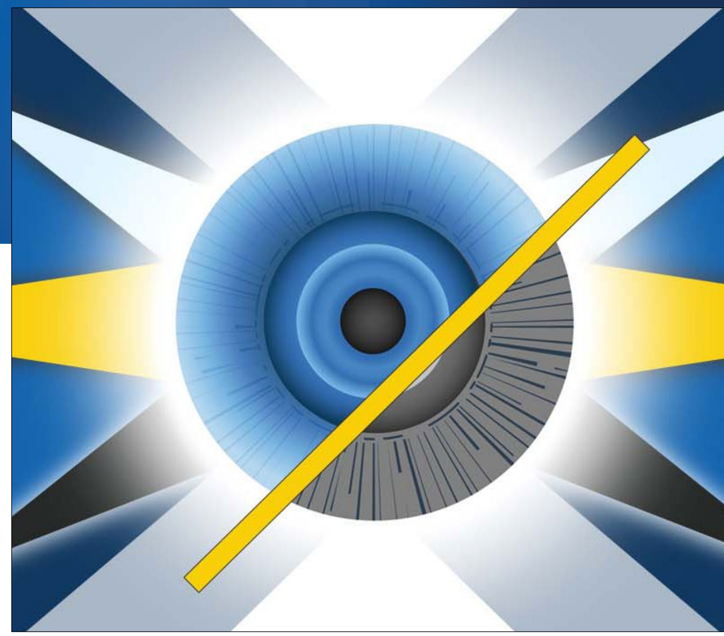
RTI Press

Conference Proceedings

April 2022



2022 Firearm and Toolmarks Policy and Practice Forum

Nicole S. Jones and John Grassel, Editors

RTI International

RTI Press publication CP-0014-2204

RTI International is an independent, nonprofit research organization dedicated to improving the human condition. The RTI Press mission is to disseminate information about RTI research, analytic tools, and technical expertise to a national and international audience. RTI Press publications are peer-reviewed by at least two independent substantive experts and one or more Press editors.

Suggested Citation

Jones, N. S., and Grassel, J. (Eds.). (2022). *2022 Firearm and Toolmarks Policy and Practice Forum.* RTI Press Publication No. CP-0014-2204. https://doi.org/10.3768/rtipress.2022.cp.0014.2204

This publication is part of the RTI Press Conference Proceedings series.

RTI International
3040 East Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
E-mail: rtipress@rti.org
Website: www.rti.org

This forum was funded through a cooperative agreement from the National Institute of Justice (2016-MU-BX-K110), Office of Justice Programs, US Department of Justice. Neither the US Department of Justice nor any of its components operate, control, are responsible for, or necessarily endorse, this forum. The opinions, findings, and conclusions or recommendations expressed in this forum are those of the presenter(s) and do not necessarily reflect those of the Department of Justice. Any products and manufacturers discussed are presented for informational purposes only and do not constitute product approval or endorsement by the US Department of Justice.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204 **www.rti.org/rtipress**

Contents

(continued)

About the Editors

Nicole S. Jones, MS, is the director of the Forensic Science Translation and Implementation Program in the Center for Forensic Sciences (CFS) at RTI International.

John Grassel, MS, is the Manager of the of the Forensic Science Translation and Implementation Program in the Center for Forensic Sciences (CFS) at RTI International.

Contents (continued)

Abstract

The National Institute of Justice (NIJ) and the Forensic Technology Center of Excellence, an NIJ program hosted a four-day symposium, January 11–14, 2022. The symposium included presentations and panel discussions on topics relevant to recent advances in firearm and toolmark examination with a focus on the future. The symposium brought together 685 criminal justice professionals to explore implementation of three-dimensional (3D) imaging technologies, best practices for forensic examination of firearm and toolmark evidence, federal initiatives, gun crime intelligence, black box studies on firearm and toolmark examination, legal challenges to the admissibility of current examination of firearm and toolmark evidence and engineering solutions that will be used in court in the future, implementation of Organization of Scientific Area Committee (OSAC) standards and reporting, uniform language in testimony and conclusion scales. The panel discussions and presentations and provided examples of how agencies implement new imaging technologies for firearms and toolmark examination, incorporate statistics to add weight to forensic comparisons, address legal issues, and operationalize forensic intelligence to improve public safety and share information with the justice community. The symposium also provided a platform to discuss a series of considerations for the forensic, law enforcement, and greater criminal justice community that could help support a successful national transition to incorporate statistics in forensic testimony and accelerate the adoption of imaging technologies for firearm and toolmark examination.

Directors

Jeri D. Ropero-Miller, RTI International
Lucas Zarwell, National Institute of Justice

Emerging Technologies Subcommittee

Wesley Grose
Heather Seubert, Chair
Andy Smith
Jody Wolf
Lucas Zarwell
Xiaoyu Alan Zheng

Intelligence Subcommittee

Hugo Barrera
Tammy Brown
William Demuth II
Basia E. Lopez
Jonathan McGrath
Mallory O'Brien
Stephanie Stoiloff, Chair

Legal Aspects Subcommittee

Jim Agar
David Angel
Amie Ely
Ted Hunt
Erich Smith
Raymond Valerio, Chair

Research Subcommittee

Michael Beddow
Sabrina S. Cillessen
Gregory Dutton
Ryan Lilien
Max Morris
Robert M. Thompson, Chair
Todd J. Weller

RTI International Planning Committee

Jeannine Bratts
Gabby DiEmma
Erica Fornaro
John Grassel
Nicole Jones
Jaclynn McKay
Jeri D. Ropero-Miller

Introduction

On average, guns are used to kill approximately 80 people and wound nearly 300 more every day in the United States.[1] In 2020, the Federal Bureau of Investigation (FBI) uniform crime reporting (UCR) statistics reported that in the United States, firearms were used in 69% of homicides, 41% of robberies, and 30% of aggravated assaults.[2] Law enforcement agencies solved only 22.2% of robberies and 31.3% of aggravated assaults that involved a firearm.[2] Debate about new policies surrounding gun violence rarely includes discussion about the actions that law enforcement agencies are taking under existing laws to reduce violent crimes committed with guns, including implementation of advanced research, use of emerging technologies, and operationalization of gun crime intelligence.

The National Institute of Justice (NIJ), through its Forensic Technology Center of Excellence (FTCoE), organized and delivered a four-day symposium, January 11–14, 2022. The symposium included presentations and panel discussions on topics relevant to recent advances in firearm and toolmark examination with a focus on the future, including the following:

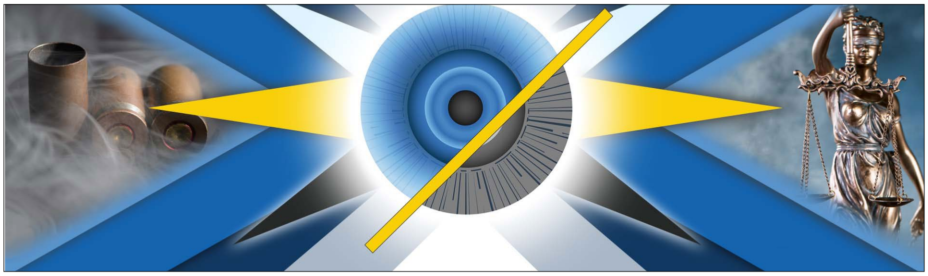
- Implementation of Three-Dimensional (3D) Imaging Technologies
- Best Practices for Forensic Examination of Firearm and Toolmark Evidence
- Federal Initiatives
- Gun Crime Intelligence
- Implementation of Research into Practice
- Black Box Studies on Firearm and Toolmark Examination
- Legal Challenges to the Admissibility of Firearm and Toolmark
- Research and Engineering Solutions That Will See the Courtrooms in the Future
- Implementation of Organization of Scientific Area Committee (OSAC) Standards
- Reporting, Testimony and Conclusion Scales

This global four-day symposium brought together over 685 criminal justice professionals including forensic examiners, laboratory directors, prosecutors, defense attorneys, judges, law enforcement, and researchers.

This report summarizes the panel discussions and presentations and provides examples of how agencies implement new imaging technologies for firearms and toolmark examination; incorporate statistics to add weight to forensic comparisons; address legal issues; and operationalize forensic intelligence to improve public safety and share information with the justice community. The summaries contained in these proceedings were created by the editors and should not be attributed as direct work products of the presenters.

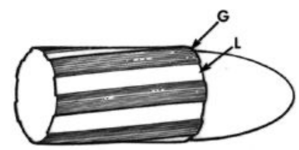
The report also provides a series of considerations for the forensic, law enforcement, and greater criminal justice community that could help support a successful national transition to incorporating statistics in forensic testimony and accelerating the adoption of imaging technologies for firearm and toolmark examination.

Firearm and Toolmarks Examination



Due to subtle manufacturing differences, every firearm produces different microscopic characteristics when it is fired, and these characteristics can change over time based on how the firearm is used and maintained. Firearm identification is the process of analyzing the bullets and cartridge cases left at a crime scene to determine whether they originated from a specific firearm. Characteristics are identified as "class" and "subclass" if they arise from the manufacturing process and "individual" if they result from environmental, post-manufacturing processes. Class characteristics are measurable features that are specific to the rifling specifications of the barrel from which the bullet was fired and are marks those manufacturers intend to imprint to brand their firearms.[3] Class characteristics include caliber, the number of land and grooves, direction of the twists of the lands and grooves and the width of the lands and grooves. Subclass characteristics are also associated with the manufacturing process but are unintentional and may carry across several batches in the manufacturing process. Individual characteristics are factors such as imperfections, corrosion or damage to the barrel which effects the rifling pattern contained in the barrel of the firearm. The uniqueness of these characteristics on the cartridge casings attributed to the firearm and its firing, make it possible to use information from the brass collected at the crime scene to either connect a firearm to the scene or eliminate a firearm from consideration.[3] A firearm is identified if there is "sufficient agreement" for source identification between the marks left on the brass at a crime scene and the marks made by the firearm in question (see Figure 1 for an example). The Association of Firearm & Tool Mark Examiners (AFTE) theory of identification defines "*sufficient agreement*" as the *"agreement of individual characteristics of a quantity and quality that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility."*[3] Although the interpretation of the toolmarks to make an identification or exclusion is subjective in nature and based on the examiner's training and experience, it is founded on scientific principles.

Figure 1. Land and groove pattern on a fired bullet



The rifling pattern in the barrel that fired a particular bullet can be determined by counting and measuring the number of groove (G) or land (L) impressions around the circumference of the bullet.

DAY 1

SCIENTIFIC FOUNDATION AND THE STATE OF FIREARMS AND TOOLMARK RESEARCH

Day 1 focused on the scientific foundation and the state of firearms and toolmark research.

This included a keynote presentation by Dr. Theodore Vorburger on the scientific foundations for firearm examination and related research, based on a study conducted by the National Institute of Standards and Technology (NIST).

This was followed by a panel that was moderated by Robert Thompson of NIST and included the following presentations:

- An overview of the National Institute of Justice (NIJ) research program by Dr. Gregory Dutton from NIJ
- 3D microscopes and their application to firearm and toolmark examination, by Dr. Thomas Brian Reneger from NIST
- Firearm and Toolmark Virtual Comparison Software Technology by Todd Weller from Weller Forensics LLC
- Computer-aided firearm and toolmark analysis by Dr. Johannes (Hans) Soons from NIST

A Review of the Scientific Foundations of Firearm Examination and Related Research

In the 2009 National Academy of Sciences report, it was documented that the process of analyzing toolmarks on bullets is inherently subjective since a firearms examiner makes the final determination of a match, not a computer.[4] Historically, the firearms examiner's opinions have been based on the AFTE Theory of Identification and not based on metrology or statistical Formulas. However, research looking at ways to incorporate more objective measurement into firearm examination continues to evolve. Several entities have called for the need for scientific foundation reviews including the National Research Council (NRC), the National Commission on Forensic Science (NCFS) and President's Council of Advisors on Science and Technology (PCAST).[4-7]

Requests to assess what data exists to support forensic science methods:

"demonstrating the validity of forensic methods"
—Recommendation 3 from the NRC Report (2009)[4]

"technical merit evaluation"
—NCFS Recommendation (2016)[5]

"establishing foundational validity"
—PCAST Report (2016)[6]

In fiscal year 2018, Congress appropriated funding for NIST to conduct a series of literature reviews to analyze and document the evidence base and scientific foundation of the methods and practice of forensic examination. These reviews aim to answer the research question, "What empirical data exist to support the methods that forensic science practitioners use to evaluate evidence?" As of this writing, NIST is conducting foundation studies on DNA mixture interpretation, bitemark analysis, digital evidence, and firearm examination. The foundation study evaluation of firearms examination, which began in October 2019, includes an evaluation of whether selected features can be characterized and are measurable, the extent to which the discriminating power of those features is known, the understanding of the factors that affect transferability of the features, and how well the persistence of the features are understood.[8]

NIST Firearm Examination Foundation Study

Goal: To evaluate the scientific foundations of firearm examinations and reliability of conclusions drawn.

Objective: To answer the question, "What empirical data exist to support or refute the claims and methods that firearm practitioners use to analyze evidence?"

To conduct a full review, NIST evaluated peer-reviewed literature, interlaboratory studies, proficiency tests, laboratory validation studies, position statements, non–peer-reviewed literature and input from the firearms and toolmarks examiner community and expert working groups. When assessing the data, NIST evaluated the retrievability, reliability, and scrutiny/respect of peer reviewers. The scope of the firearm examination included studies of comparison methods; comparison microscopy as applied to both bullets and cartridge casings; statistical approaches; algorithmic comparison methods; and studies that looked at specific regions of interest including tool working surfaces including breech face, firing pin, barrel rifling, chamber, extractor/ejector, and magazine lip, as well as subclass characteristics and manufacturing marks (see Figure 2 for nomenclature of firearm components). The scope did not include nonfirearm toolmark evidence, firearm classification such as the determination of barrel length and caliber, shooting scene reconstruction, gunshot residue, trace metal profiling, or automated investigation methods such as the National Integrated Ballistics Information Network (NIBIN).

Figure 2. Firearm components

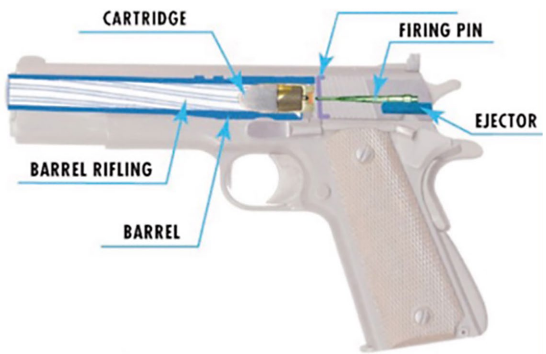


Image Courtesy: Dr. Theodore Vorburger presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

The review sought to study the following claims and subclaims of firearm examination:

Foundational Claim: "A conclusion of common origin between two compared toolmarks or a pair of marks can be made when there is sufficient correspondence of distinctive (or unique) features called individual characteristics and these conclusions are extremely accurate when rendered by a competent or qualified examiner."

Subclaims:

1. The surfaces of firearm parts produced by manufacturing tools are unique. The parts of interest are those that interact as tools on ammunition components.
2. Upon loading and firing, these parts of interest cause marks on the surfaces of ammunition components that are unique to the firearm.
3. Upon loading and firing, the marks on the surfaces of ammunition are also reproducible from one shot to the next and for different ammunition.
4. With normal use, the unique surfaces on the firearm parts are stable over time and over many firings resulting in reproducible, unique marks imparted onto ammunition over time and over many firings.
5. Although there are limits to accuracy due to human performance, to the resolution of optical microscopes and variations in the dynamic process of firing, the error rates are low when conclusions are drawn about common origin.
6. The AFTE theory of identification and its use of the term "practical impossibility" are consistent with measured error rates.[8]

The draft report includes an evaluation of error rates for publicly available data, scientific criticisms noted in previously published reports, consecutively matching striae (CMS), automated algorithms, statical concepts, subclass characteristics, the AFTE theory and foundational claims, and factor space assessments. The factor space assessments use a platform for evaluating strengths and weaknesses of published firearm validity studies, proficiency tests, and error rate studies. The factor space considerations that were assessed in this study included the following:

1. Study classes: independent vs correlated questions
2. Region of interest
3. Questions only versus questions and knowns as related to the availability of known matches
4. Declared versus blind testing
5. Consecutively versus nonconsecutively manufactured firearms and the effects on the potential for subclass characteristics

6. Study size including the number of participants and number of different questions posed
7. Test difficulty and the properties that affect the difficulty including subclass influence, reproducibility, new/different manufacturing methods

To assess these factors, NIST developed a firearms examination validity study reference chart and reviewed firearm studies for their study design, classification, and statistical evaluation of the results, conducting difficulty ratings based on the toolmark quality and the quantity of observations in the assessment. In the PCAST report, the committee evaluated eight studies that were classified into "open" or "closed" study designs of which only the Baldwin study was viewed as a favorable and produced an average false positive error rate of 1.5%.[6] By contrast, the NIST study evaluated 31 studies or reviews and classified them into five design categories using data from 20 of them to estimate the error rate.

