consistent size-wise
    - In Q vs. K, do analysis of Q first, then do K independently, then compare
        - Helps with objectiveness
    - Mitigate extraneous information
    - When we do a procedure less often, we stick to what we know
    - Introduce no new unknowns
    - Independent replication
    - Avoid contextual bias
- Knowledge Gaps
    - Error in lab - have to track back to see where it went wrong. Was it examiner or was it other? Do we have to retrain?
        - We don't want to blame a broad group but rather a single person
    - Not blaming someone if they fell victim to unconscious bias or use of language that was acceptable at the time

    - Field of human factors to consider
    - We need Transparency
    - We're not calibrated
- Challenges/Steps to Move Forward
    - The bitemark cases that first resulted in convictions followed by exonerations should be examined
    - Do you have a mandate for independent verification/second review?
        - There should be a requirement
    - Time/cost of doing a verification
    - Process maps/complex decision trees
        - Could be more transparent
    - Root cause analysis has been requested, but cooperation has been minimal
        - Internal resistance

If we go to root cause analysis - other big industries are required to investigate, would that happen in BM

Group B

Discussion:

- Key Approaches
    - Some research in fingerprint (5 persons casts)
        - Difficult to find foils that are similar enough to the suspect and the criteria for selection of the foils are not well defined
        - Some validity but difficult to apply
    - Closed populations - exclude all but one, then pretty indicative
    - 5 close to suspects, 5 similar to BM, or 5 random?
    - Blind peer review
    - Need to disclose what we can in reports and on the stand
        - Science - it is our job to articulate what we mean
        - Whats intuitive and obvious to one practicioner, is not to another
        - If the bitemark isn't unique, and the suspect dentition isn't unique then you cannot compare
        - When your only conclusions are exclusion/inclusion the Prosecution will try to turn the words
- Knowledge Gaps
    - Reducing the bias
    - Qualitative analysis is subjective by nature
    - Qualification review of examiners
        - Some kind of certification?
    - We should not use the term 'match'
        - Non-exclusion is often given more weight than is meant
    - OSAC study
        - Rough attempt to take qualitative assessments and quantify them
    - Not a bitemark issue, but a forensic science issue
    - Locked into questions that are not all encompassing
    - We don't have gold standard data in Bitemark
    - Contextual information management

- Challenges/Steps to Move Forward
    - Studies
        - Should be studies with judges who have presided over bitemark cases
        - Error rates must be determined
        - Post orthodontic changes
        - Reliability studies
            - Need ground truth data
            - Only seeing if consistent
    - Judges should be doing gatekeeping on expert testimony including BM and deciding on whether or not it will be allowed
    - ABFO: Not STANDARDS but recommendations/guidelines/best practices
        - Standards are broader, need something more defined

Group C

Discussion:

- Key Approaches
    - Examiners should make a determination of the quantity and quality of the data in the Q before looking at K. Is the Q data sufficient for comparison purposes?
    - Dental line-up
        - No standardization
        - 5 models? Why 5? How are the 5 selected? No standards
    - Have you already looked at the bitemark to decide what you need?
    - Suggested by ABFO to have an independent reviewer
        - It should be a technical review - more about QA
    - What is the quality of the information in front of us?
        - Baseline considerations: Scale, adequate photo documentation
- Knowledge Gaps
    - Is the lineup even needed? Research should be done to validate it! - Validation of "dental line-up"
    - Is the data sufficient? How do you determine and validate?
    - Is it capable of being compared?
        - "This is a BM" how accurate is it?
        - 99% accurate? we know we're doing it right
    - No piece of research to support racial differentiation based on dentition
    - From exonerations, concepts from a closed population were found to be not accurate
        - Population may not be properly defined
        - Have a threshold, where we can independently determine quality and value
- Challenges/Steps to Move Forward
    - Biggest challenge - subjectivity
    - Similarity index needed
    - Funding/resources
    - Need an error rate established
    - Establish a threshold "sufficient" data
    - It is a subjective comparison and it always will be

3.0 Moderated Discussions of Science Questions

3.1 Science Question 1 - Are there measurable characteristics or features in human dentition that vary among individuals and are persistent within an individual?