The five study design categories included (1) independent questions; (2) compilations of consolidated testing service (CTS) exams, independent questions but correlated sub-questions; (3) batch sorting; (4) matching columns with nonmatches; and (5) matching columns with all matches. The first included studies where examiners were presented with known item and the participant must decide if the question item is from the same firearm or different firearm as the known item and then the next question item in the test is completely independent from the first item. The second is the design used by the CTS exams where each test item is independent but there are correlated sub-questions so for example the question items may not match the known items but multiple question items may be from the same firearm. In batch sorting, the participant is given many question items and the examiner must sort through the batch to determine which items came from the same firearm. This is a challenging and realistic scenario for a firearm examiner, but each decision is correlated with other decisions. In the last two study designs, the examiners are given a list of knowns and a list of question items, and they are challenged to match the items to each other. The difference is that in design 4 there are some question items that will not have a known match in the dataset whereas in study design 5 all question items will match a known and there are no close nonmatches. The false positive error rate was calculated based on the number of false positive identifications divided by the number of nonmatches. Since study design 5 did not have nonmatches, the data in those studies was not used to calculate the observed false positive rate. The studies that were evaluated can be found in Table 1 where asterisks denote whether the study examined bullets only, cartridge cases only, or a mixture of bullets and cartridge cases. The number in parenthesis is either the size of the study, the number of responses to the question items or the number of items sorted for batch sorting studies.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Table 1. Classes of firearm examination studies (interlaboratory comparison, laboratory validations, or proficiency tests)

Independent Questions	Correlated Questions			
	Compilations of CTS proficiency test results	Batch sorting	Matching column with nonmatches	Matching column with all matches
Pauw-Vugts et al., 2013 (637)***,9	Peterson & Markham, 1995 (2,106)***,10	Bunch & Murphy, 2003 (80 Sorted)**,11	Fadul et al., NIJ Report, 2013 (1,650)*,12	Lyons, 2009 (178)**,13
Baldwin et al., 2014(3,268)**,14	Murphy 2010 (11,349)***,15	Defrance & VanArsdale, 2003 (63 Sorted)*,16	J.A. Smith et al. 2020 (1,480)*,17	Fadul et al. AFTE 2011 (3,225) *,18
Walters et al., 2017{Walters, 2017 #5741} (588)***,19	NIST review (unpublished): CTS data from 2014 to 2020 (18,144)***	E. Smith, 2004 (160 Sorted)***,20		Mayland &Tucker 2012 (192)**,21
Keisler et al. 2018 (2,268) **,22		T.P. Smith et al. 2016 (744 Sorted)***,23		Fadul et al. AFTE 2013 (3,255)**,24
Lilien et al. 2019 (1184)**,25		Kerkhoff et al. 2018 (53 Sorted)**,26		Cazes & Goudeau 2013 (544)**,27
Mattijssen et al. 2020 (1,620)**,28		Thompson & Casarez 2020 (≈25 Sorted?)***,29		Stroman 2014 (75)**,30
Mattijssen et al. 2021 (7,008)**,31				Hamby et al. 2009 (7,605)*,32
Chumbley et al., 2021} (8,502)***,33				
Law & Morris 2021 (340)**,34				

Other: Monkres et al., 2013*,35; Hamby et al., 2016**,36; Gutowski, 2005***,37; Wilson Wilde et al., 2017***,38
*Bullets only; **Cartridge cases only; ***Mix of bullets and cartridge cases

The difficulty factors that were examined to assess toolmark quality in the region of interest included the categories and types of examination area such as breech face, firing pin aperture, firing pin, extractor, ejector, chamber, the type of projectile rifling, and the size and surface texture. The difficulty factors that were examined to assess the studies' quantity included the numerical sample size, the number of features within the sample that were provided for comparison, and corresponding measurable features. After evaluating the quality and quantity, an assessment of the proficiency test and validation study design complexity was evaluated to contribute to the overall assessment of the study or test design. This included an evaluation of whether the test or comparative items were presented with few total comparisons containing typically encountered regions of interest. For example, the breech face impressions of two fired cartridge cases are compared with each other to determine if the source was the same firearm as an example of a noncomplex test (Figure 3). An evaluation was also conducted to examine if the test or comparative items were presented as numerous firearm sources or numerous comparison items to determine the source. An example of a complex test would be a test fired cartridge casings that represent numerous firearms that are to be compared with numerous fired cartridge cases or bullets for firearm sources in an open ground truth design. See Figure 3 for examples of such comparisons.

Figure 3. Comparison microscopy of firearm toolmark regions of interest

Comparison of a bullet fragment (left) and a test fired bullet from a suspect firearm (right).

Comparison of breech face detail on two cartridge cases.

Comparison of detail found in firing pin impressions.

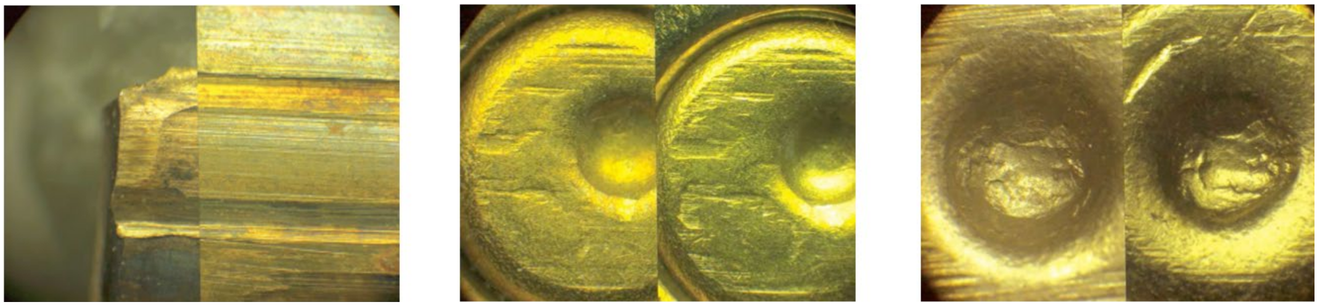


Image courtesy: Micrographs from R.M. Thompson, Firearm Identification in Forensic Science Laboratory (2010), https://www.nist.gov/system/files/documents/forensics/Firearms_identity_NDAAsm.pdf

Other factors that were examined included comparative items that were presented to the examiner in the form of numerous tests for comparison that are received and returned over a multiple month period where tests can be reintroduced later in the study; whether a combination of known and questioned samples that were produced by consecutively manufactured firearm surfaces are presented to examiners to increase the possibility of a false identification; and whether the study provided the examiner with edited digital or virtual 3D images rather than physical samples to reach their conclusions. Using this difficulty rubric, NIST evaluated studies that occurred between 1995 and 2021 for evaluating the difficulty of the study and the error rate of the examiners in the study. Based on the data from the 20 studies that were used to calculate the false positive rate, the observed false positive rate averaged approximately 1.5%.

National Institute of Justice Research Program

NIJ is the research, development, and evaluation agency of the US Department of Justice. NIJ is dedicated to improving knowledge and understanding of crime and justice issues through science. NIJ provides objective and independent knowledge and tools to reduce crime and promote justice, with a focus on state and local jurisdictions. Within NIJ, the Office of Investigative and Forensic Sciences promotes justice through sound science, scientists, and forensic practice. Based on data from the 2014 survey of publicly funded forensic crime laboratories, there are 409 publicly funded laboratories in the United States that received 3.8 million requests for forensic services, of which 4% were for firearms and toolmarks.[39] Of the 409 agencies that participated in the survey, 55% perform firearms and toolmarks examination.[39]

NIJ has three focus areas that include the (1) research and development (R&D) programs, (2) dissemination and technology transition, and (3) coordination of federal partners and the stakeholder community. R&D is important for supporting innovation and the constant need to push forward new capabilities to advance efficiency by using research from the broader sciences for forensic applications. R&D is also important to establish the scientific foundations, accuracy and reliability of forensic methodologies, the development of quantitative methods of comparison and black box studies of examiner performance. Research also factors into the Daubert standards of admissibility and the federal rules of evidence. The NIJ R&D process is a six-step process that includes (1) identifying needs of the field through stakeholder engagement, (2) developing the research agenda, (3) implementing research, (4) conducting post award activities, (5) evaluating research results, and (6) disseminating findings to the field. To identify the needs, the NIJ convenes a group of forensic science partitioners and then publishes those needs to inform researchers. NIJ will also point to research needs from other groups like the Organization of Scientific Area Committees (OSAC) to inform research proposals.

Before 2009, NIJ released discipline-specific solicitations that did not include firearms and toolmarks, but starting in 2009, in response to the NAS report, NIJ broadened their solicitations to include fundamental research in impression and pattern disciplines. Since 2014, NIJ has released two solicitations annually, one that focuses on R&D in forensic science for criminal justice purposes and one that focuses on the research and evaluation for the testing and interpretation of physical evidence in publicly funded forensic laboratories. Since 2009, the NIJ forensic science R&D program has funded 578 projects at a total value of $255 million with an annual average of $21 million to support approximately 48 projects. Descriptions of each project and links to publications can be found on the NIJ website. Approximately 20% of the total funded research projects concern impression and pattern evidence and 6% concern firearms and toolmarks. Since 2006, 48 projects have been funded in firearms and toolmarks research at a total of $16 million, with an average of 3 projects funded per year for a total annual average of $1 million. These research projects are diverse, but generally include research focused on instrument development, standards, quantitative measurement of similarity, weight of evidence, establishing error rates, and virtual comparison microscopy. These have included four continuing lines of research that have had significant contributions to the field. One of the first efforts to apply 3D imagining to firearms and toolmarks was undertaken by Intelligence Automation, Inc., which started in approximately 1997.

NIJ has supported research at Ames Laboratory that examined toolmark measurements of similarity and synthetic toolmarks. NIJ has supported NIST firearms and toolmarks research through what was formerly the Office of Law Enforcement Standards and continues to support projects at NIST through the competitive award process. This has included two awards (2016-DNR-657–1 and 2013-R2R-4843) to support the development of the NIST Ballistics Toolmark Research Database (NBTRD), a database to provide a foundation for applying statistics to source identification of ballistic evidence.

Another project (2016-DNR-6257–2) at NIST involved the development of physical reference standards and standard reference materials for 3D imaging that includes imaging calibration standards to measure step height, roughness, 2D grid arrays and reference flat to ensure that the image measurements are calibrated to acceptable standards regardless of manufacture or model of the instrument used. NIJ has also supported an ongoing line of research at Cadre Research that started with instrument development, validation of virtual comparison microscopy and error rate studies (2012-DN-BX-K058, 2013-R2-CX-K005, 2014-DN-BX-K012, 2015-DN-BX-K032, 2016-DN-BX-0182, 2017-IJ-CX-0024, 2018-DU-BX-0216, 2019-DU-BX-0012, 2020-DQ-BX-0028). NIJ currently has nine active firearms and toolmarks research projects that include three on virtual comparison microscopy, three on error rates or weight of evidence and three on quantifying similarity. The Forensic Technology Center of Excellence (FTCoE), a program of NIJ, provides scientific and technical support to NIJ's R&D efforts by facilitating transfer and adoption of technology into the criminal justice system.

To measure the impact of the NIJ R&D program, NIJ tracks the products that are produced by the research awards as well as the impact measures such as citations, scientific literature, court documents, the number of technologies fielded, workforce supported, and students trained. In 2015, the final grant deliverable was changed to a shorter summary format to allow researchers to focus on publishing in the peer-reviewed literature with an enhanced expectation to generate scholarly products. To measure impact, NIJ tracks bibliometrics of the R&D program including the number of publications and their citations as well as the journals where they are published. Since 2003, NIJ has funded 43 firearms and toolmarks projects that resulted in 47 peer-reviewed articles, 664 citations, 215 conference presentations, and 4 patents. These projects also have an impact on the workforce by supporting 28 unique principal investigators, 149 researchers, 22 undergraduate students, 62 graduate students, and 2 postdoctoral research associates. Tracking these metrics allows NIJ to determine the median impact of a project as a measure of what projects have had the most impact and to inform future awards. There is a challenge to measuring real-world impact such as technology adoption, court citations, and facilitating information sharing through open access publication because this information is not easily captured or traceable back to basic or applied research.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

3D Microscopes and their Application to Firearm and Toolmark Examination

2D conventional microscopes have been used for firearms and toolmarks comparison for many years. These include optical comparison microscopes, stereo microscopes and microscopes that are attached to searchable database systems like NIBIN. These microscopes produce an optical image that represents the object that is produced from contrast that is a function of slope, shadowing, reflection, optical properties, and the direction that light is introduced to the target object. This gives local height variation indirectly and is sensitive to lighting conditions. While conventional microscopy is easy to use and images can be acquired quickly, the lighting intensity and angle are critical for reproducibility.

Figure 4. Topographical image of a pair of fired bullets

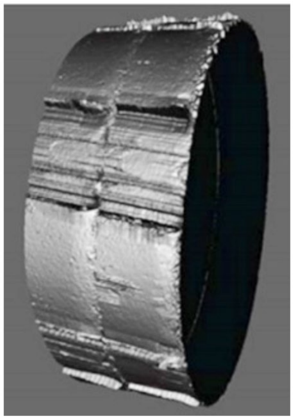


Image Courtesy: National Institute of Justice Forensic Optical Topography Working Group, https://forensiccoe.org/private/5d937bdece0c6.

Due to advances in computer and imaging technologies, optical instruments are now able to produce precise measurements on 3D surface topographies for ballistic identification. 3D optical microscopes can include focus variation, photometric stereo, confocal, and interferometry. The most common of these technologies are confocal and interference microscopes. 3D microscopes measure surface topography and local height variation directly and independent of illumination and shadowing effects which can facilitate objective comparisons but can be more time consuming to measure and subject so signal-to-noise interference and data dropouts (see Figure 4 for an example). These newer instruments have significantly altered the way that crime laboratories conduct firearms analysis and provide better quantitative examination of firearms.

The two general principles for acquiring 3D measurement data are scanning microscopy and sequential illumination. Confocal, focus variation, and interferometric microscopy are all forms of scanning microscopy in which the sample or the objective is scanned as it is moved through the focal plane, and a series of images are taken at different locations. Surface topography is calculated when the images are put together.

In confocal microscopy, the objective is scanned vertically through the focal plane to capture sequential images during the scan. This type of microscopy uses pinhole apertures; only in-focus light is returned to the detector, while out-of-focus light is rejected. The computer calculates the height of each pixel based on the intensity distribution.

Focus variation uses a standard bright-field optical system and scans the object vertically through the focal plane. Using image stacking, the height of each pixel is calculated based on the best sharpness approach that uses a highest contrast algorithm.

The interferometer splits the light between an object path and a reference path and then forms an interference pattern when these two beams of light are superimposed upon each other, based on the differences in the optical paths which creates a fringe pattern. The detector then analyzes this fringe pattern, and the computer calculates the height of each pixel based on the fringe pattern.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Photometric stereo uses sequential illumination consisting of a camera system with multiple light sources where light is passed over the image at different angles and a series of images are captured to produce surface topography. The computer then calculates the 3D surface topography from changes in the light and shadow regions of the object of interest. The position of each light source relative to the camera and the object is measured and calibrated.

Digital imaging systems, including both 2D and 3D microscopy, have resolution limits. Resolution can be improved by using higher powered objectives, but this limits the field of view. One way to maintain high resolution while providing a large lateral measurement range is to use image stitching, which records multiple images of the object of interest and then stiches them together to create a full image at a higher resolution; however, this can increase the acquisition time. One potential issue with measurement stitching is misalignment of the individual images which can cause distortion.

3D optical microscopes need to be calibrated to ensure the accuracy of the measurement results and is a critical component of making sure that the measurements are precise and consistent with other instruments and other laboratories. It is also important to establish metrological traceability to the international system of units of length. Many components of a microscope can be calibrated or checked. These can include the vertical (z) scale amplification factor, lateral (x,y) scale amplification factor, scale linearity, instrument noise, flatness errors, lateral resolution, mapping errors, optical distortions, and image stitching errors.

Quality control is also necessary to ensure that measurements are consistent over time, regardless of environmental conditions or operators and maintain the reproducibility and repeatability of measurements. Control charts are a key diagnostic tool to demonstrate quality control over time. OSAC and the Technical Working Group on 3D Toolmark Technologies (TWG3D2T) are two organizations that are working to develop documentary standards and guidance documents to help with validation, quality control and implementation. Also, the International Organization for Standardization (ISO) has standards for 3D Optical Hardware (ISO 25178) including ISO 25178–600:2019 on Metrological Characteristics for Areal Topography Measuring Methods; 25178–604:2013 on Coherence Scanning Interferometry; ISO 25178–606:2015 on Focus Variation; and ISO 25178–607:2019 on Confocal Microscopy.

Firearm and Toolmark Virtual Comparison Software Technology

Traditionally, forensic laboratories have used light comparison microscopy (LCM) for the comparison of toolmarks left on fired ammunition. In 1925, Cornel Calvin Hooker Goddard adapted medical microscopes to allow for the simultaneous microscopic examination of bullets.[40] Examiners began using this technology in the early 20th century to allow them to to simultaneously compare the striations or impression marks found on two separate bullets or casings. Over the last century, these comparison microscopes have evolved into sophisticated instruments with powerful optics. Comparison microscopy works by optically bridging two microscopes and two objects. Light is shined on the objects to illuminate different surfaces of interest. A mirror then reflects the two individual images back to a single eyepiece so the objects can be compared in the same field of view, divided by a prism line. This allows the examiner to look for similarities for side-by-side comparison and the field of view can be captured through a photograph that is attached to the comparison microscope.

Virtual comparison microscopy (VCM) is the digital comparison of the 3D measured or scanned surface topography of bullets or cartridge casings. First documented by Nicola Senin and colleagues in their 2006 paper in the Journal of Forensic Sciences,[41] VCM incorporates two components: the acquisition subsystem and the analysis subsystem. The acquisition subsystem includes hardware to measure the physical specimen and acquire a 3D surface measurement topography. The analysis subsystem allows the examiner to analyze the features of the surface topography. VCM can be used to digitally compare and archive evidence, allowing for instant access to remote and historic data. An example of this was recent work by NIST and the National Archives where bullets from the John F. Kennedy Sr. assassination were digitally preserved by NIST and then made available through the National Archives to any examiner to be able to download the digital images and analyze the physical evidence. In casework, evidence from one crime can be compared with evidence from a second crime without having to obtain the physical evidence from the submitting law enforcement agency. The VCM software allows the examiner to access multiple digital images for comparison. The images can be rotated to align the areas of interest, and the digital light source can be manipulated to provide better visualization of the areas of interest. The software also allows the examiner to annotate notes and conclusion in the digital image itself as the examiner is conduction the examination increasing productivity and data quality. These features allow for a simplified workflow in which the comparison of the digital images, annotation, and notetaking are all captured by a single software.

VCM can also be used for research to calculate error rates, validation studies, and proficiency tests to analyze examiner conclusions and supporting annotations, as in the 2018 study published by Duez and colleagues.[42] In the Baldwin error rate study using LCM, 218 participants were sent 15 test-sets with 4 samples per test-set.[14] This required the collection, labeling, packaging, and shipment of over 20,000 physical specimens. By comparison, error rate research studies using VCM do not require as much initial effort to obtain and send out the test-sets. For example, the 2020 study published by Chapnick and colleagues included 107 participants that were sent 40 test-sets with 3 specimens per test-set.[43] This study used 120 physical specimens that were digitized and sent to all the participants via the internet. The researchers had the examiners annotate the areas and features of interest in the digital images as they were making their comparisons. This allowed the researchers to use the annotations in the software to analyze examiner consensus regarding the features and areas of interest used to reach conclusions. The researchers then overlaid the annotated images over each other to create the image shown in Figure 5, which provided insight into the examiners' decision-making process and the source of errors.

Figure 5. Annotation heat map of examiner consensus

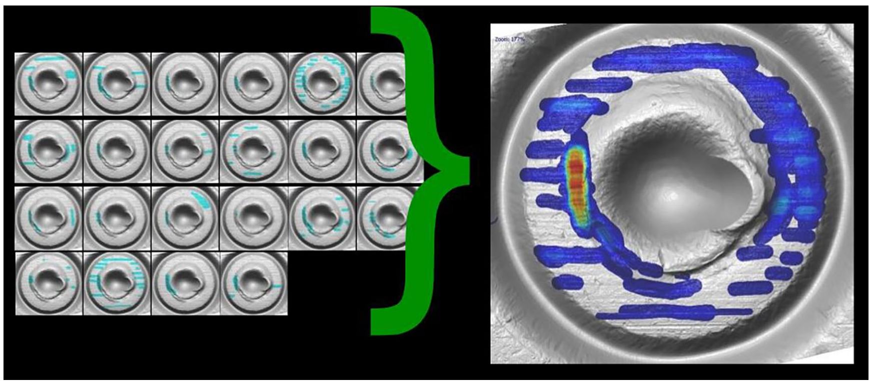


Image courtesy: Todd Weller presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

The annotations can also be used to examine false eliminations as is the case in Figure 6, where the ground truth is that Known A, Known B, and the Unknown were all fired from the same firearm. However, the trainee arrived at a false elimination by identifying a different area of interest from the other examiners in the study. Being able to see this information is important both to understand what areas of interest the examiners are focused on and also to teach trainees which features are not reliable for reaching a conclusion.

Figure 6. Example of examiners annotation on a false elimination

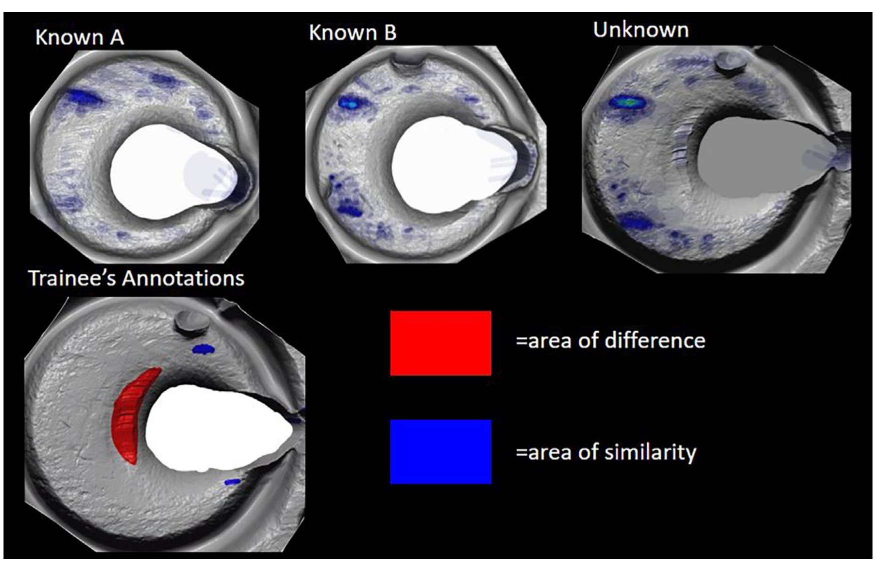


Image courtesy: Todd Weller presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum (adjusted for clarification).

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

The digital files can also be used for future computer-aided verification that uses comparison algorithms as a scoring function and to serve as an independent check to support the examiners' conclusions. With computer-aided verifications aim to determine a statistical score of surface similarity, comparable to a likelihood ratio.[44] A possible workflow for computer-aided verification could include a scenario in which the examiner makes a comparison and arrives at a same source conclusion while the computer independently runs a comparison algorithm to arrive at a confidence score that supports a same source conclusion. If this score was used only in cases of the highest confidence score, it could in theory speed up the analysis process because it would make a secondary verification by another analyst redundant, as the confidence score could support congruence and report the conclusion. In cases where the examiner reaches a same source conclusion but the confidence score produced by the comparison algorithm does not support a same source conclusion, then a second examiner could be brought in for a secondary verification as is typically done in casework. It is important that the features the comparison algorithm uses to calculate the confidence score and how the computer algorithm aligns with the two samples are available to the examiners so that they can interpret the score. This allows the examiner to evaluate the areas of interest that the computer is using to determine if they agree or whether the computer is using subclass characteristics that should not be used for comparison.

In conclusion, the implementation of VCM acquisition and visualization software components for use in casework requires validation and training. Another important consideration for VCM is the quality of the 3D measurements being generated by the 3D optical instruments. Algorithms that are used for statistical analysis require validation, statistical process controls, and routine checks to verify performance. VCM software technology has great potential because it allows the user to integrate comparison notes and documentation into the digital images allowing for instant recall of virtual specimens and insight into the features used to support the examiners' decisions. This feature also allows the firearms and toolmarks examiner community to share digital images for training to expand the knowledge pool. Ultimately, VCM provides a way of measurement that can support the development of comparison algorithms to provide statistical weight to examiner conclusions.

Computer-Aided Firearm and Toolmark Analysis

A key step in firearm and toolmark identification is the comparison of individualizing marks. This is typically done using a comparison microscope. The examiner must assess whether the similarity between marks exceeds the best agreement demonstrated by marks from different tools and whether the similarity is consistent with the agreement demonstrated by marks from the same tool. This process is currently subjective and relies on the skill and expertise of the examiner. As with other types of pattern evidence, the current practice is under scrutiny as documented in the 2009 NAS report and the 2016 PCAST report.[4, 6] Although there may be disagreement on some of the findings in these reports, there seems to be a consensus on two key issues. The first is the need to move from subjective comparison methods and criteria to more objective methods, and the second is the need for statistically sound measures that characterize the reliability or evidentiary strengths of the comparison results.

Addressing these concerns, however, is difficult because there can be significant variability in marks produced by the same firearm, even when shooting one cartridge after the other using the same ammunition. Likewise, there can be similarity between marks produced by different firearms if they were, for example, manufactured using the same tool. Thus, understanding the firearms manufacturing processes is a key part of examiner training. The manufacturing methods of firearms also affects the characteristics of the individualizing features. While firearms examiners look for scratches, peaks, and valleys in surface topography when trying to determine whether two cartridge cases or two bullets were fired from the same firearm, the nature of individualizing these features is not as well defined as, say, that of the minutiae in fingerprints. Finally, there is no consensus on objective comparison metrics, and it is difficult to obtain the amount of data from the number of comparisons required to estimate error rates or other measures of reliability.

To address these challenges, the quality of toolmark data can be improved by measuring the topography of the samples instead of relying on reflectance microscopy images that are significantly affected by reflectivity, focus, and lighting conditions. Using topography enables examiners to measure the samples once and then use the topography data to perform virtual microscopy comparisons on a computer in a manner similar to traditional comparison microscopy. Comparison algorithms are being developed that incorporate a series of computational steps to assess both the level of toolmark geometric similarity (typically expressed by a score) and the degree of certainty that the similarity results from a common origin such as an applicable error rate or likelihood ratio.

The ANSI/ASB Standard 062 for topography comparison software considers three categories of comparison.[45] Software category zero algorithms generate a ranking score, typically using a proprietary algorithm. These algorithms are used for database searching, usually during the investigative phase, and result in a list of samples in the database that are ranked according to similarity between the toolmarks on database samples with submitted samples. Typically, these algorithms do not report a match or exclusion, they simply issue a rank according to the similarity. Since the algorithm used to issue the score is proprietary, the score should not be mentioned in reports that are used during in court procedures; however, the search rank of the data in the database search can be reported with no statistical significance of the rank. Category one algorithms give an interpretable scoring function where the examiner can describe the principle of the score. Therefore, the implied degree of similarity can be reported but a confidence metric cannot be reported with the results. Category two algorithms generate a statistically validated scoring function that establishes statistical confidence for the result. Category two algorithms allow the examiner to report a measure of reliability, such as an error rate or likelihood ratio, based on citable studies that are relevant to the specimens being compared.

The first step in processing evidence using a comparison algorithm is data exchange through the transfer of topography measurement data. Currently, ISO XML 3-D Surface Profile (X3P), which was developed for exchanging topography data, is the preferred data format for exchanging topography measurement data of toolmarks. The X3P format is described in ISO Standard 225178–72 and conforms to both ISO 25178 and ISO 25178–27. ISO developed X3P for many different applications; therefore the Open Forensic Metrology Consortium (OpenFMC)—which is a group of firearm forensics researchers from academia, industry, and government whose aim is to establish file formats, means of data exchange, and best practices for researchers using metrology in the forensic sciences—is currently developing an extension to the X3P data format that describes metadata that are relevant for toolmark analysis. Examples of this metadata include information about the case, the arm, the ammunition, and the region of interest that is being captured.

Once the data are in the computer, image segmentation occurs. In this process, the image is divided up into regions of interest, such as the land engraved area (LEA) on the bullet or a breech face impression on a cartridge case, which are compared separately. Other regions of interest include the firing pin impression, firing pin aperture shear, and the ejector mark. This segmentation step is critical because some regions on the sample, such as the primer roll off, are not informative but may significantly affect similarity results if they are compared. Furthermore, different mark types, such as striations and impressions, require different analysis methods. Also, the relative position and orientation of marks from different firearm components (e.g., the ejector mark or the firing pin impression) may not repeat over time. Therefore, marks cannot be analyzed as one image; rather, multiple images focused on each of these individual marks must be compared.

Figure 7 shows an example of a topography measurement of a land engraved area on the bullet fired from a Ruger P89 Pistol. The measurement data are presented in false color. The yellow areas are high, representing the peaks and the blue areas present valleys. The white areas are areas where no measurement data was collected (e.g., if the slope was too high or there was dirt on the surface). For a striated marks, such as land engraved area, the examiner is interested in the surface topography profile in a plane orthogonal to the striae.

Figure 7. Topographical image of the stria on the land engraved area of a bullet

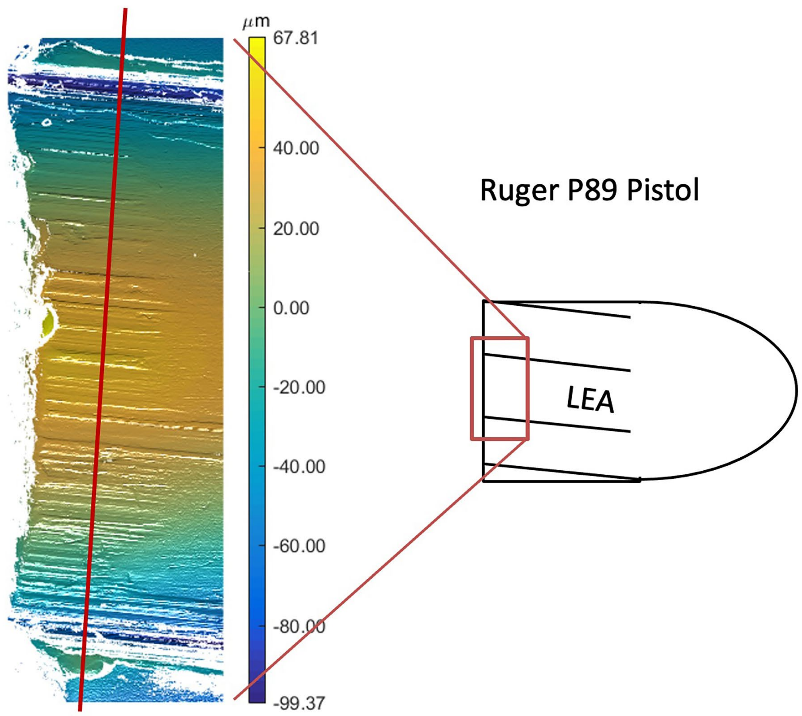


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 8 shows the same data, but the image has been rotated 90 degrees, and an edge detection algorithm has been applied that identifies strong striae. More information about the striation detection can be found in the 2011 paper by Chu and colleagues.[46]

Figure 8. Edge detection to identify strong striae

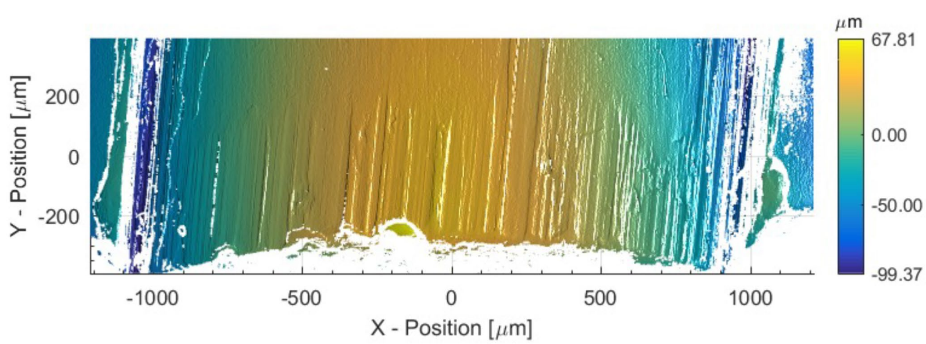


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

The surface heights in each pixel column for those areas that have strong striations are then averaged, and that yields a cross-section profile that is further processed by removing the shoulders at the left and right of the image, as shown in Figure 9.

Figure 9. Cross-section profile of image segmentation of a land engraved area

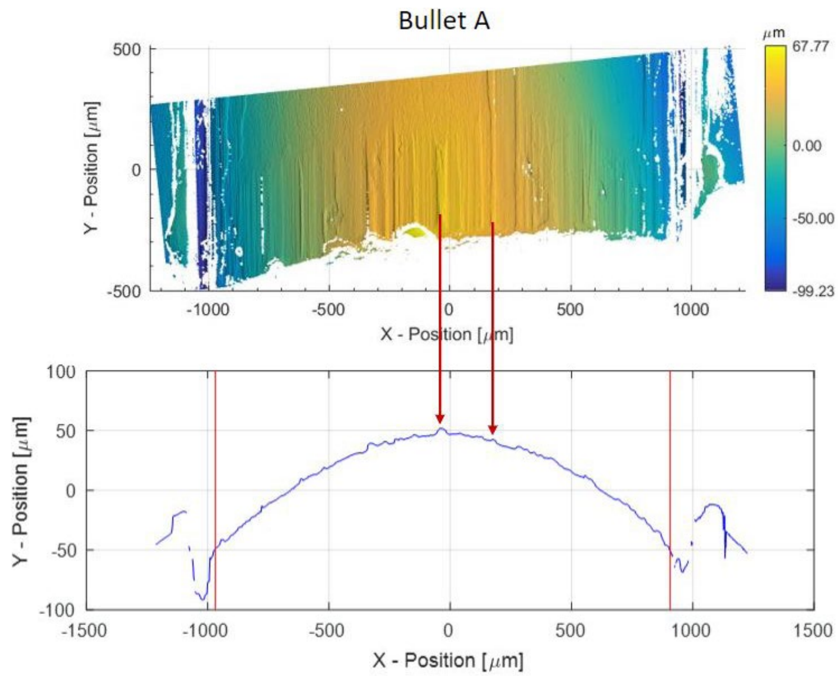


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

In some cases, there may be significant deformation of the striations due to impact of the bullet. In those cases, an algorithm is applied that corrects for the distortion of the striae to accurately determine the cross-section profile. Figure 10 shows an example of this. More information about this can be found in the 2020 paper by Chen and colleagues.[47]

Figure 10. Distortion correction before column averaging

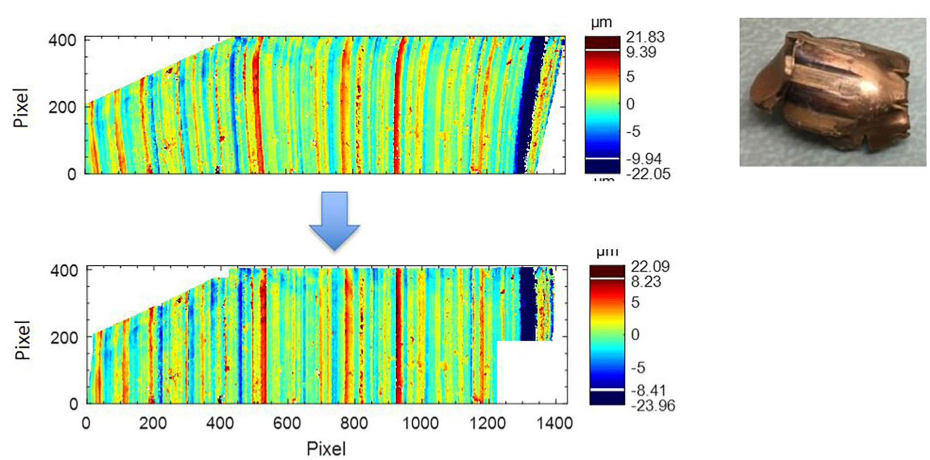


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

After applying the algorithm to correct for distortion on two bullets that were fired from the same firearm, two profiles are created of the land engraved area of the bullets. While these profiles, shown in Figure 11, look very similar, the overall circular shape is not individualizing. The measurement profile or topography data consists of several components ranging from the overall shape or form, which is characterized by very large wavelengths, to noise with very short spatial wavelengths. Other components include waviness and roughness. Therefore, the profile components that are superimposed on that circular shape need to be compared for the individualizing marks.