3.1.1 Focus Area A Summary - What measurement method(s) provide the best information for capturing reliable information about the detention?

- 3D scanning is the best measurement to capture dentition information
- Taking measurements and accuracy of measurements are not an issue (standards may exist for capturing dentition information – e.g., Metadata resolution; precision – ADA?)
- 2D applied to 3D dentition is a gap
- Resolution
- Using overlays to compare 2D and 3D is a gap
- Doesn't account for things such as height, chips, incisal height (horizontal height)/above or below the plane)
- Protocol needed for taking measurements
- Ability to determine force or pressure
- Steps to Move Forward
- Need to know measurement variability to define requirement of replicas
- Understand the variability/reliability associated with 3D scanner
- Determine how comparable casts are to in situ 3D scanning
- Characterize differences in 2D vs. 3D

3.1.2 Focus Area B Summary - How do we appropriately collect information to create population databases that can be used for scientific and statistical analysis of human dentition?

- Development of 3D databases of human dentitions (i.e., cadaver dental impressions to avoid IRB)
- Large population studies including statistical consideration
- Present databases can be leveraged to get information: Invisalign 3D, OrthoPics, existing casts dentist's offices, not representative of populations
- Need to agree on what characteristics to collect
- What variables are worth collecting
- Need to secure funding
- When considering ultimate goal of comparisons and variability of bitemarks is the database even necessary?
- Disagreements: a less than ideal database is still better than what we have now
- Cannot be excluded finding does not prevent the need for statistical information gained from database analysis – will not be easy NHANES to do dental scans due to public health mission and extreme cost
- Too many variables in dentition and bitemarks to gain usable information
- Gaps:
    - Money for database collection
    - Understand database design
    - Representative populations

  - Statistical interpretation has to vary between 3 sets of dentition: child, teen, adult
  - What is rate of variability
  - Identify and classify class characteristics
  - Identify and define characteristics
  - Natural vs. post-orthodontic dentition
  - Understand racial differences in dentition
  - Representative population issues
  - Cohort database (longitudinal study)
  - Representative population issues (dental practices are not inclusive)
  - What to do first – database or decision of variables that are important to measure (chicken vs. egg)
- Steps to Move Forward:
  - Follow anthropologists lead on identifying and relating variables of bone structure
  - Design and implement 3D scan and/or model databases
  - Engage other experts/fields to identify process to identify variables to look for in database
  - Identify factors of variability
  - Leverage existing databases for forensic purposes (Invisalign, OrthoPics, molds)
  - Interval dental practitioners for a list of variable features
  - Leverage dental schools for information on those that do not go to formal dentist
  - CDC and NHANES to get dental scans or create NHANES like program

3.1.3 Focus Area C Summary - What are the most probative features/parameters to use and what are the limits associated with each?

- Features/Parameters Identified
  - Physical characteristics (shape, horizontal height, length, width)
  - Arch form/intercanine distance
  - Occlusion (how the teeth come together) – we don't know how people bite (position of their jaw)
  - Range of motion – dynamics of biting (class 1, class 2, class 3 occlusion)
  - Underbite, overbite, etc. (jaw geometry) – cannot be obtained just from the cast alone
  - Pathology/injury
  - Restoration/modification (grinding, cosmetic)
  - Relation between teeth (spacing, presence and absence, alignment, rotation, supernumerary)
  - These features may change over time
  - No racial differences in dentition – according to one paper mentioned by John Morgan
  - If midline is not anatomically possible, then you would exclude?
  - There is such variability in lower jaw movement – and we don't know the mechanism
- Bitemarks are a bit of bruising on a flexible piece of skin
- How many people are including/excluding bitemark evidence in court? Is this worth putting forth further work? (would this be funded? Is it worth it to the justice system?)