Figure 11. Image segmentation of the land engraved areas of two bullet fragments

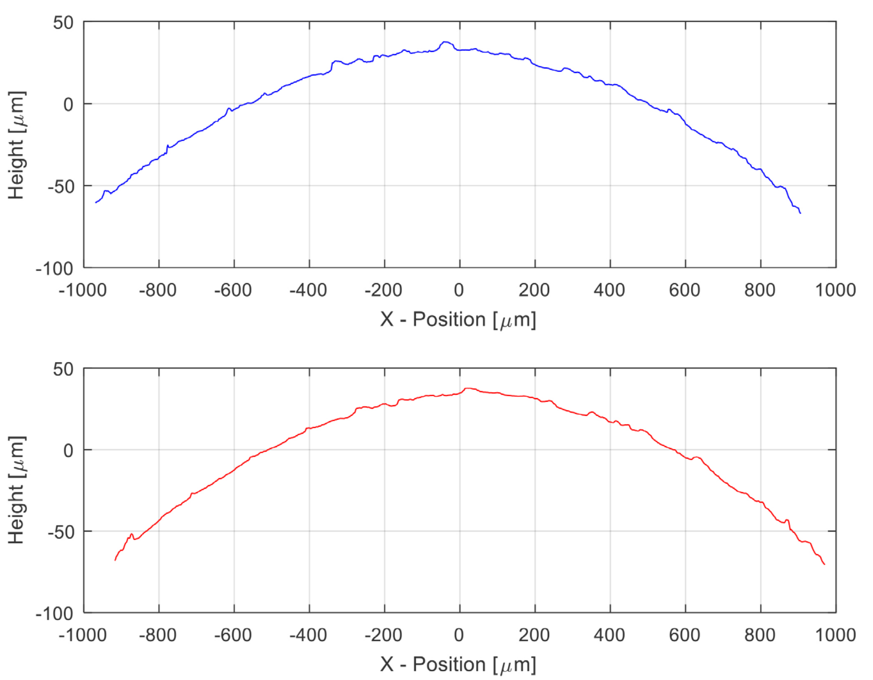


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Repeatable individualizing marks are often found in the roughness region, which is characterized by small to mid-spatial wavelengths. To extract these, a bandpass filter is applied to the measurement profile that only allows frequencies within the range of the roughness profile and attenuates or rejects the frequencies for noise, form, and waviness. The results of this filtering can be seen in Figure 12, where the image on the top shows the roughness of bullet A and the bottom shows the roughness profile of bullet B. While several individualizing features can be seen by the human eye, these features are more difficult for an algorithm to identify because the similar features do not occur at the same position, therefore the image being compared must be rotated so that the individual features on the two images overlap. This process of aligning the peaks and valleys to maximize similarity between the two image profiles is called image registration, and the resulting images appear in Figure 13.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Figure 12. Roughness profiles of bullet A and B resulting from bandpass filtering

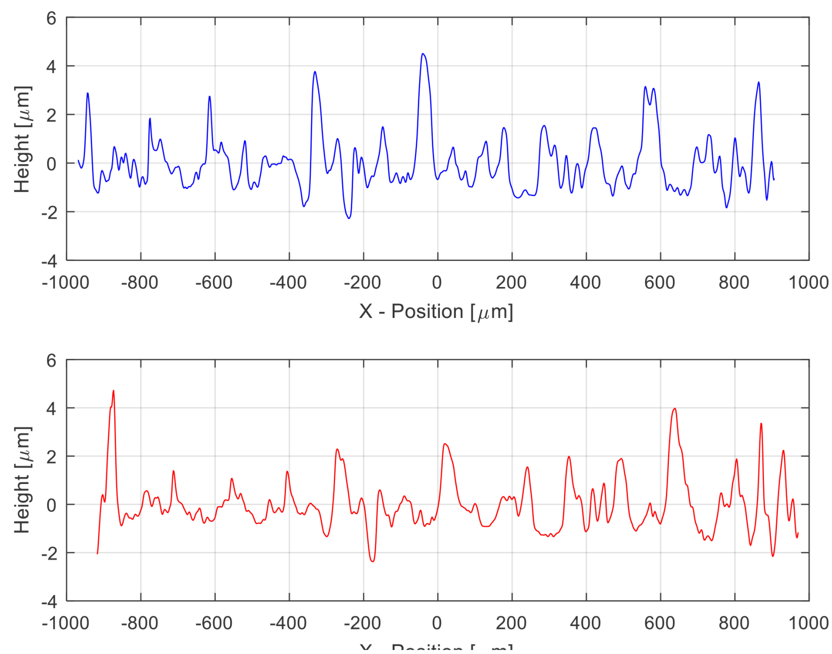


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 13. Image registration to maximize the overlap of individualizing features

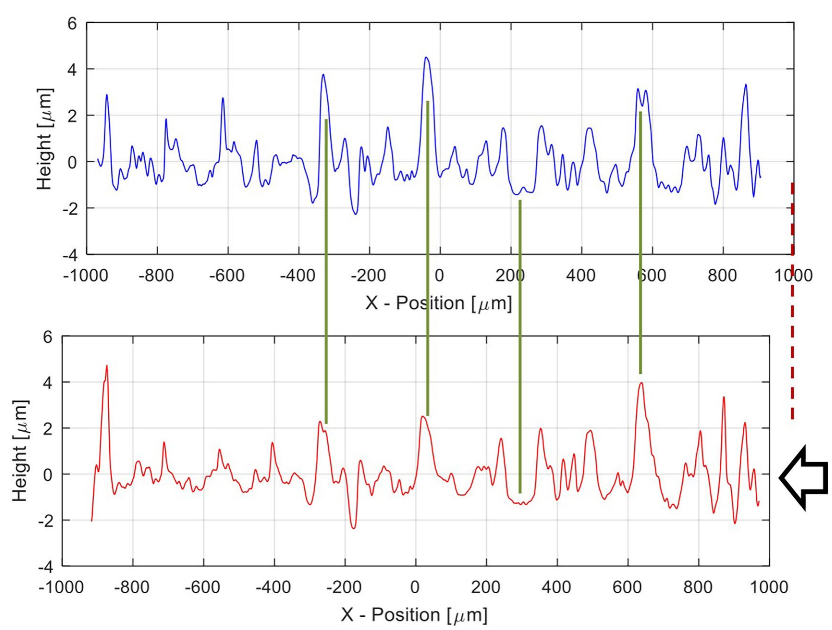


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

There are several methods to measure objective similarity. One involves calculating a feature-based score that is based on the identification and comparison of features. This method works in a similar manner to how minutiae in fingerprints are analyzed. An algorithm would first evaluate the topography data of each surface that will be compared and identify features, which are typically peaks and valleys with a distinct lateral position.[48, 49] The score is the ratio of similar features whose difference in location can be explained by the rotation and translation of the compared surface. The advantage of this feature-based scoring method is that once the features are identified, the evaluation of this number or ratio of similar features can be done extremely fast, making this method well suited for database searches and mimicking how a human examiner would compare two exemplars.

The second approach to calculating an objective similarity metric is to calculate a profile or area similarity score by evaluating the height variations of the image profiles after registration and attenuation of nonindividualizing components. Some of the similarity metrics that can be applied include the mean absolute difference, the root mean square difference of surface heights, or the correlation coefficient. The correlation coefficient is not affected by the mean of each image profile nor the height scale of each image profile. When calculating the correlation coefficient, a value of one means the two profiles are the same, a value of –1 means the two profiles are mirror copies, and a value of 0 means the two profiles are unrelated and there is no correlation between the height variation of the two surfaces.[50]

Figure 14 shows the raw data files of the breech face impression on two cartridge cases that were fired from the same firearm where the yellow indicates peaks and blue indicates valleys. The hole in the center is the firing pin impression where no measurements were obtained. The image data is then trimmed and leveled to the region of interest, as shown in Figure 15.

Figure 14. Same source samples—raw data

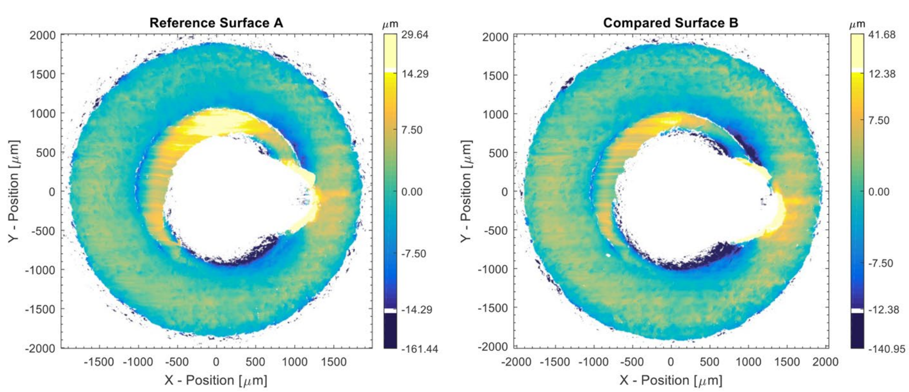


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[51]

Figure 15. Same-source samples—trimmed and leveled

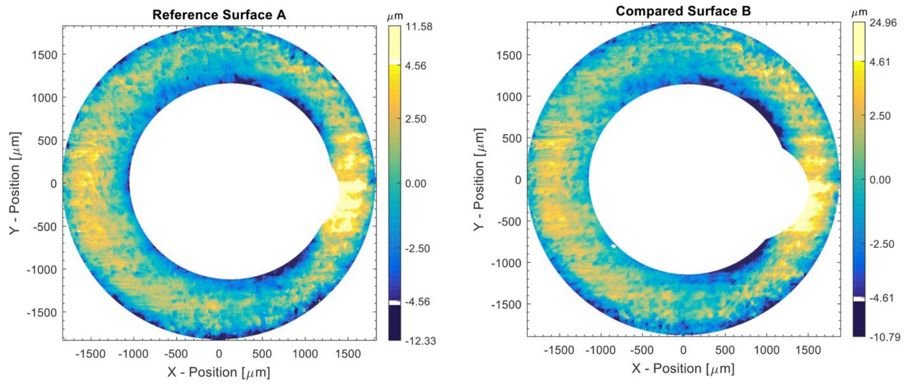


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[51]

A bandpass filter is then applied to highlight the individualizing features, the images are registered, and the comparison image is rotated and translated to maximize the correlation coefficient between the two images shown in Figure 16. In this case, there are visual similarities in both the granular as well as the striated features on the breech face impression, and the calculated correlation coefficient was 63.9%. By comparison, Figure 17 shows different-source samples with a correlation coefficient of 18.7%.

Figure 16. Same-source samples—filtered and registered

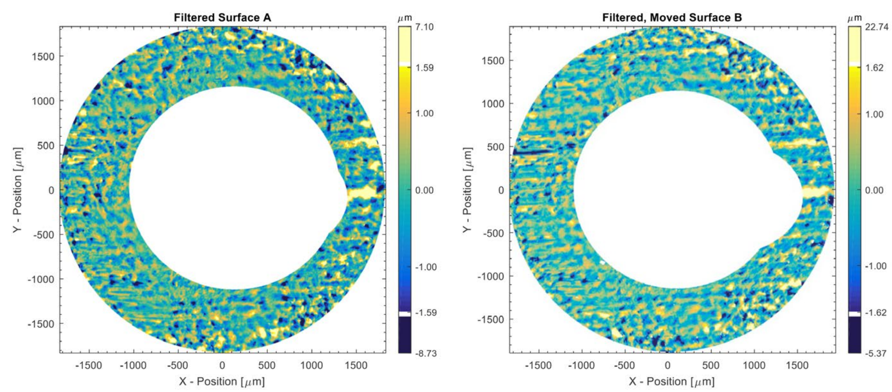


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[51]

Figure 17. Different-source samples—filtered and registered

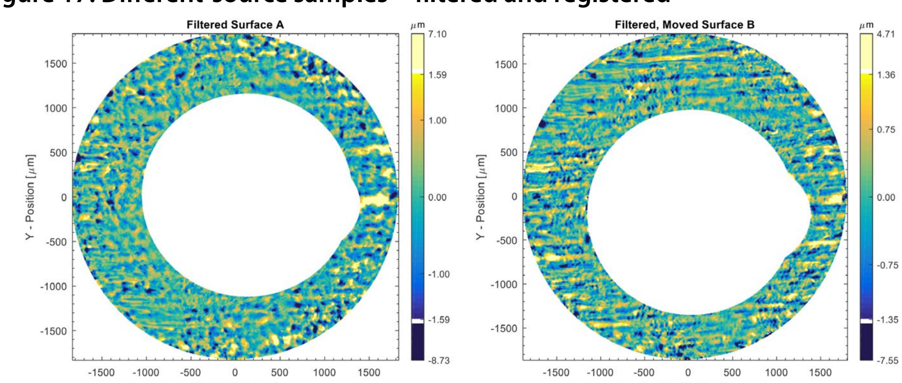


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[51]

The Lighthouse study examined the persistence of subclass characteristics on the breech face of exemplars fired from 10 consecutively manufactured Smith and Wesson Sigma Pistols. In the study, three cartridge cases were fired per firearm resulting in 30 same source comparisons and 405 different source comparisons.[51] The data from this set of comparisons are graphed in Figure 18, where the horizontal axis is the correlation coefficient and on the vertical axis is frequency. The red bars are comparisons of cartridge cases fired from the same firearm, and the blue bars represent cartridge cases fired from different firearms. The similarity score that was obtained in the previous example of 63.9% is consistent with scores obtained for known matching comparisons and significantly higher than that of known nonmatching comparisons. In theory, a threshold line can be postulated between the two distributions and classify a comparison result as a match if similarity score exceeds the threshold value and an inconclusive or exclusion if it is below the threshold value. However, the two distributions are relatively close together so it would be expected that some false positives and some false negatives could occur near the threshold.

Figure 18. Frequency of comparison scores

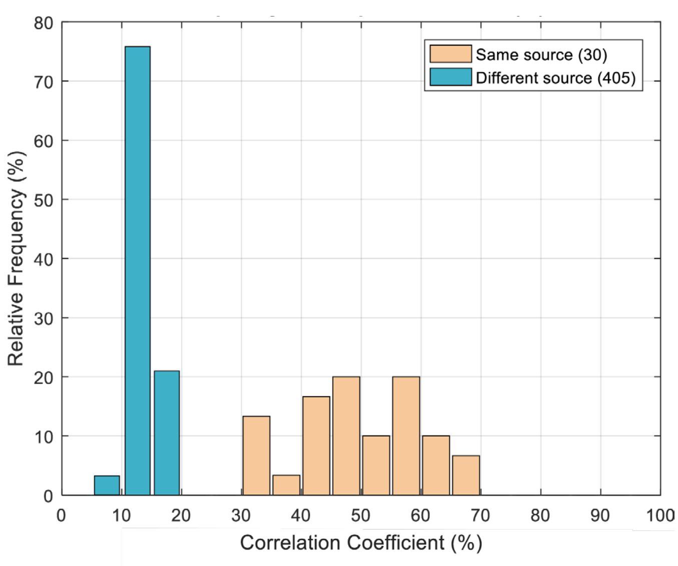


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[51]

Jun-Feng Song, a mechanical engineer at NIST, developed the congruent matching cells (CMC) method, which addresses some of the challenges of area-based comparison methods, in particular the possible presence of areas with poor mark reproduction.[52] This approach combines feature-based methods and area methods by dividing the surfaces into cells. In Figure 19, the reference surface is on the left-hand side labeled as sample A, and each cell scans the compared surface in sample B to find the location where the similarity is the highest.

Figure 19. Example of congruent matching cells

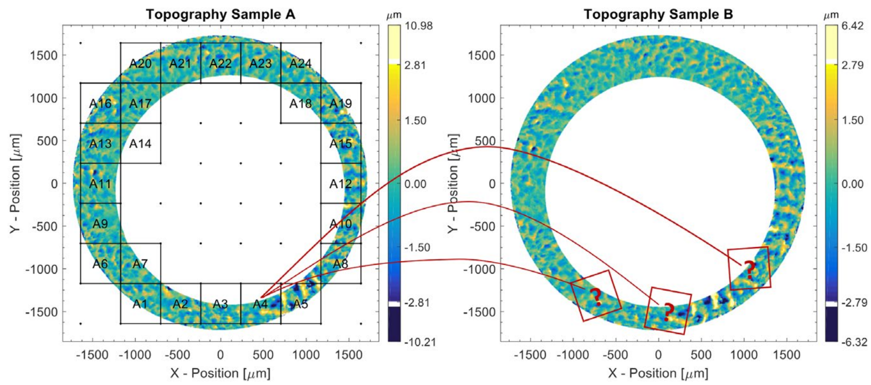


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[52]

Figure 20 shows the comparison of two breech face impressions on cartridge cases fired from the same firearm. The location of the cells on the compared surface is not the same as location of the reference surface due to the rotation of the compared surface.

Figure 20. Comparison of two breech face impressions from the same firearm

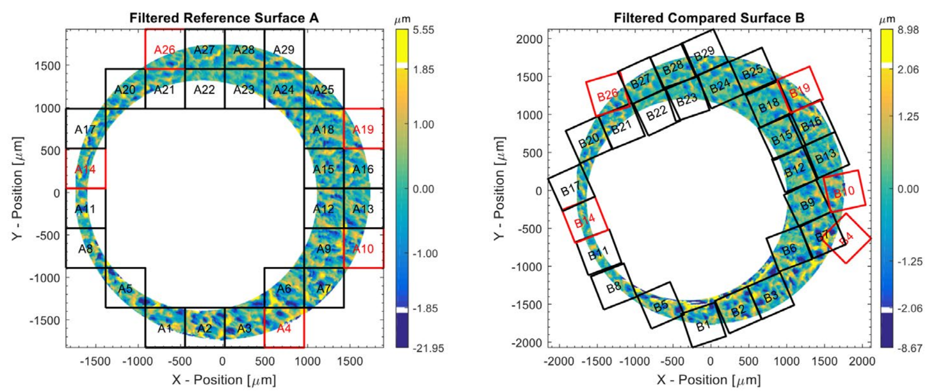


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[52]

If the same comparisons are performed for two samples fired from different firearms, as shown in Figure 21, then cells may find a location on the compared surface that has high similarity but cannot be explained by a rotation or translation of a sample. Therefore, the similarity metric for this method is the number of cell pairs that have both a high similarity and a congruent registration position such that Figure 20 of the same source comparison has 24 CMCs, and Figure 21 with a different source has 0 CMCs.

Figure 21. Comparison of two breech face impressions from different firearms

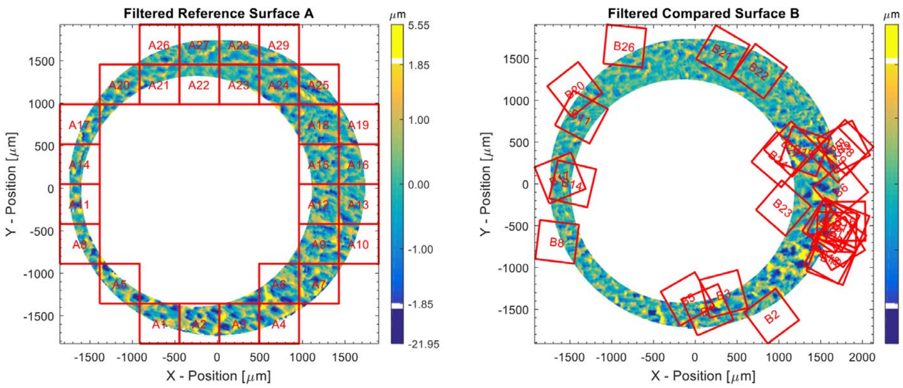


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[52]

When the CMCs method is used for the same set of comparisons from the Lightstone dataset shown in Figure 18, there is now a larger separation between the known matching and the known nonmatching score distributions (Figure 22).

Figure 22. Breech face impression dataset comparison applying CMCs

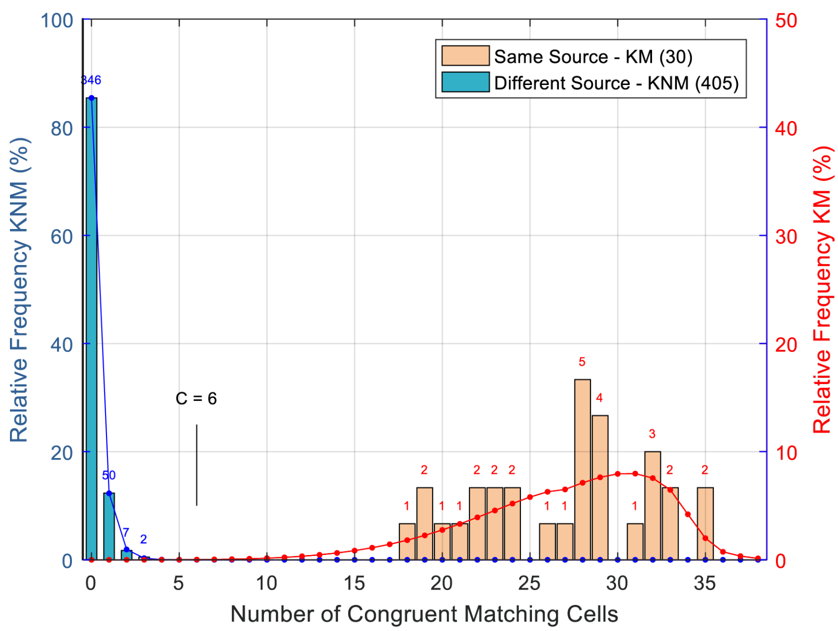


Image courtesy of Dr. Johannes Soons presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.[52]

When this method was applied to additional datasets, some overlap between matching scores in red and nonmatching scores in blue were observed. However, the nonmatching scores were consistently low, indicating low false positive error rates. In these examples, only one region of interest is being compared, and only one scoring method is being used. In practice, an examiner conducts comparisons of different regions of interest such as the firing pin and impression and the breech face impression, and it may be beneficial to combine different scoring methods. The correlation of similarity scores with a declared match or nonmatch gives a clear answer but recent research has investigated presenting the score as a value of the strength of the evidence as opposed with an identification or exclusion. One approach is to calculate a likelihood ratio, a ratio of the likelihood of obtaining the comparison score if the samples were fired from the same firearm versus the likelihood of obtaining the comparison score if the samples were fired from different firearms. In this scenario, the examiner would testify to the strength of the evidence for each scenario or the ratio of the strengths of the evidence for each scenario to the jury instead of a classification of identification or exclusion.

For the last two decades, algorithms have been used for database searches, such as NIBIN searches during the investigative phase. However, these algorithms only provide a ranked list of possible matches based on similarity. Virtual comparison microscopy is currently being introduced in court, and in VCM algorithms have a relatively small, supportive role in improving feature visibility such as highlighting similar areas, but their role is relatively minor. Algorithms for objective identification and characterization of evidentiary strengths are much more difficult and still require more research before they can fully be implemented into the workflow. A key challenge to overcome is determining consensus on the best scoring metrics or processing parameters. In addition, more research is needed to enable translation of human expertise and skill into algorithms, especially regarding the identification of subclass characteristics. These differences in firearm characteristics and ammunition can have a major effect on comparison score distributions and therefore on estimations of the strengths of the evidence. Training is also needed for examiners to know how to express these scoring metrics in courts, and the application of these algorithmic techniques will require careful consideration to avoid confirmation bias.

DAY 2

EMERGING TECHNOLOGIES

Day 2 of the Firearm Policy and Practice Forum focused on emerging technologies for firearms and toolmark examination. This included a keynote presentation by Dr. Ryan Lilien of Cadre on the impact of advances in computer science, statistics, and engineering on the science of firearms and toolmarks examination. This was followed by a panel that was moderated by Heather Seubert of the FBI and included the following presentations:

- An overview of the FTCoE, a program of the National Institute of Justice research program by Dr. Jeri Ropero-Miller of RTI International
- Law enforcement perspective of Evidence IQ Ballistic IQ imaging software system by Lacey Oden from the Escambia County Sheriff's Office
- Ballistics IQ for the crime laboratory by Deion Christophe from the Plano Police Department Firearms Examination Unit and NIBIN Center
- Evofinder®: from validation to implementation by Michael Beddow from the Phoenix Police Department Crime Laboratory
- Reference Population Database of Firearm Toolmarks (RPDFT) by Xiaoyu Alan Zheng from NIST and
- Cadre verification and use in casework by Sabrina Cillessen from the Virginia Department of Public Safety

The Impact of Advances in Computer Science, Statistics, and Engineering on Firearms and Toolmarks Examination

When a firearm and toolmark examiner receives evidence, they examine the discharged bullets and cartridge cases from crime scenes and from suspect firearms to identify toolmarks that have been transferred to those items. These toolmarks can be used by an examiner to determine source conclusions; that is, they can determine if a specific test fire was fired through a specific firearm. Examiners may also reach conclusions regarding the number of firearms present and the makes and models of those firearms. Test fires from recovered firearms or items from a crime scene may be imaged and searched against a database to generate a list of possible hits by comparison to images previously collected from other incidents. These connections can identify putative links between current and past incidents that can serve as investigative leads.

As the cartridge is fired through the firearm, one metal surface scrapes against another, thereby scraping itself. Figures 23 and 24 show examples of the types of comparisons that a firearms and toolmark examiner will make when looking at these microscopic surfaces. Figure 23 shows an example of a known match comparison on a traditional 2D comparison microscope where there are images of two cartridge cases that were fired through the same firearm. The horizontal striation lines on the unknown image on the right appear to be a continuation of the lines on the known image on the left; the lines go straight across, and most of them align.

Figure 23. Known match comparison on a traditional 2D light comparison microscope

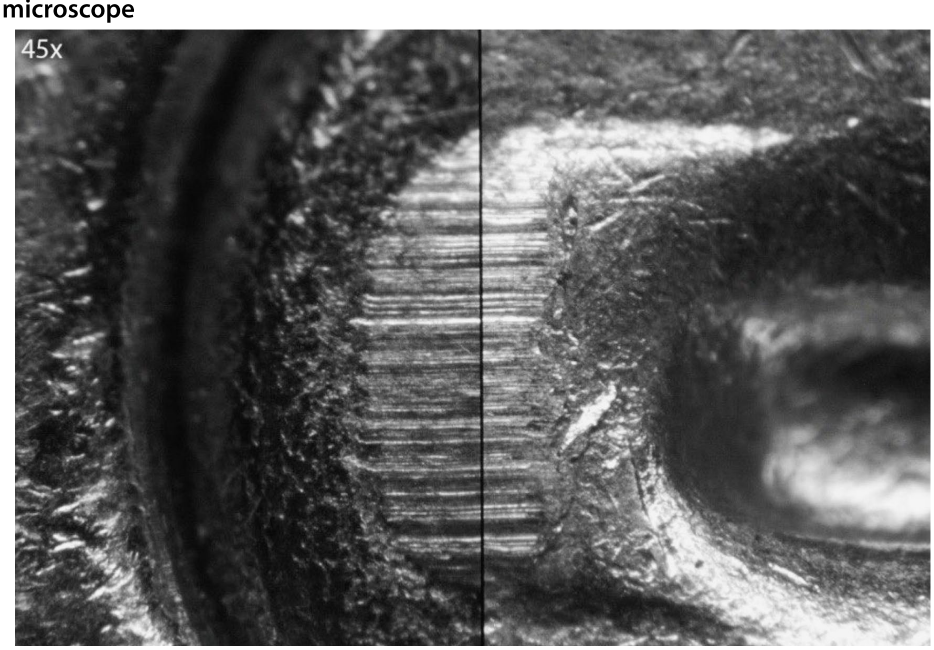


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 24, by comparison is an example of a known nonmatch where the two specimens were discharged through different firearms. Usually there some lines of commonality just by random chance, but most of these striations on the left do not line up with the striations on the right, and therefore this is a nonmatch.

Figure 24. Known nonmatch comparison on a traditional 2D light comparison microscope

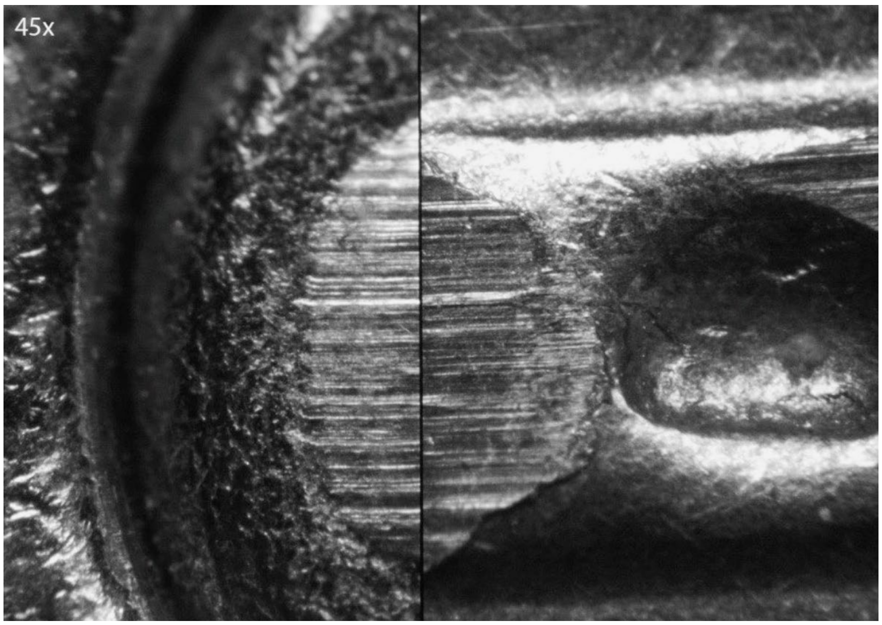


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Around 30 years ago, the FBI worked with a company called MSI to develop the DRUGFIRE systems, one of the first digital networks for automatic 2D comparisons of toolmarks. Around the same time, ATF was working with a company called Forensic Technology to develop the Integrated Ballistics Identification System (IBIS) system, which evolved into the NIBIN system. For the past century, the workflow consisted of an analyst examining microscopic surfaces, then comparing two at a time, and then comparing those images to historic cases to find connections. In the past 30 years, with the efforts of the FBI and ATF, this comparison has become automated.

The term 3D virtual comparison microscopy was first introduced in a paper published in a 2006 issue of the Journal of Forensic Science by Senin and colleagues.[41] This paper describes a virtual comparison microscope as a system comprising two components: an image acquisition component and an image analysis component. The acquisition component uses a 3D microscope to collect measurement data, which are then transferred to a computer where an examiner can view those surfaces, analyze them for comparison of features, and reach a conclusion in the virtual environment much as they would using a traditional microscope. However, instead of examining the physical cartridge case, the firearm examiner is studying 3D measurements of that cartridge case or bullet. Figure 25 shows the 3D scan of the NIST standard cartridge case, Standard Reference Material (SRM) 2461. This image is scanned at about 1.4 µm per pixel.

Figure 25. 3D scan of the NIST SRM 2461

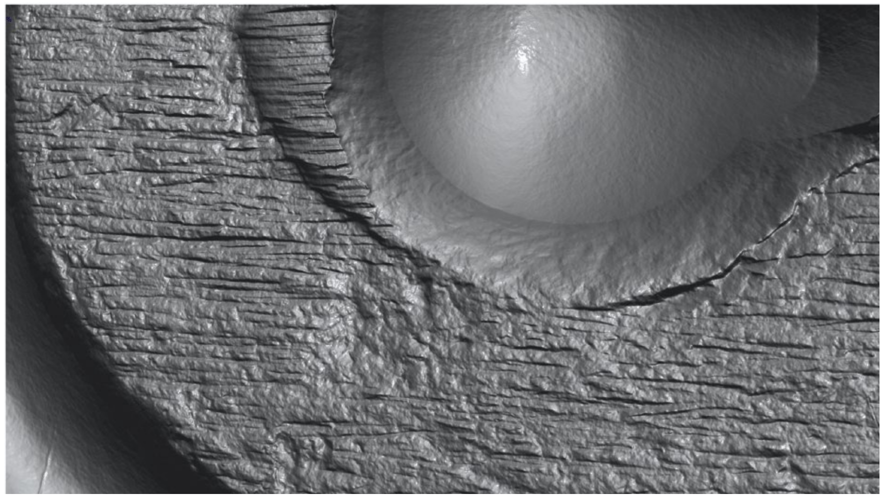


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 26. Known match comparison on a traditional 3D comparison microscope

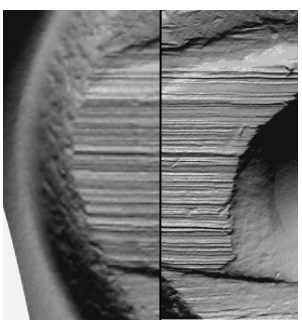


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 26 shows the comparison of aperture sheer toolmarks from two cartridge cases fired through the same firearm. Compared with the impression shown in Figure 23 and 24, the image in Figure 26 is much sharper and shows the marks in a lot more detail and easier visualization.

Some of the strengths of 3D microscopy compared with 2D microscopy is that 3D microscopy high-resolution surface measurements (on the X,Y,Z axis) accurately represent the object of interest independent of lighting conditions. This contrasts with 2D, which is highly dependent on the light because it is does not capture the z measurements. Because 3D measurements are not light dependent, if they are collected in a traceable manner then they can be compared from one microscopy system to another or from one site to another, using a common file exchange format (X3P ISO 25178–72). 3D microscopy also allows the examiner to rotate, translate, zoom, change the z-dimension scaling, as well as adjust lighting, contrast, and brightness to augment the comparison.[42, 43, 49, 53, 54] The precise measurement data that 3D microscopy provides allows the development of statistical models to develop accurate scoring methods that is more difficult with 2D microscopy.

Despite these strengths, 3D microscopy also has some weaknesses such as slower acquisition speeds and the need for specialized equipment. In addition, 3D microscopy generates high-resolution data files (between 20 and 100 megabytes for each single image) that require more storage. These weaknesses are generally simple to overcome and are outweighed by the strengths of 3D microscopy.

Data from these 3D images can be used in several ways, including VCM, in which the visual examination of a digital 3D microscopic representation is conducted virtually, without the need to physically access the specimen. With a virtual microscope, once that scan is acquired, that physical specimen can be returned to the archive, which could be at the originating agency in a different city, county, or state or in a different building. One could also examine the data by using an algorithm to compare specimens using their 3D surface topographies primarily to establish a quantitative measure of similarity.

Some of the potential uses of virtual microscopy include

1. comparing and archiving evidence;
2. training;
3. validation studies and proficiency tests;
4. verification with annotations, blind verifications, technical review; and
5. computer aided verifications with a generated statistical confidence.

Many of these use cases are already in practice in laboratories across the country. When comparing and archiving evidence the examiner has instant access to remote or historic data, and can use this access to reach source conclusions, to write reports or document those conclusions and findings; to perform triage, for example, sorting specimens; to find the number of firearms that might be present; to identify the best exemplars to submit to national databases; to confirm database hits; and to annotate visualizations as part of the documentation process.

Using VCM for training allows the laboratory to have access to the best training samples, even those that are not typical of what an examiner may have locally. For example, a firearms examiner training at one location in the country will have access to all the test fires that are available at that agency. That agency may not have a relatively rare test fire that another laboratory across the country may have and vice versa. With virtual microscopy, if an agency were to scan those images and a database was set up to collect these types of datasets, then any examiner could download that 3D data and compare this relatively rare phenomenon that is not included in local samples. VCM also provides insight to corrective action or corrective training if an error is made during the training process. The Cadre software allows recording surface annotations indicating areas of interest and features used to reach conclusions. Virtual microscopy can also be used for validation studies and proficiency testing. One of the criticisms of many validation studies and proficiency tests is sample-to-sample variance since each examiner who participates in a non-VCM proficiency test or validation study receives a slightly different set of test fires as they are physical items.[42, 43, 55] With virtual microscopy, every examiner examines the exact same files.

VCM can also be used for verifications, blind verifications, and overall technical review. Not all agencies perform verification or blind verifications for all casework. Verification is the process by which a second examiner independently reworks part or all of a case that was worked by the first examiner. The conclusions of both examiners are then compared, and if the independent conclusions agree the first examiner's conclusion is confirmed and if they disagree on their findings, then a third party is brought in to resolve that conflict. This can be done as a blind verification, where the second examiner does not have any of the notations from the first examiner, or it can be done in a nonblind way, where the second examiner has the first examiner's notations. Blind verification can reduce bias; thus, the ability to perform blind verifications on all conclusions is the best scenario. VCM also allows the second examiner to perform the verification at a remote location or another agency. These features could be beneficial for organizations with multiple laboratories. Finally, an emerging use of VCM involves establishing statistical confidence through computer-aided reviews.

Cadre's TopMatch virtual microscopy interface has up to six interactive view panels, allowing the examiner to compare up to six items at the same time. In addition, the examiner can link specimens that are believed to have a common source by color coding them in the software. Figure 27 shows an example where the examiner believes that the three blue scans are from the same firearm and the two red scans are from a second firearm. The examiner can also document their work through surface annotations and written notes.

Figure 27. TopMatch VCM dashboard

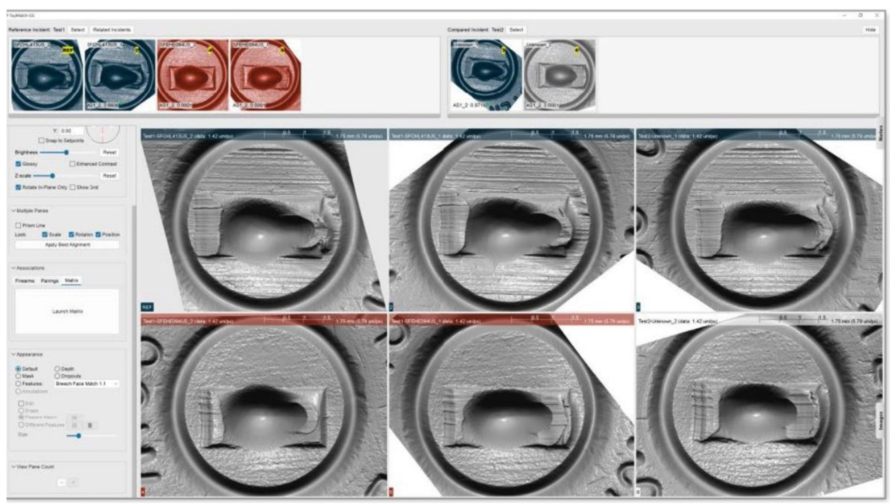


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 28 shows where some areas of the surface have been highlighted by clicking and dragging with the mouse. These areas can be highlighted in different colors to indicate areas of similarity and difference as part of the documentation process. For verifications, all of these annotations can be hidden from the second examiner. This blinding would be much more of a challenge for physical specimens.

Figure 28. TopMatch VCM dashboard annotation of similarity and dissimilarity

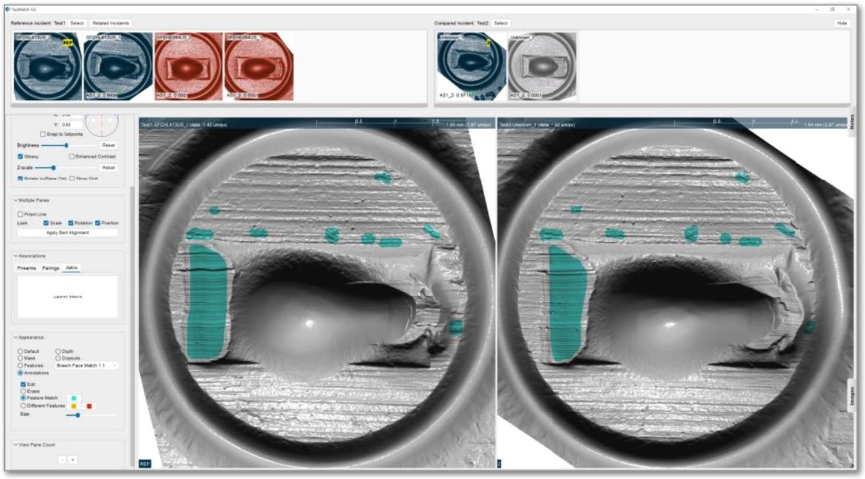


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

VCM software can also extract surface profiles. Figure 29 shows the aperture shear marks; the examiner can extract a cut profile, as exhibited in the red and blue curves. The examiner can use this to compare the profiles to determine how well the profiles match up for the two specimens.