- Knowledge gaps: collection and analysis of data regarding the features discussed (frequency, class vs. individual, independent vs. dependent variables, features over time), in creation of a database (private companies with existing 3D proprietary data, privacy and legal/ethic issues)
- Challenges: studies to collect and analyze 3D data, uniform methods to collect, analyze and interpret data, available funding

3.1.4 Open Discussion

- Regarding occlusion: We don't know what position the jaw is in
  - Occlusion can be representative because the position may affect the reflection of the BM on the skin
  - Range of motion/dynamics of jaw
  - Occlusion incorporates aspects of the jaw itself - is this dentition? Or is it something more?
  - Occlusions classed 1, 2 or 3
  - Physiological features before the movement
  - When people bite, the mechanism/dynamic might change from bite to bite
  - Occlusion important in cases, useful consideration but not diagnostic
  - Are we saying we would base an exclusion/non exclusion on occlusion?
  - Should be a FACTOR
  - There is such variability
- Long list on things specifically for dentition - don't all apply to bitemark
  - Placing all this into context: how many people are actually using these factors in cases? What happens if it stops completely?
  - Important to collect all the data for CJ? because there isn't controversy in excluding people, but there is in INCLUSION
- Are there privacy concerns on dentition from deceased?
  - Make an arrangement with ME to collect
  - Deceased are not necessarily representative of the population
  - Not a human subject anymore, but there are ethical issues related to cadaver studies; MEs are not necessarily going to just agree - not as simple as it sounds
  - Some states, against the law to take images of deceased
  - To the extent that dentition issues, they still apply to postmortem identification
- Dentition images could be applied to odontology more broadly

3.2 Science Question 2 - Do bitemarks transfer measurable characteristics of the dentition to the substrate?

3.2.1 Focus Area A Summary - What imaging and measurement methods(s) provide the best information for capturing reliable and reproducible information about the bitemark?

- Limitations (knowledge of gaps)
  - Not capturing as much discriminating information as desired
  - Don't know what information we cannot see on the contusion
  - Does imaging inform dentition variation?
  - What does depth of damage tell us?
  - Important to capturing mark and dentition similarity (3D) – can we image marks with 3D reliably?
  - Is impression the limitation?
  - Limit of time series of mark/contusion – don't know changes over time or impact
  - Injury distortion is relevant – teeth correlation to injury then becomes a challenge
  - Image quality does not make something more reliable
- Does information being collected provide useful/relevant to question
  - Contusion measurements can be improved
- Lack of knowledge that the impact tool has
  - Need to determine what characteristics human bitemark (orthodontics or other impacts? time series valuable?)
- Steps to Move Forward
  - Develop ground truth database to assess contusion variability
  - Variation in upper (drag) and lower teeth (more distinct) marks may determine if it is human (can exclude) – need to a study to assess variability
  - Explore imaging possibilities (filter/extract data)
  - Determine what class and individual characteristics determine a human bitemark (what is sufficient? What is relevant to determine human bitemark? What else produces similar marks?)
  - Study inter-rater and intra-rater reliability of measurements
  - What is the impact of teeth brittleness on tooth loss? (kids that are got teeth later have harder teeth – natural fluoride)
  - There may be information that is not being imaged (Raman could help recover these images?)
  - Important expertise that is missing: organizational behavior
- We don't know enough about distortion
  - Bush studies are useful information was questioned – do not accurate reflect because on how primary distortion is defined
  - Primary distortion is unknown…
  - When skin rebounds, distortions
  - We don't know impact of distortions over time
  - We don't know what is necessarily most representative
- Once you started to build something (for Congress) – you cannot kill it
  - How would we think about it if we were building a plane? What do we think we need before we let every pilot fly it?