Figure 29. TopMatch VCM dashboard surface profile extraction

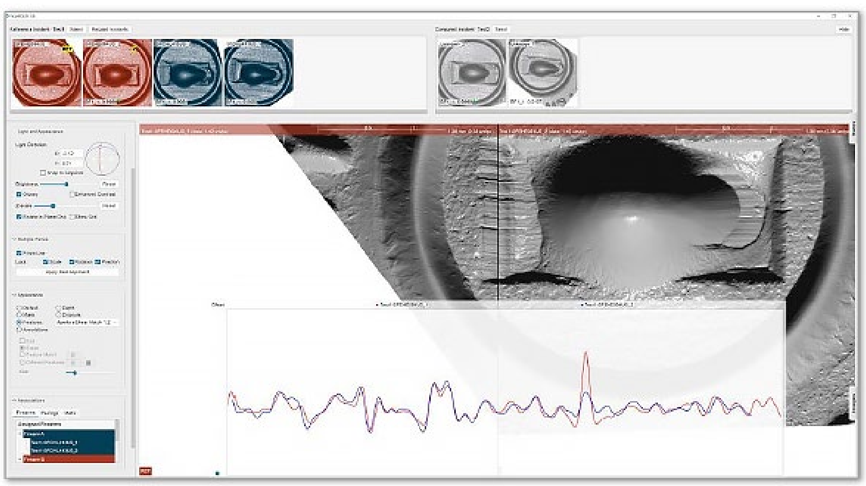


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 30 shows two test fires from the same firearm. The software uses an algorithm to color the surface a darker shade of purple/blue where there is greater geometric similarity identified by the algorithm and with lighter or no shading when little or no similarity is identified. This helps draw the examiner's attention to these features, something that cannot be done on a traditional comparison microscope. In the documentation process, the VCM software allows the examiner to describe particular features of the scan by adding annotations next to that surface image, and directly tying observations to the object of scrutiny. These structured case notes are all typed, rather than handwritten, making them easier

Figure 30. TopMatch VCM dashboard algorithmic identification of geometric similarity

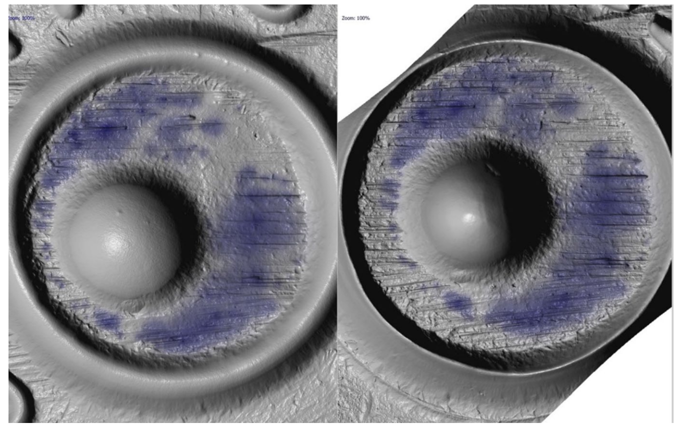


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

to read. The VCM case notes can be added to the laboratory information management system (LIMS) for archival purposes and can be digitally signed and distributed.

It is worth noting the use of algorithmic based analysis is optional and can be disabled during the VCM examination process. In many cases, examiners will conduct and complete their examination before activating any algorithmic analysis. Careful use of any algorithm can provide the benefits of a second review and statistical score while minimizing the likelihood of unwanted bias.

Before 3D virtual microscopy can be implemented and used in casework, an agency must validate both the hardware the software, and every examiner who will use the hardware or software should be trained and competency tested. Several validation studies have been conducted including that developed by the FBI Firearms/Toolmarks Unit (FTU); several others have been published in either the AFTE journal or peer-reviewed journals such as the Journal of Forensic Science.[42, 43, 55] The FBI FTU went live using 3D VCM in 2017; before the FBI could use VCM in casework, the agency created a validation study for examiners that consisted of previous casework and items selected from their reference collection. The cases were blinded, and the examiners worked through all scans passing all internal quality controls and producing results that were deemed as good or better than traditional light comparison microscopy. Cadre published their first peer-reviewed validation study in a 2018 issue of the Journal of Forensic Science.[42] In this pilot study, 46 trained examiners from 15 laboratories examined two, seven-test fire-sets and reported no errors on their analysis.[42]

Another much larger peer-reviewed study, the virtual comparison microscopy error rate (VCMER) study, was published in a 2021 issue of the Journal of Forensic Science and presented at several conferences.[43] This study had more than 100 global participants and each participant examined 40 test-set triples, consisting of two knowns and one unknown, with the examiners tasked to reach a conclusion regarding common source using a five-point range of conclusions.[43] This type of study design is favored because it resembles casework; the error rate in this study was between 0.1% and 0.2%.[43] This reported error rate is on the lower end of the range typically reported for traditional light comparison microscopy studies with error rates between 0% and 1.5%. These results support the conclusion that virtual microscopy is a valid technique for rendering source conclusions.

In 2018, the Royal Canadian Mounted Police (RCMP) completed their internal validation study, which was published in a 2022 issue of the *Journal of Forensic Science*.[55] In this study, Knowles and colleagues had 13 firearms examiners in the RCMP, comparing 50 pairs of test fires, concluding that virtual microscopy produced results as good as or better than traditional light comparison microscopy.[55]

In addition, Cadre recently conducted a virtual comparison microscopy topography resolution (VCMTR) study researching the effect of scan resolution on inconclusive rates. The study involved more than 100 participants. Results have been presented at multiple conferences, and a manuscript is currently being prepared for publication.[56, 57] In the future, it is anticipated that more VCM studies will be published by laboratories implementing the technique, academia, industry, and industry collaborations.

One of the limitations of traditional approaches to training and validation studies is that a physical set of test fires has to be sent to every participant, which means that thousands of test fires have to be created and subsets sent to every participant by mail. This poses a challenge because in some countries it is illegal to send cartridge cases in the mail. In addition, with a traditional (non-VCM) study there is concern about sample-to-sample variability, as the test taken by one participant is not exactly the same as the test taken by other participants. With VCM, a set of test fires can be scanned and distributed electronically so that all participants get the same identical 3D surfaces. In addition, validation studies conducted with Cadre's VCM offer insight into the decision-making process as information about how the examiner reached their conclusion because of the surface annotations attached to the virtual image; thus, these black box studies become white box studies in which researchers gain insights into the examiners' thought processes. This can be captured by asking the examiner to color the surface to indicate the areas of similarity and difference that they used in reaching their conclusions. Figure 31 shows examples of surface annotations where the blue-green indicates areas of similarity. After the participants completed their annotations, the researchers examined the unknown for each test-

Figure 31. VCMERS surface image annotations of similarity

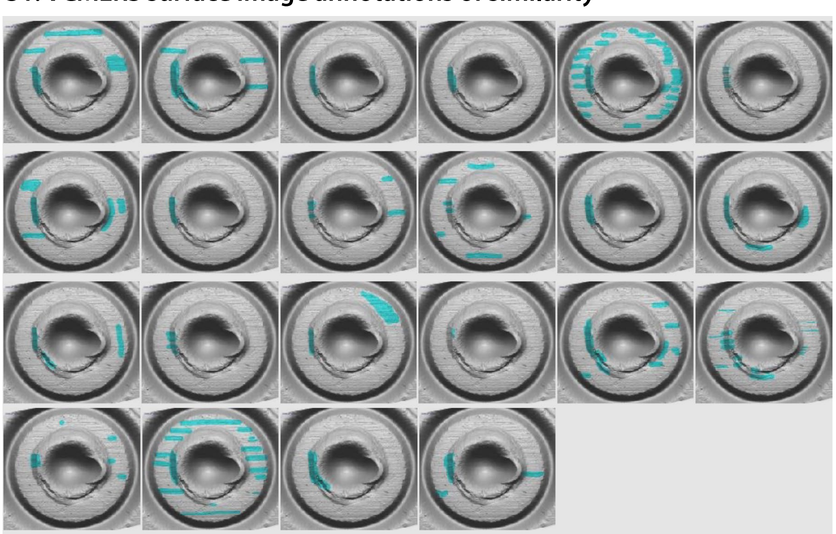


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

set from all the participants and combined them into a summary annotation map (shown in Figure 32).

Figure 32. VCMERS summary annotation map

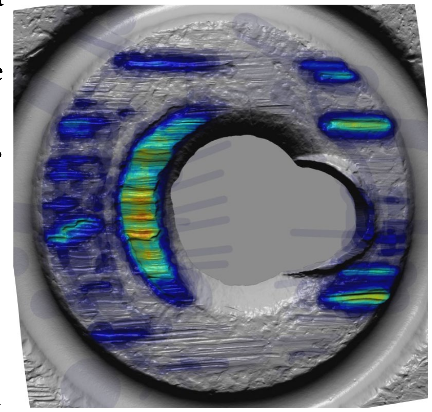


Image courtesy of Dr. Ryan Lilien presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

In Figure 32, areas of the surface that are not colored are those that no participants marked as being significant when reaching a source conclusion. Areas that are marked in cool colors, are areas that a small percentage of participants marked, and those that are marked in the warmer colors such as yellow, orange, and red are areas that many people marked, with red being the most common. These maps can be made to examine similarity, which colors the areas people marked as similar, or for the parts of the surface that individuals marked as different, creating a difference map. In Figure 32 most of the similarity was noted in the region of the aperture shear. One interesting aspect of these maps is that they show strong agreement among annotations made independently by different examiners. The maps demonstrate consistency within the examination process.

Developing comparison algorithms will be useful for addressing a few main uses. A toolmark comparison algorithm is a mathematical model that takes surface topographies as input and returns a score that quantifies geometric similarity. The score may also represent the likelihood that the compared specimens were fired from the same firearm. In addition to a score, algorithms can also give an alignment that best orients the two surfaces. Algorithms can also be used to identify areas of the surface that have the most or least agreement.

Database searching is essentially a recall function in which evidence is scanned on a machine, the images are compared to a database of other scans, and the algorithm generates a list of candidates as the output. The examiner can then work down that list of candidates, comparing the items that matched against the query using VCM, if it has been validated in the laboratory, or by going back to get the physical specimens to compare with a traditional comparison scope. Firearms examiners have been performing database search and recall over the past 30 years using either the DRUGFIRE or IBIS NIBIN system. An examiner does not need to use virtual microscopy to take images, search a database, get hits, and then correlate, but using 3D VCM across the entire workflow is likely to improve the accuracy and efficiency of these processes.

The second use of algorithms is calculating pairwise statistical scores to tell the examiner how likely the two surfaces are to have come from the same firearm. A pairwise statistical correlation would allow examiners to make quantifiable statements with statistical support for the strength of evidence and the likelihood of same source match probability in reports and testimony. To do this, a large population database is needed that includes a full range of firearm makes and models, manufacturing types, ammunitions, and markings that are seen in casework. There are currently several reference collections that have been scanned by multiple different organizations including the FBI FTU, RCMP, state laboratories, NIST, and Cadre. Many of these scans have been submitted to the NIST Ballistic Toolmark Reference Database for public access.

Another use of algorithms is to facilitate the examination process by using algorithms to triage and sort items regarding potential source. For example, if the computer was given 10 different cartridge cases, the algorithm would analyze them and sort them into groupings. For example, it might determine that 4 cartridge cases are from one firearm, 3 cartridge cases are from a second firearm, and 3 cartridge cases are from a third firearm. This could be performed as part of the verification process, or it could also be performed by a technician before any of the examination takes place to select items or give some triage information. Then an examiner would perform a visual comparison to see if they agree with the computer algorithms' assessment.

Algorithms can also be used to facilitate examination by aligning two images at a micron-scale level. Once aligned the examiner can adjust a virtual prism line, which is a divider line, to reveal more or less of each of the two surfaces. Since the algorithm has aligned these images at the micron scale, the examiner can really see how the features that are reproduced from one cartridge also appear on the other cartridge. Another visualization mode can be used for comparing aperture shears; the examiner can look at the extracted topographical profile that corresponds to the striated marks and move the indicator line through the profile to highlight different areas of the surface that correspond to the peaks. This type of view is something that could never be visualized with a traditional 2D comparison image.

In summary, 3D VCM gives an accurate microscopic representation of an object and is moving from the research laboratory into the crime laboratories. Several laboratories are now using virtual microscopy in casework. 3D VCM can be used for several applications, including its primary function to determine source conclusions, training, proficiency testing, verification, and statistical scoring. There are several things an examiner can do with virtual microscopy that they cannot do with traditional comparison light microscopy. The software can allow an examiner to save multiple different scan orientations or alignments and then come back the next day and restore the surfaces as they were when they left off. VCM also saves surface annotations and linkages, allows for remote access for verification, and saves structured case notes, all of which improve the overall work product. Completed validation studies support the use of VCM; these studies included over two hundred participants. These validation studies have supported the hypothesis that virtual microscopy has similar or lower error rates than traditional microscopy and much lower inconclusive rates than traditional microscopy. Finally, algorithms that search a database for determining statistical scores or facilitating examination should be generalizable to work on real world data. This requires that scans be collected from many of different types of firearms and lots of types of ammunition.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

The Forensic Technology Center of Excellence Transition of Emerging Technologies

Since 2011, RTI International and its collaborating partners have operated the FTCoE, a program of NIJ. The FTCoE is tasked by NIJ to act as a bridging organization within the forensic community to support the adoption of technology solutions. The FTCoE helps NIJ evaluate how well a solution addresses the intended functional requirements. This determination can range from market analysis to real functionality testing for mature solutions. Ultimately, the FTCoE works with NIJ to facilitate the introduction of the solution into practice. Thus, the FTCoE plays a critical role by identifying promising R&D solutions that may substantially impact the practice of forensic science and warrant further NIJ investment. The FTCoE recognizes that transition of research into forensic applications demands more than just raising awareness, supporting acceptance, and facilitating adoption by the forensic community, but rather that the technology or process must stand up in a court of law based on scientific merit. Thus, considering the barriers that grants must overcome to have impact for both forensic researchers and practitioners is critical.

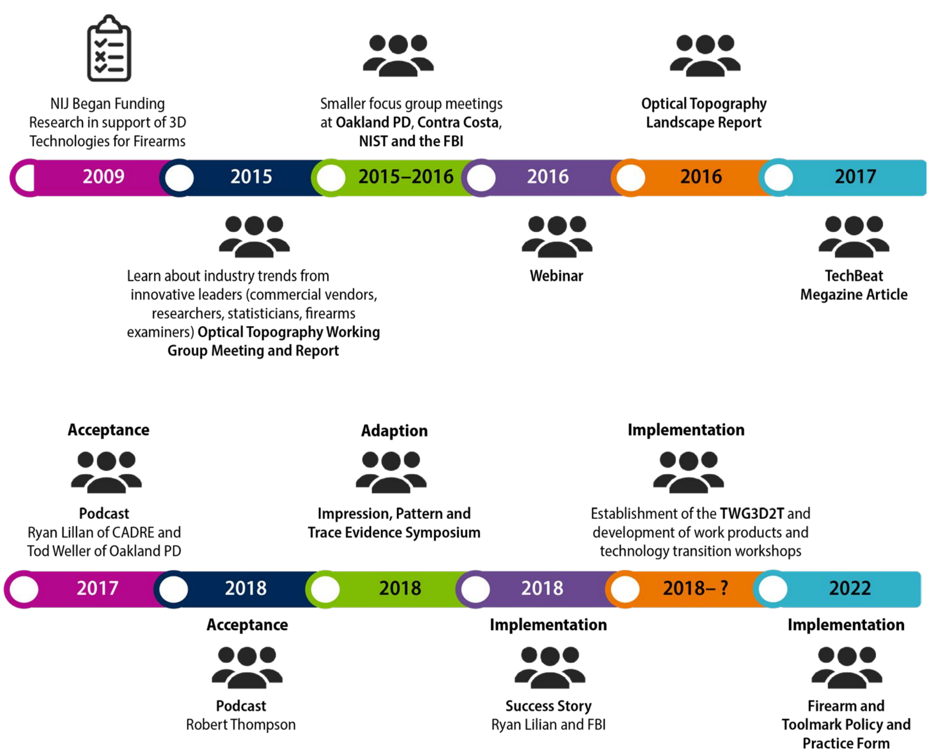
Transition and implementation of 3D imaging technologies for firearm and toolmark examination is one example of an emerging technology that the FTCoE has been facilitating. NIJ began funding research in 3D imaging technologies for firearms examination in 2009, and the FTCoE began efforts to facilitate the adoption of this technology in 2015. Some of these activities included convening a Forensic Optical Topography Working Group of firearm examiners, researchers, and instrument manufacturers from the United States and beyond. The working group included members from federal agencies (NIJ, NIST, FBI, ATF), state, and local agencies (Contra Costa County Office of the Sheriff, Illinois State Police Joliet Forensic Science Laboratory), universities (South Dakota State University, Iowa State University), international forensic agencies (Netherlands Forensic Institute, Belgium Nationaal Instituut voor Criminalistiek en Criminologie), and manufacturers (Alicona, Zeiss, Leica Microsystems and FTI, Inc.).

The working group sought to establish the applicability and validity of optical topography to forensic investigations by examining optical topography instruments, methods, data systems, and analysis from a practical perspective for ballistic and toolmark identification, including requirements for systems that may be deployed in crime laboratories. The working group also produced publications or training materials that can be accessed by the entire forensic community and that will provide guidance to practitioners on applications and recommendations for further R&D and capacity assistance. The group reviewed current and past efforts to implement optical topography in the crime laboratory, including the application of confocal microscopy and published a final report summarizing the working group discussions. Noting the value of the comparison microscope to identification decisions in current practice, the working group determined that it was unlikely that optical topography would supplant the comparison microscope as the primary tool for the forensic examiner in the near term. Instead, optical topography is likely to be a confirmatory tool or a method to examine very difficult comparison cases.