- Bitemarks in exclusions to exonerate people – science says it is just as difficult to exclude as include (therefore we cannot say that bitemarks can reliability be used to exclude)
    - There is no evidence to support exclusions – there is no science to support exclusions
    - Using something as an investigative tool would be okay (suspect with two teeth, bitemark with 6 marks)
- Post-orthodontic results in almost perfect dentitions

3.2.2 Focus Area B Summary - What contributes to the variability in bitemarks from the same dentition and how can the variability be determined?

- Variability from the same dentition (same set of teeth biting…)
    - Variables affecting bitemarks
    - Location of bite on body
    - Angle of bite
- Disagreements: disagreements on conclusions of what can be reached on bitemark aging (living vs. dead), may be too many variables to create a controlled experiment
    - Skin vs. other substrates should be in the same discussion
    - May not be possible to experimentally recreate actual bitemark dynamics
    - Can you get IRB approval?
    - David Senn's model is not realistic
    - "Bitemarks are like snowflakes"
- Ease of finding volunteers for bitemark testing?
    - Difficulty in bitemark to bitemark comparison (unknown to unknown vs. bitemark to bitemark)
- Limitations of models – people know they are being bitten and not reacting to being bitten
    - There is no perfect model
    - What the Bushes have done is the most definitive thing we have so far
    - Never figured out how to mathematically model the bitemark (not worth modeling porcine if this cannot be done)
    - All biting is done on a fixed hinge – this is not realistic (Dr. Senn's patterns disappear quickly) – concern with reliability of these studies – caution in looking at any of these studies!
    - By eliminating all of the other variables, and the bite cannot be reproduced, then this whole area is problematic
- We keep combining two issues: (1) is it a bite or not? (2) a comparison – we need to separate these
    - If we cannot identify sometimes if something is a bite or not, then we should not be doing it
    - If you bite 10 times and the results are all different, then should we even be doing this?
    - Exclude or inconclusive (don't know if it is a bite or not) – can these two categories be a minimum common ground?
    - We don't even have the baseline to start the process – is it a bitemark?
- What do we know? Do we have the empirical evidence that we can in the variety of experiences described that we can call something a bite? We don't have a building!

- We don't have perfect models – what do we do now?
- Texas Forensic Science Commission – moratorium on judges on bitemarks; we cannot keep going on without empirical data; can be powerful tool to misdirect
- We have to be courageous enough to say we don't know, and we should do the work first to gain confidence (we are not making widgets here – we are putting people in prison or letting people out)
- Steps to Move Forward
    - Need experimentation (identify 9 to 12 variables to limit effort), historical case review
    - Determine what variables interact to help design experiments, design better biting devices and experiment on living people
    - Use approximations in experimental design
    - Work to determine what is the relative contribution of each variable
    - Embrace new ideas (different perspectives) because if all the answers are known we do not need to have this meeting

3.2.3 Focus Area C Summary - What data collection techniques are sufficient to collect evidence of pattern injuries on human skin

- Collection techniques for pattern injuries on skin
    - Photographic (color, ALI, UV, IR), distance, with scale, over time
    - Tissue impression/3D scan
    - Postmortem tissue excision, transillumination (not commonly done or recommended)
- Knowledge gaps:
    - Reliability of pattern injury transfer to skin (skin as an unreliable impression medium, reliability of measurement)
    - Availability of advanced equipment (scanners, camera)
    - Needs for standardized training and protocols
    - Bias by examiners
    - Lack of data on known sources of pattern injuries
- Challenges:
    - Funding for studies and equipment
    - Examination of current protocols compared to current scientific literature
    - Standardized education and training for forensic odontologists
    - Standardized workflows for data collection
    - Identify metrics to interpret pattern injuries (e.g., human bitemarks)
    - Threshold requirements for technology to collect pattern injury data
    - Ground truth studies needed to establish reliability and repeatability
    - Raman to examine and characterize subsurface injuries

3.2.4 Open Discussion

- Injury information being collected vs. bitemark pattern? We just don't know
- Including things other than bitemarks in studies
- Veracity - being able to predict in a long-term basis (persistence over time)
- Are some teeth more likely to be lost due to wearing, susceptibility, care, etc.?

- May be able to predict - take an impression and predict if they may lose teeth in ten years if you retook a BM
- I don't think its relevant
- Doesn't matter how well we can image injuries as the injury is so distorted - how much relationship is there to the actual teeth
- Something CAN distort, doesn't mean it ALWAYS distorts
    - Shouldn't assume taking a picture makes it 100% accurate
- Structure of conversation useful
    - Building the plane from nothing what do we need before we let every pilot
    - Get stuck in the 'where are we now?"
- Just as difficult to exclude as to include
    - What's the ability of bitemark to exclude?
- Bush studies - distortion creating by biting
- Unknown if primary distortion is distortion at all - maybe it's the most accurate?
- Removing the bush's data - where is the science on exclusion?
    - Don't have it yet, but science that we could do to support exclusion
    - Other things to prove innocence
- Differences are what we are looking for
    - Post ortho work are nearly identical
- IRB - cannot break skin
    - If you're going to make a claim that a specific model
- Cadaver models have been used for 40 years - used for foundation of literature model
- Limitation with current model - know they will be bitten, not pushing away, fighting or movement
- There's no perfect model
- Bush is most definitive
    - Attempts on pigs would be valuable
    - Haven't been able to mathematically model the BM
- How do you deal with variability in forensic evidence?
    - More variability biting 3 different subjects vs. 3 bites on the same substrate
    - Need to address mechanisms of biting in human jaw
- By eliminating other variables, creating a bite on one subject over 4 periods
- Combining two issues
    - Is it a bite or not?
    - Comparison of BM to dentition
- Writing a standard - bite or no bite?
    - Need both sides to agree on that at least
    - Only have two things - exclude or inconclusive
- Talking about writing a standard, but about what we are looking for
- Can we define a BM, what to look for?
- We're doing this backwards
    - Put the building aside
    - We need to say, "what do we know??"
    - Do we have the empirical evidence so we can see something that has a limited variation that we can call a 'bitemark'?
- There's a reason we suggested a moratorium to TX judges

    - We have to have the data to support the whole community's ability
    - Cannot continue to go on as we are WITHOUT the science backing
    - People's lives and their freedoms are on the line
- Wide variety of people of expertise, Bonner's idea of Raman studies dismissed immediately, but I see the value.
    - Embrace ideas and at least see if there is validity
- Imaging the actual tissue damage using RAMAN
    - Once we know it is an actual bitemark can help examine the injury
- Pattern injury on human skin - not necessarily a BM
    - Pattern information vs. injury vs. BM specific
- Pattern injury on live - 'injury' itself doesn't occur after death
- Injury to skin - can we discern patterns - and then in patterns, is it a BM?
- Raises the question if this should be done by forensic paths not odontologists
- See a bruise, do a visualization technique, then see a pattern
    - Don't want to conflate general pattern inferences with BM specific
- Do experiments to include things other than bitemarks
- Visualization techniques are based on the injury
    - Better chance to get funding if focused more broadly on injuries including BM rather than only BM
- Could apply to crash injury emergency medicine – could provide data for other areas beyond forensics
    - Emergency medicine, multiple injuries in the ER - differentiating between accidental or violence
- Not seeing a formalized understanding of what the threshold might be
- There are extreme examples (2 teeth vs. 6 bitemarks)
- Not comparing an injury with every set of teeth in the world – looking for potential outliers
- What is the chance of selecting two people at random with a particular set of teeth?
- We can be deceived by patterns (contextual and cognitive bias)
- We need to find a way to make decisions with rational data
- Was the pattern useful information?
- This isn't about a belief system
- ABFO still allows single arch data
- AAFS talk was not published

3.3 Science Question 3 - What interpretation strategies (techniques and practices) produce the most accurate and reliable results?

3.3.1 Focus Area A Summary - What defines sufficiency to establish reliability in the association of bitemarks to dentition?