Following the working group meeting, a practical review of examination methods at the FBI laboratory was conducted by a subgroup to establish consensus on the application of optical topography for ballistic identification with respect to examiner practices, instrument requirements, training, and analysis. An additional review was held at the Contra Costa County Office of the Sheriff to discuss the challenges that they experienced with implementation and the firearm research they participated in at their laboratory. As a follow-up from this meeting, and in partnership with the working group, the FTCoE conducted a landscape study of forensic optical topography. This landscape study compared available instruments, including those more commonly used in industries other than forensic; a discussion of the barriers to implementation; practical and technical considerations for adopters and provides an overview of applicable international standards and technology developments. The FTCoE is currently working on an update to this 2016 report that will be published in early 2022.

Figure 33 presents a timeline of FTCoE activities from 2009 to 2022.

Figure 33. Timeline of FTCoE activities to promote the implementation of 3D imaging technologies



Following the working group meeting, the FTCoE facilitated several smaller group meetings with agencies that had implemented 3D technologies into the firearm case flow, including Contra Costa, the Oakland Police Department crime laboratory, and the FBI. In 2018, the FBI established the Technical Working Group (TWG) for 3D Toolmark Technologies (TWG3D2T) with administrative support from the NIJ FTCoE. The TWG3D2T was established to "support and promote the advancement of the forensic application of 3-Dimensional (3D) optical topography instruments in the Firearms/Toolmarks community through the development and dissemination of consensus-based standards, guidelines, best practices, and recommendations." The objectives of the TWG3D2T are to:

- Define the scope and practice areas of the discipline(s) of firearms/toolmarks (F/T) specific to 3D measurement technology and its application to the discipline of F/T.
- Recommend standard practices, protocols, reports, limitations, and terminology.
- Recommend standards for data interpretation and wording of conclusions.
- Recommend education, training, and continuing education requirements.
- Promote and disseminate research and development priorities for the community.
- Collect and distribute discipline-specific information on scientific foundation.
- Seek international recognition and harmonization of appropriate TWG3D2T work products.
- Establish relevant toolmark population statistics for F/T.
- Promote the harmonization of TWG3D2T documents with applicable laboratory accreditation requirements and quality assurance.

Law Enforcement Perspective of Evidence IQ Ballistic IQ Imaging Software System

Evidence IQ's Ballistics IQ is a system that can be used to scan casings and submit the cases for sorting and triage, providing insight into how many firearms were used, type of firearms used, and potential links between cases. The Escambia County Sheriff's Department (ECSD) purchased the Ballistics IQ system in April 2020 and implemented the system in May 2020. ECSD serves a population of just over 300,000 and collects approximately 600 firearms and 930 shell casings on an annual basis. ECSD does not have a firearm examiner in house, so all ballistics evidence is sent to the Florida state laboratory for examination. Currently the Florida state laboratory backlog is approximately three to six months. Therefore, ECSD needed equipment that could provide a quicker turnaround time to produce intelligence and investigative leads to help solve these crimes.

Before incorporating Ballistics IQ into the ECSD workflow, the agency completed a two-day training and then developed technical standard operating procedures. At ECSD, every casing that is turned into the evidence unit, regardless of the circumstances (e.g., homicide, property crime), is rerouted to the crime scene unit. When a crime scene technician receives the casing, it is documented and then swabbed for DNA evidence and processed for fingerprints. Once that is completed, the information is uploaded into Evidence IQ's Ballistic IQ imaging system software. If ECSD has a request to test fire a firearm, the request is submitted into the system; the crime scene techs will process the firearm first for DNA or fingerprints. The firearm is then transferred into the custody of a sworn officer, who will then take that firearm to the local ATF agency, and it is test fired there. The casings are then brought back to ECSD where the test fire is uploaded into the Ballistics IQ gallery. Every single submission generates a report and is also put on a spreadsheet kept for statistics and documentation purposes.

Each casing that is submitted is not a full case submission, because with Ballistics IQ, the agency pays approximately $260 per case submission. ECSD has a three-year contract with Evidence IQ, which includes 100 cases a year, so ECSD saves case submissions for homicides or any type of major shootings. Therefore, not every single casing that is uploaded into the ECSD library is submitted for correlation, but every casing is uploaded into the library. If ECSD issues a case submission, two reports may be sent back to ECSD. First, a crime scene analysis report indicates which casings should be entered into NIBIN, the type of firearms that could have fired the casings, and whether these casings have any potential links to any other casings in the system. If there is a potential link, a second report will be generated with this information. The investigator is then notified of the potential links so they can follow up on the lead.

As of November 18, 2021, ECDS had uploaded 34 case submissions and scanned 1,971 cartridge casings. The turnaround time for analysis is 2 hours and 12 minutes, but some results are reported in 45 minutes or less. Of the 34 case submissions, 27 potential links have been reported, and 157 unique firearms have been identified. Although 1,900 casings is not a lot, ECSD has only had the Evidence IQ Ballistics IQ since May, so it will take some time to build up the library working through the backlog. ECSD currently has about 2 years of backlog in the library.

One of the benefits of ECSD use of Ballistics IQ is timely results. If evidence is sent to the Florida state laboratory, turnaround time is 3 to 6 months, but with Ballistics IQ, the results are reported in less than 2 hours. Ballistic IQ is cost-effective and can build the contract around an individual agency's budget. A smaller agency may only need 20 or 25 cases a year so the statement of work can be customized to the agency's budget and needs. Ballistics IQ will also use an algorithm to triage evidence for submission into NIBIN, which can save an agency time and resources. The agency also has complete control over when they want to utilize Ballistics IQ or not.

The software also can identify unique firearms and provide insight into the type of firearm used to fire the uploaded casings. In addition, it can provide immediate intel and is portable and rugged, so it can be used in the field at crime scenes. While crime scenes are processed, investigators can upload casings, and the software will report what type of firearms the investigator should be looking for. This is particularly helpful if a suspect in possession of a firearm is at the scene, allowing the investigator to immediately take the suspect into custody. The investigator can then take that firearm to test fire it and upload the test fired casings into the system to see if the test fires match the casings found at the shooting. The investigator can then immediately use this information to go back to the suspect and present that information to them so that they can explain why the firearm in their custody came back matching the casings on scene. This immediate turnaround is not something an investigator can get with any other technology currently available.

ECSD also uses Ballistics IQ's option for shared networks. With this feature, the software will compare an agency's case submissions against those in any other nearby agencies that are also on the network. As previously mentioned, ECSD has had 27 potential links found using Ballistics IQ. One example case was a homicide that occurred in Okaloosa County, which is nearby. That agency also has the Evidence IQ Ballistics IQ imaging system and had uploaded their homicide casings into the system two days prior. ECSD officers were in pursuit of the suspect, shots were fired, and the suspect was shot and killed. The suspect had a firearm in his possession that was test fired. The casings were uploaded into the ECSD gallery and came back as a match to the casings found at the Okaloosa crime scene. Another example was a homicide in which over 50 casings were found. After those casings were submitted, the results came back in 1.5 hours; those casings had been fired by three different guns, indicating that this was not a single-shooter event. ECSD was able to create a sketch with the layout of the casings, using the information from the crime scene analysis report from Evidence IQ, and were able to determine where each of the possible shooters were located when those casings exited the firearm, giving the investigators a general idea of reconstruction of the scene. In addition, ECSD received three potential links in that case, and the firearms that were used were linked to three other previous shootings. For those previous shootings, ECSD had possible suspects that they were able to follow up with, resulting in ECSD finding the suspect in this homicide.

As far as limitations, Ballistics IQ is a presumptive tool for generating investigative leads. Confirmatory testing must be sent to a laboratory. Evidence IQ is planning to provide confirmatory testing in the future. ECSD has testified in court on its use of Ballistics IQ and has not had any issues to date. Ballistics IQ is not NIBIN nor is it a competitor of NIBIN. Ballistics IQ is meant to be used in conjunction with NIBIN because it can triage casings for entry into NIBIN. Currently, analysts have to manually orientate the casings in the system, but in the future the system will be able to do that automatically, which will speed up the process.

Ballistics IQ for the Crime Laboratory

The Plano Police Department Firearms Unit (PPDFU), currently serves over 63 agencies in the area and provides forensic firearm services and analysis to law enforcement agencies throughout the Greater North Texas region. Established in 2004, the unit serves as a regional ATF NIBIN entry point. Under the umbrella with ATF, PPDFU provides critical analysis and support to any of the investigations and prosecution of criminal activity involving the use of firearms in their jurisdiction. In recent years, the Texas's state laboratory is no longer equipped with NIBIN, so many of PPDFU policies and procedures have had to be adjusted to support agencies in need. Much of PPDFU's success has directly reflected their ability to provide timely examinations comparable to those of the state laboratory using local government resources. Since 2016, the PPD laboratory has seen a 300 percent increase in evidence submissions, prompting PPDFU to look for solutions to enable the agency to maintain processing this increase in cases.

In April 2021, PPD implemented Ballistics IQ. When PPDFU obtained the system, the agency had several objectives that they were seeking to investigate. The first was whether it is suitable for Ballistics IQ to triage both evidence exemplars and test fired exemplars for entry into NIBIN. PPDFU also wanted to evaluate the system's accuracy, specificity, reproducibility, and timeliness regarding cartridge case triage. Ballistics IQ would support the NIBIN technicians because the system uses an algorithm that identifies the most suitable cartridge for NIBIN entry. Even more importantly, PPDFU wanted to know whether the system's triage algorithm coupled with the virtual correlations center (VCC) would include a known match cartridge case and exclude a known nonmatch cartridge case when most appropriate.

During implementation, PPDFU also wanted to evaluate whether there was risk associated with adopting the system as a laboratory triage aid. Because it had been recommended that both aspects of the system be used, PPDFU wanted to validate the system in its entirety, including validation of Ballistics IQ's performance based on the manufacturer's recommended operating procedures assessing the quality or the performance of the algorithm independent of the VCC. The validation study began in April 2021, and PPDFU went live with Ballistics IQ in September 2021. During the validation, PPDFU used several different validation studies that were previously known to the firearm examiner's community. They also used Collaborative Testing Services (CTS) proficiency tests, and the results indicated that Ballistics IQ successfully provided investigative information in a timely manner. On average, PPDFU received crime scene analysis triage reports returned in less than two minutes using the system's algorithm. PPDFU also received VCC approved triage reports in under three hours.

During the validation, PPDFU found that the system effectively identified the number of firearms present among groups of cartridge cases. In addition, the Ballistics IQ algorithm was able to identify and associate unknown matches to the firearms present within groups. The algorithm also has restrictions and triage CIs that would group cartridge cases in high confidence, low confidence, and in undetermined groups with a notification to push the user to utilize the VCC as a supporting element. Based on the validation, PPDFU changed their existing policy to triage all firearms and toolmarks evidence that came into the laboratory based on their observation that the Ballistics IQ system exhibited optimal performance when coupled with the expertise of the VCC.

PPDFU has a three-tiered analysis, in which all the evidence is entered into Ballistics IQ, is transferred to microscopic analysis, and is verified. Therefore, all evidence is reviewed three times on different levels of analysis. Additionally, PPDFU has external administrative oversight with their NIBIN technician. Any time evidence is entered into the Ballistics IQ system and submitted to the VCC, the agency is notified immediately of any errors in entries, orientation issues, or anything that may be wrong in the documentation, or the data based on the information that is contained the entries. This has resulted in increased efficiency in time management across the laboratory, primarily because of a consistent workflow. PPDFU has also experienced an increase in the dissemination of intelligence and investigative leads from the crime scene analysis (CSA) triage reports and the approved VCC reports. As soon as the reports are returned to PPDFU via email, they are reviewed, and notification is immediately sent to the investigative agency and the detectives assigned to that case. This increase in both internal and external communication has sparked questions regarding the intelligence in the packet of information that is now submitted. Agencies are intrigued by the amount of information they receive in these emails and motivated because now they know how many guns are involved, what cartridge cases exist in groups associated with that particular firearm, providing them a wealth of information that they can use in their investigation or in their approach a particular case.

Another advancement for PPD was the addition of multisystem links. PPDFU has had several cases that have been tied to one another and triaged through Ballistics IQ. Once a cartridge case has been selected for NIBIN entry the agency will get a NIBIN lead outside of their jurisdiction. These multisystem links have confirmed that Ballistics IQ and NIBIN complement and validate each other. PPDFU has also seen a backlog reduction. Before the implementation of Ballistics IQ, even when PPDFU would try to control the amount of evidence taken in, PPDFU noticed that it could not keep up with the amount of cases coming in. Now, PPDFU can focus on cases that have been triaged. The agency receives a notification for those cases that need a court written report for the detective. The laboratory can focus on completing those requests while the evidence is triaged through the Ballistics IQ system. This workflow has led to a reduction in turnaround time from approximately 12 days to 7 days and continues to decline.

In 2021, there were more than 39,000 gun violence deaths and more than 629 mass shootings.[58] Northeast Texas is an area of concern because of the increase in gun crime, and one of the elements identified in combating gun crime is strong forensic firearms investigative capability. Through research and application, the PPDFU can help in advancing investigations, assisting in proper utilization of administrative resources, and decreasing judicial processing time from crime scene to courtroom. From a laboratory perspective, every agency should explore having a solution to enhance firearms investigative capability, whether it is case backlog reduction, attempts to decrease turnaround time, proper utilization of administrative resources, or simply enhancing the quality of investigative intelligence.

PPDFU recently had a case for Gainesville, Texas Police Department in which there were four suspects, and 39 cartridge cases were submitted across three different incidents, all of various calibers. While the evidence in these cases had been collected, none had been examined at the time of submittal. Through triage, PPDFU was able to link the weapons that were recovered in a previous search to the cartridge cases found across all three incidents in less than three hours of submitting to Ballistics IQ. In addition, the email notifications that were compiled based on those Ballistics IQ reports were able to be disseminated the same day as analysis. The initial information received from the investigators at the time of submission was that these individuals may have been involved in several crimes, but they only had evidence connecting a string of different events. The Ballistics IQ link ultimately connected a drive-by shooting, a traffic foot pursuit with shots fired, and a previous aggravated assault with a deadly weapon incident. While it may be difficult to disrupt the cycle of violence, appropriate use of technology, intelligence, resources, partnerships, and effective communication all play a major role in assisting in the effort to solve crime.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Evofinder®: From Validation to Implementation

For more than a decade, researchers, both forensic and academic, have investigated the application of existing surface metrology and 3D surface topographical microscopy methods to the field of forensic firearms examination as it pertains to bullet and cartridge case comparisons. The application of these techniques has resulted in a new technique known as virtual comparison microscopy, or VCM. This process involves comparing high-resolution 3D surface topographical images of toolmarks present on fired cartridge cases and bullets as compared with traditional optical light microscopy, or LCM. The examiners conduct the same type of analysis as traditional comparative microscopy, and still following standard methodologies within the field of firearm and toolmark analysis; however, these comparisons are conducted on a computer screen rather than through a microscope.

Why should a laboratory pursue VCM? As with most all processes, techniques, and equipment, traditional LCM has advantages and disadvantages. However, until the implementation and deployment of VCM equipment and software, LCM was our only method for comparing toolmarks. In the event a particular set of toolmarks were difficult to observe using LCM, there were no additional methods available to attempt to resolve these toolmarks. Certain toolmark types are inherently difficult to examine with LCM, such as fine granularity on a highly reflective surface. Often these types of toolmarks are difficult to resolve due to hot spots created by the intense reflectivity from the surface.

With the deployment of VCM technologies as a new tool to assist forensic scientists in the comparison of toolmarks, there is the potential to increase the examiners' ability to observe more of the toolmarks present with sufficient clarity to render an opinion. Although Evofinder is not immune to reflectivity issues, this problem is not as pronounced as with LCM. Furthermore, Evofinder has an infinite depth of field, allowing the examiners to observe more of the toolmarks in a single view without continued focus adjustment. This idea of providing examiners with a new tool was one of the primary reasons the Phoenix Police Department Crime Laboratory (PPDCL) decided to pursue VCM with the implementation of Evofinder. By expanding the number of technologies, the firearms and toolmarks examiners have at their disposal, the greater the likelihood the laboratory can reduce the number of inconclusive results reported. Although Evofinder has the potential for toolmark databasing and computer-based comparisons, PPDCL does not intend to utilize it for these purposes and are not currently using these features. The PPD has made a large investment in the NIBIN IBIS program for firearm and toolmark databasing purposes. Therefore, the deployment of new equipment for databasing was not necessary at this time, and PPDCL wanted to deploy the equipment and begin VCM before investigating any of the computer-based conclusion support.

Before purchasing Evofinder, PPDCL began investigating into the possibility of implementing VCM in the laboratory and examined technologies available from multiple vendors who provide VCM hardware and software. This process included discussions with vendors, hands-on demonstrations either within the laboratory or at multiple AFTE conferences, as well as presentations and workshops hosted by specific vendors about their technology. The primary criteria when evaluating the available systems included (1) operator ease of use; (2) system versatility (i.e., the system's ability to scan or acquire toolmarks both on bullets and cartridge cases); (3) image resolution and quality; and (4) speed of acquisition. The agency wanted to ensure that the selected system would be versatile and usable by the average examiner. The various VCM systems available on the market use a range of different microscopy methods and technical or technological methods. The Evofinder system uses focus variation topographical microscopy to create a 3D rendering of the surface topography of the scanned items. Evofinder's ability to scan the toolmarks present on cartridge cases and bullets, coupled with relatively quick acquisition times and exceptional image quality, made it an ideal candidate for validation within the laboratory as a virtual comparison microscope.

At the time PPDCL began implementation, very few of the laboratories across the nation had completed validating a VCM system. The validation team set aside a significant amount of time to design and plan the validation scheme to ensure that upon completion, they would be confident in the examiners' abilities to visualize toolmark patterns in VCM and in the reproducibility of the 3D images produced by Evofinder. This was accomplished with the NIST standard bullet and cartridge case, previously completed externally produced proficiency test samples, samples including both bullets and cartridge cases created in house, and bullet and cartridge cases from previously completed consecutively manufactured toolmark studies. The validation included the completion of hundreds of comparisons across numerous sets, including both bullets and cartridge cases from multiple calibers, toolmark types, and cartridge case and bullet construction. All comparisons were completed by eight practicing firearm and toolmark examiners.

Additionally, numerous samples were selected for repetitive entry into the system to test the reproducibility of the scans and images. Upon completion of the validation study, no false positives or false negatives were reported. The quality, resolution, and reproducibility of the scans demonstrated no concerns with feature dropout or inclusion that would affect the examiner's ability to use the images for comparison purposes. The process of system selection and validation was a valuable learning experiences for the examiners as both VCM validation designers and VCM validation participants. Historically, it has been thought that comparison conclusions must be made by physical observation of the toolmark by means of LCM and should not be reached by looking at pictures of the toolmarks; however, the quality and resolution of the images being produced by the VCM systems currently on the market has changed this mindset. VCM does present a new set of challenges and learning curve for examiners that are accustomed to LCM. One of the first things noted when training on the system and evaluating the validation results to include interviews with the validation participants was what was referred to as "information overload." The volume and detail that is seen by the examiners in a single view is much greater with VCM than with LCM. This drastic increase in information presented to the examiner all at once takes some mental conditioning to train the eye and the mind to see the patterns within the evidence. This information

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

overload and the learning curve with visualization of the VCM images caused examiners to be more conservative in their conclusions until they became more confident in this new format.