- Limitations – Knowledge of Gaps
  - Sufficient experimental data to establish (provides external validity)
  - No minimum number of features to include/exclude
  - Lack minimum number of criteria/features needed to ID as bitemark
  - Pattern has limited quality and quantity to identify class and individual characteristics of dentition
  - Don't know bounds/threshold for agreement on minimum quantity and quality
  - Don't know distortion impact (variability aspects, need to know range of variability, automating doesn't help)
- How to Move Forward
  - Need to work toward universal agreement on class or individual characteristics
  - Published error rate studies
  - Do reliability studies then accuracy studies
  - Study dentition as a starting point
  - Generate simulated bitemarks in silico – vary; distorted to begin to evaluate bitemarks
  - Basis for a study: ABFO prescribes features to apply to pattern as a guide
  - Construct validity study to dentist to evaluate agreement on findings (black box study)
  - Need to limit/qualify testimony/statements
  - R&D to define agreements on quantity and quality
  - Evaluate pattern matching software (clustering of data)
  - Collect known bitemark cases (analyze potential to utilize a secondary ground truth source; needs agreement that those are bitemark features)
  - Machine-vision versus human-vision (what radiologists do)

3.3.2 Focus Area B Summary - What other data is relevant to bitemark examination and analyses?

- Consensus
  - Want to limit bias in analysis and comparisons (only collect the pertinent information, the analyst making comparisons should not know anything about suspect(s))
  - Blind/randomized comparisons
  - General need to validate and create a standard approach to association strategies (identification of a bitemark)
  - Other forensic tests such as DNA and salivary amylase (useful but not pertinent to bitemark exams)
  - Photos should include different lighting techniques and scales
- Disagreements
  - Potential for use of oral microbiomes in comparisons

  - Problem with non-consistent use of blind dental lineups consisting of random but best-known non-matches
  - Use of living victim statements (useful but may impact bias)
  - Knowledge set needed for spectral interpretation from data created by non-traditional spectral techniques (Raman and hyper-spectral techniques)
  - Is a dental lineup needed? (if so, what is the magic number of exemplars in lineup?)
  - What information needs to be submitted (time of death/occurrence, time of photographs, event information, position of body)
  - Use of salivary amylase?
- Gaps
  - Blind submission of exemplar impressions
  - No minimum standards for identification and comparison of bitemark
  - Need for standardized separate collectors to reduce bias
  - 3D scanning on bitemarks (exploring research on virtual autopsy techniques, non-traditional spectroscopy)
  - Quality of forensic documentation
  - Position of body and postmortem movement
  - Time of death vs. first observation of wound pattern
- Steps to Move Forward
  - Blind submission should be used for exemplars (approach should be validated, and standardized, dental lineup should be composed of best-known non-matches)
  - Use ABFO guidelines
  - Need to secure significant money and resources for novel bitemark imaging techniques before forming opinions (joint ventures with other fields to get funding easier)
  - Minimum acceptance criteria
- No information is desired regarding age, ethnicity, gender from the suspect
- Example given of Asian teeth having a shovel shape on upper teeth
- No racial information should be derived from a bitemark
- Age group can be important

3.3.3 Focus Area C Summary - What are the key approaches to take in bitemark analysis that will ensure the comparison is objective and if the dentition is not excluded, the significance of a match is accurately reported?

- Key Comparison Approaches
  - Bitemark analysis is a subjective exam
  - Examiner is independent from evidence collector
  - “dental lineup” utilized
  - Contextual information management to reduce bias
  - Peer review/independent verification (replication)
  - Qualification of examiners – this comment should come out about ABFO certified
  - Use of appropriate data for comparison (do no harm to evidence)
  - Analyze Q (questioned evidence) independently for quality, quantity, and significance of data

    - None of these things move us to "objectivity" - all it does is to open a second subjective box
    - Objective analysis (choosing what to measure is subjective) but subjective interpretation
    - Overlays of Q and K
    - Difference between bitemark analysis and bitemark comparison
    - Potential for distortion – this is a pattern matching exercise not a measurement exercise
    - "Is it a bitemark" is still a subjective analysis?
    - Can we make it objective?
- *There is no evidence that there is significant improvement with ABFO certified examiners!*
- Knowledge gaps
    - Validation of and standard protocols for "dental lineup" – no other disciplines does this
    - Quality and quantity of data threshold needed to support determination of bitemark (scale in photo, both arches, etc.)
    - Reliability studies to establish data consistency
    - Lack of error rates (bitemark vs. non-bitemark, human vs. non-human, etc.)
    - Lack of root cause analysis on cases of error
    - Standards for exam protocols and testimony (include peer review)
    - More detailed/documented decision tree of analysis
    - Judge and juror studies regarding bitemark testimony (what does "cannot exclude" mean)
- Mention of 4 studies on effects of lineups (hair, handwriting, and 2 in fingerprints) – see Glenda Cooper's review article?
- Question on how things are framed – as a scientist, can you make a statement that you call "scientific" in the absence of any error rate?
- Unless you have seen the thing being committed, you cannot know with certainty that the mark is a bitemark
- Steps to Move Forward
    - Funding for studies (peer review/independent validation studies, ground truth studies for error rate, data reliability studies, root cause analysis studies)
    - Establishment of a threshold for "sufficient data"
    - Creation of uniform standards and protocols (e.g., photographs)
    - Establish greater consensus within the field
- In the absence of error rates, the report should state that we do not know if it could be another child

3.3.4 Open Discussion

- Metadata: we aren't saying to include these things, correct?
    - Concern about if this data is reflected in the BM anyways
- Race from teeth
    - Shovel shaped incisors
    - Danger extrapolating backwards from bitemark
    - Why do you need to know race if you can estimate from the teeth?

        - Studies pertaining to race or ethnicity are dated - need modern data
    - No single diagnostic criteria
    - Studies are outdated
    - Make it clear that this isn't accepted truth right now
        - Can be useful for ID not bitemark
- Unidentified perpetrator - it's a suggestion, not an inclusion
- No racial or sex information should be derived from a bitemark - AGREED
    - Age group is still important - Adult could be a crime whereas a child may be non-criminal
- Age of victim, Reason it's important: BMs on a baby might heal faster than on an adult
    - In terms of how to take photographs, etc.
- All of them are good things, none of them move you to objectivity all it does is add a second objective opinion, or open the box
- Argue that the analysis is not subjective, but the interpretation
    - Can make measurements, look at certain features - is objective
- Choosing what to measure is what is subjective
- Overlay is just a shortcut
- Agree these are all smart things - big fan of independent verification - primary examiner gets info they can't help knowing but secondary can
- Issue with terminology
    - Bitemark analysis - is it a bitemark?
    - Bitemark comparison - distinguishing an individual from the bitemark
- Looking for systematic differences that match the pattern
- Even the analysis of "is it a bitemark?" is subjective, no?
- All we can do is discuss the extent of which it fits the characteristics of a BM
- Can we make BM objective?
- They make it better, but not objective, still subjective
- Need to be board certified - don't agree
- If you're a dentist, better than non-dentist?
- What do other scientists in other fields do?
- Minimum standards of training
- GAP: Not sure if qualification of examiners if accurate
- Idea that other disciplines - no lineup, just a question - I'm picking the one that's the most similar and can't exclude the one - this is the wrong way
- No one else does dental lineup
    - Do them sequentially
    - Research in fingerprints - when you have the lineup you have more false positives
    - Four studies - hair, handwriting, 2 fingerprint - show that using a lineup decreases false positive
    - If we were to have a dental lineup with missing tooth, find all the people
    - Database publicly avail (110,000) for missing tooth
- Long list, some are nice to have
    - Lack of error rates - every other discipline has error rates
    - Whether or not as a scientist, can you make a statement in the absence of an error rate

- All of these things continually harp back to the ability to determine if something is or isn't a BM
- Transparency of inclusion/exclusion - our responsibility to be completely transparent

4.0 FINAL MODERATED PANEL DISCUSSION – CONCLUSIONS AND TAKEAWAYS

- How will this be reported out?
    - SNA to CSAFE to NIST: NIST Internal Report will just be published
- Pitfalls of bitemark analysis - would you be amenable to showing a few injuries of Bitemark as refresher
- Baseline understanding to utility of using bitemark today
    - ~100 people, how many a year? How many go to court? What type of case? How many total cases? How many times was there enough evidence? How many bitemarks did you identify?
    - How useful will this be in the future?
    - Go thru ABFO and other agencies to collect data
- In the last 5 years how many cases have you done? 38/100 said they routinely testified to it
- Saw the dead babies and know it has to be done, but also heard the underpinning
    - Do we stop bitemark analysis right now until we have the science?
        - Can we put that to the group to vote?
- Frame the question differently - from everything we've learned
    - High but unknown probability of bitemark vs. non-bitemark
- A group of academics - not unreasonable to make a view as a group on whether or not this is safe
    - To continue to practice bitemark analysis is a public health risk
    - Urgency of the situation
- It is out of scope to take the vote right now - could be debated for days
- You've asked a lot of my time to fix something where learning all I've learned makes me not want to be associated
- NIST has been asked to define the scientific foundations - not about the courts
    - Learn what kinds of answers
- If its unsafe: priorities may be different
- Notion that skin can be replicated
- Dangerous topic - the technology not sufficiently
- Our job is to tell about the reliability
    - We can't answer this question - in agreement
- Disappointed that courts have played such a role in the talks
    - Science not discussed as much as the impression that was given
    - I fully understand the other points of view
    - Uncomfortable with the vote b/c not about viewpoint, but open forum about science and not politics
    - Gave the impression that it was one thing and turning into something different
- Should different thresholds be established for investigative lead value versus information that could be provided in court?
- Is there a scientific basis for these thresholds to be established?
    - From what we know right now? - No
    - I want to see the data that substantiates these thresholds
    - Why do we need data to determine what an "injury" is?
- Can there be a minimum threshold for determining?

    - Injury to skin?
    - Pattern?
    - Human bitemark or not?
    - What are they?
        - Studies (reliability evaluations, error rates,
        - System to identify and it works
- Human bitemark
- Steps where human decision making is taken out
    - To make subjective process more reliable/rigorous
    - Can make the process rigorous with decision thresholds
- One unpublished study does not equate to the entire truth
- All studies are important, we need error rates first
- Are there some aspects of the science that has not been discussed the past two days?
- I want to get resolution - I hope that both sides can sit down because we know that there are issues, but how do we resolve it going forward?
    - First time we're all in the same room - big accomplishment and I don't want it to stop
- "I don't know" "there's no data" "we need the research"
    - What message does this send?
- Dentistry is not a statistical practice, it's an experience-based field
- How do we set minimum standard?
- Require mathematicians and statisticians to think differently
- I know:
    - When someone bites skin, we know there is going to be an alteration
- Hope here is to start to address the "i don't knows"
- Implore everyone to remember that there are people whose lives can be ruined if this type of information is introduced in a way that includes errors and can lead law enforcement in a wrong way, leading to wrongful convictions
- Suspect that there is consensus about the fundamental questions regarding the studies
- Question regarding that the fundamental research needs to be done?
- What are the steps to move forward? What are the recommendations for accomplishing these tasks?
    - Do we need:
        - Common terminology? (individual characteristics)
        - What research/validation is needed? Is research feasible/possible? (line-up studies? Error rates?)
        - Minimum acceptance criteria of characteristics?
        - Minimum standards?
        - Others?
    - What are the priorities?
- Conclusion: Prioritize Fundamental Research Studies:
    - IS THIS A HUMAN BITEMARK?
    - Preregister this - w/o knowing the outcomes - what do we need to make this happen? Sample size, sample type, outcome measures