Figure 34 shows two cartridge cases fired by the same gun as viewed through an LCM at 22× magnification using oblique fluorescent lighting. Typically, the toolmarks on these cartridge cases will be compared at a magnification between 30× and 40×. This view through the LCM has excellent clarity and resolution of the toolmarks present but does not contain as much information as the same two cartridge cases viewed through the Evofinder system shown in Figure 35. This image shows the same two cartridge cases, with the screen zoom set of 100 percent on the Evofinder system to simulate as closely as possible the view seen in the previous image captured through LCM. Notice the increased volume of fine detail seen throughout the image. This forces the examiners to look at much more detail in the same view. Until an examiner becomes used to seeing this new format, the increased detail can be mentally and visually overwhelming.

Upon completion of the validation study and approval by the laboratory administration and quality manager, PPDCL developed an implementation plan to ensure that the examiners were competent and comfortable with the use of the Evofinder system in casework. The plan involved a 6-month implementation window in which VCM was used in conjunction with LCM. The plan required that all comparison cases be completed with LCM first; a minimum of 50 percent of all those cases were duplicated on VCM with the duplication of at least one sample of each level of conclusion. This allowed the examiners to complete the comparison optically and then see the same toolmarks virtually. With this implementation, the examiners were required to show both LCM and VCM data and images in their analysis notes. If an inconclusive LCM duplicated on Evofinder resulted in a definitive conclusion, including either an identification or an exclusion, any additional inconclusive comparisons completed on LCM would then be completed on VCM. PPDCL concluded this implementation window in late 2021 and is now in the process of establishing Evofinder as a standalone microscope to be used at the examiner's discretion. If all examiners are confident enough with the use of their system after training,

Figure 34. Two cartridge cases fired from the same source with optical fluorescent lighting at 22× magnification

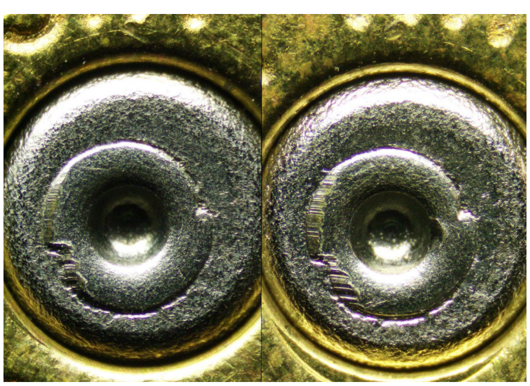


Image courtesy of Mr. Michael Beddow presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 35. Evofinder image of two cartridge cases fired from the same source at 100% screen zoom

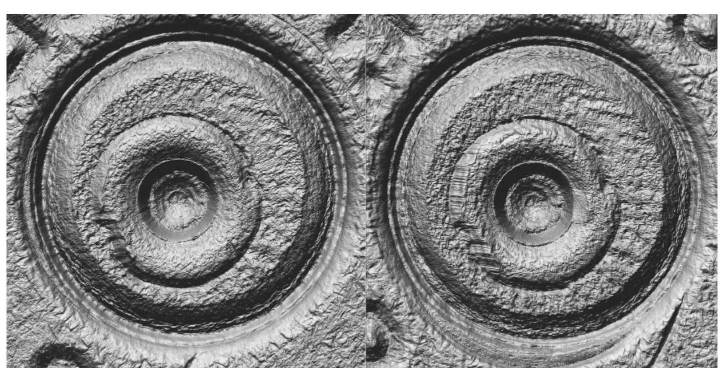


Image courtesy of Mr. Michael Beddow presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

practice, participation, and validation, a laboratory may be able to go straight to casework without the required coupling with LCM.

For PPDCL, this implementation period has been beneficial for the examiners and has allowed them to build confidence in using the system by providing the examiners an opportunity to see actual casework samples with both LCM and VCM. This process allowed examiners to gain real casework experience with VCM and provided an opportunity for the customers and the courts to become accustomed to seeing the additional information and image type from VCM within the case notes. Evofinder and VCM are not the end-all, be-all; some sample types prove to be best suited for LCM, some for VCM, and a few benefit from both. The conclusions thus far look promising in that some inconclusive LCM comparisons have been able to be successfully identified or excluded with the use of the Evofinder. VCM is not suitable for all sample types, either due to the condition of the sample or the practicality of using VCM versus LCM. Some excessively damaged bullets or comparisons of firing pin impression sidewalls are not well-suited for Evofinder. This may change with time; however, the practicality of using VCM for all samples over LCM is not there yet, such as comparisons involving Glock firing pin, aperture shear, or Hi-Point breech faces. These comparisons are typically so quick and simple with LCM, it may not be worth the extra time to scan these types of items into Evofinder.

The next few figures show images of evidence as examined using VCM and LCM. Figure 36 shows a Glock aperture shear with the LCM image on the left and the VCM image on the right. While both images demonstrate good correspondence of the stria, those on the VCM image on the right are more defined and crisper to the eye.

Figure 37 shows the relatively smooth, hemispherical firing pin impression. In the LCM image on the left, there are some unique small individual characteristics at the center of the impression. They show up as a small curve and a few dots beneath it. As the cartridge case was rotated, more features like these were observed. When viewed on Evofinder, these features were more difficult to resolve due to the nature of the type of lighting and the way the image is being captured.

Figure 36. Glock aperture shear with the LCM image on the left and the VCM image on the right

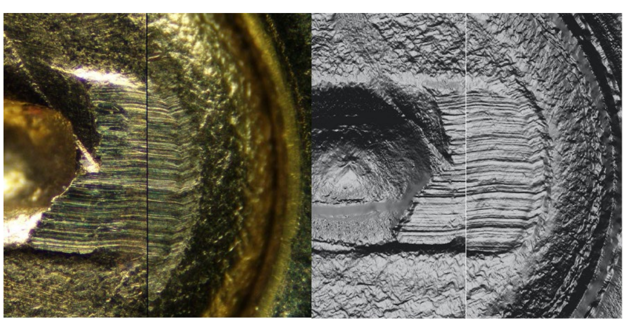


Image courtesy of Mr. Michael Beddow presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Figure 37. Firing pin impression with LCM image on the left and the VCM image on the right

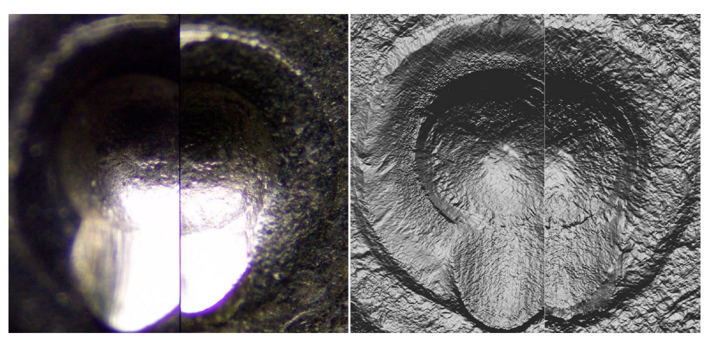


Image courtesy of Mr. Michael Beddow presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Figure 38 show toolmarks characteristic of a bunter tool impression. This is an example where the toolmarks are visible and comparable within both comparison methods while neither of them necessarily being better than one another. There are differences in contrast, but not in clarity or visibility for use by the examiners.

Figure 38. Bunter tool impression with LCM image on the left and the VCM image on the right

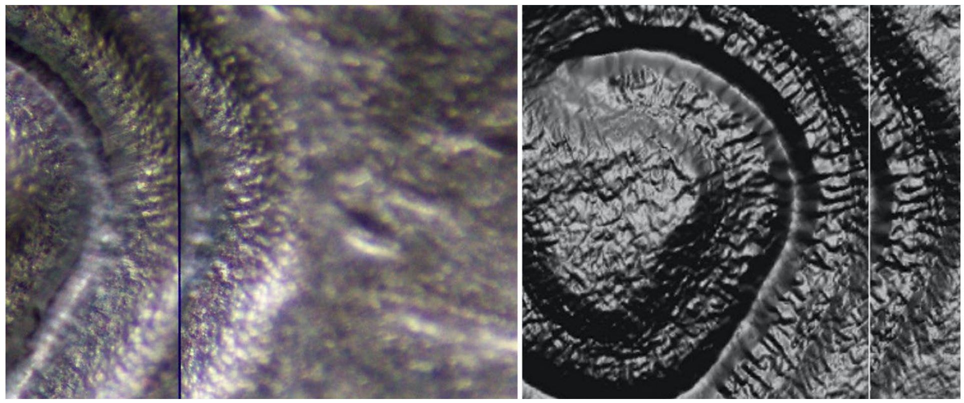


Image courtesy of Mr. Michael Beddow presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Increase in casework efficiency was not considered throughout the validation or implementation of Evofinder. However, it is something that can be tracked in the future. Since implementation, PPDCL has recorded which cases are complete with LCM, VCM, and/or both through fields within the LIMS system at the laboratory, so it would be possible to track usage efficiency. Now that PPDCL has concluded the implementation period and have Evofinder set up as a standalone microscope, they will continue to collect data related to the use of the instrument to help track which toolmark types are most suitable for VCM with Evofinder and any potential increases or decreases in efficiency, and any decreases in inconclusive results. In the future, PPDCL hopes to investigate the potential implementation of computer-based comparison results to help support visual conclusions.

Reference Population Database for Firearms and Toolmarks

The last decade has seen innovative progress in the development of algorithms, measurement instruments, data, and methods to facilitate objective analysis of firearm and toolmark comparisons. The development of 3D technologies has provided repeatable data and feature based correlation methodology was vital for the development of area and feature based correlation algorithms.

The first research study to incorporate feature-based correlation methodology into firearm examination was the 1959 study of consecutive matching striae (CMS), which is a numerical description of a toolmark used to describe the observed pattern of land and groove impressions. CMS was initially proposed by Al Biasotti upon analyzing 720 known nonmatch comparisons of land and groove impressions and finding no instances in which the CMS exceeded four.[59] Biasotti and John Murdock published a follow up study nearly 50 years later.[60] CMS has received some criticisms that counting striations is also subjective and has been referred to as a probability model not an identification model due to the inability to account for barrel changes.[61]

Recently, NIST also introduced the congruent matching cells (CMC) method for ballistic identification and estimation of error rates using 3D topography measurements on area-based and feature-based correlation cells.[62, 63] In 2008, NIST developed an XML data standard to allow for the interchange of fingerprint, facial, and other biometric information. Recently NIST has promoted the adoption of an open format for storing 3D topographical images of bullets in a format called X3P (XML 3-D Surface Profile) that conforms to the ISO5436–2 standard. This format provides a simple, standardized, ISO compliant way to exchange 2D and 3D data. This format was adopted under the Open Forensic Metrology Consortium (OpenFMC) framework for interoperable sharing of ballistic identification data on a national basis. The development of OSAC standards for toolmark topography comparison software, standards for implementation of 3D technologies in forensic laboratories and standards for measurement systems and measurement quality control along with the development of a common file format (XPs) for imaging technologies have all contributed to the development of the Reference Population Database of Firearm Toolmarks (RPDFT).

In 2018, NIST, the FBI, and the Netherlands Forensics Institute (NFI) started developing of RPDFT, which is a system that consists of a user interface, reference database of firearm toolmarks (impressed and striated), data processing modules, quantitative similarity metrics, and statistical weight of evidence calculation protocols. In the past three years, a reference database infrastructure has been developed to index test fires generated by the FBI according to their class characteristics.

The reference database consists of ground truth known match and known nonmatch comparisons and is designed to be filtered according to matching/ relevant class characteristics of the evidence being analyzed. NIST and NFI's objective toolmark analysis algorithms and statistical protocols have also been integrated into the database. These allow for objective and quantitative one to one comparisons of firearm toolmarks which are then used in conjunction with the relevant reference populations to calculate its statistical weight of evidence and provide a statistical statement of certainty. The database also provides reference data for continued innovation of correlation algorithms; however, the primary goal is to provide firearms and toolmark examiners the ability to support their testimonies with objective similarity values and statistically sound quantitative expressions for the weight of the evidence.

RPDFT is a national framework that is maintained by the FBI with the help of NIST and the TWG3D2T to maintain and update the reference data and distributions as needed. Laboratories will submit their correlation results along with the metadata to RPDFT and request the weight of evidence (WOE). RPDFT will then pull the relevant population to build the background distribution and report the WOE back to the requesting laboratory. Vendor's correlation algorithms will also be used in the RPDFT reference datasets to build vendor specific statistical distributions. To participate, commercial vendors will be required to conform to the OSAC standards with respect to hardware, software, and reference geometric standards, and bullets/cartridge case replicas will be used to confirm the minimum specifications. This will drive innovation and keep the commercial market open and competitive. The reference data are stored in the X3P format, so the data are readable by all vendors.

The development of the reference database builds on the NFI Scratch software platform and the NIST ballistics toolmark research database. The database is designed to encompass a variety of existing and future comparison metrics. Measurements are performed at NIST, and the FBI curates the data before they are imported into the database to create a large population of correlation scores. This database then allows examiners to compare features to a relevant population to calculate a likelihood ratio that the evidence in question came from firearms. The RPDFT system accomplishes this by building a relevant population using the observed class characteristics from the evidence. As Figure 39 shows, the examiner enters the hierarchical metadata for the firearm such as manufacturer; model; manufacturing methods for the breech face, firing pin, and barrel; and the number of land engraved areas (LEAs). For the bullet, the examiner will enter the manufacturer, model, caliber, primer and cartridge case material, the bullet weight, twist, and surface material. The more hierarchical data the examiner can enter, the better the population fit. If the sample size is too small, RPDFT moves up in the hierarchy to include more samples and scores, but this can affect the statistics due to the loss of specificity. The user interface also allows the examiner to focus in on a specific region of interest.

The system builds a relevant population using the class characteristics, and once the examiner enters the metadata, the database will calculate the likelihood ratio for the known match and known nonmatch based on the class characteristics entered. As an example, if the examiners receive firearms, they will perform a test fire and then compare the breech faces of the two cartridge cases to obtain a CMC correlation score. That score is then entered into RPDFT and compared with the background reference population which in this case would be the Glock reference population based on the metadata. Figure 40 shows the comparison score histogram for this example. Based on the examination the comparison

Figure 39. RPDFT data entry screen

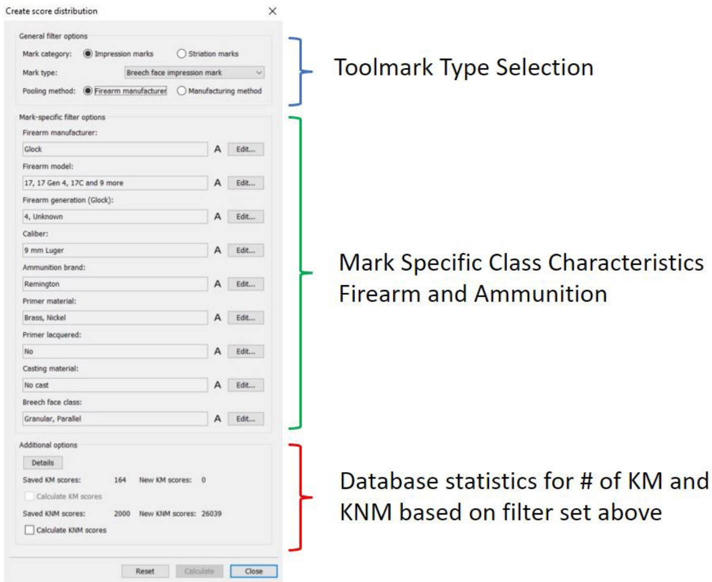


Image courtesy of Mr. Xiaoyu Alan Zheng presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

resulted in a CMC correlation of 19 cells. The areas under the red and blue curves demonstrate the background population of 164 known matches of similar types of Glocks and 2000 known nonmatches of similar Glocks. When the CMC correlation score of 19 is entered the system can then calculate the likelihood ration based on the distribution in Figure 40. Since likelihood ratios (LR) can range from one to infinity, researchers at NFI are currently examining ways to limit the LR to fit the population dataset.[64]

Figure 40. Known match/known nonmatch comparison score histograms for a Glock 9 mm with nickel primer

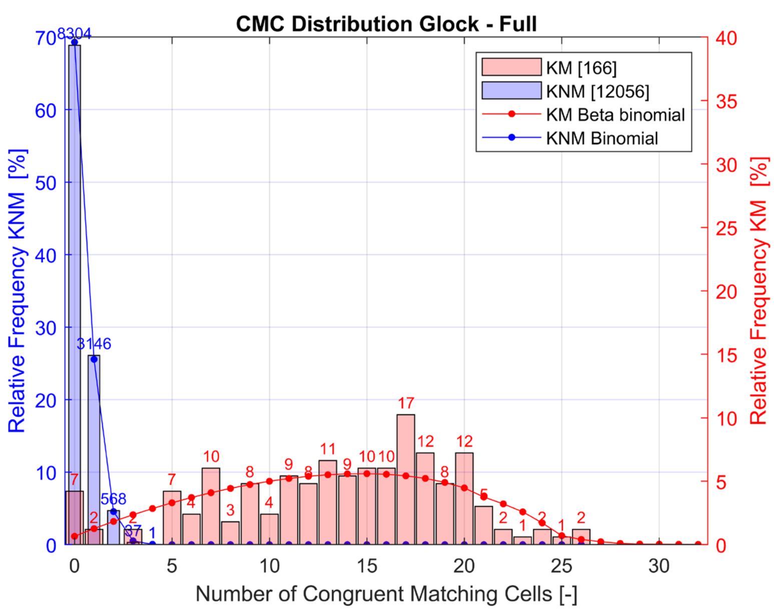


Image courtesy of Mr. Xiaoyu Alan Zheng presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

There are currently 393 Glock firearms in the database and 314 Ruger firearms. The data have been processed, and over 11,000 known matching scores have been generated with the potential of generating 1.3 million known nonmatching scores. NIST is currently working with statisticians at both the FBI and NFI to discuss strategies for statistical model selection, LR calculations and LR calibrations. By the end of 2022, the pilot test to validate the process from start to finish at the FBI should be finished. The pilot test will be conducted on old case work, parallel case work, and proficiency test-sets. Once the pilot validation study is finished, a portal will be built to allow forensic laboratories to access the population database.

Cadre Verification and Use in Casework

In 2021 the Virginia Department of Forensic Science (VA DFS) implemented the Cadre 3D scanner and virtual comparison microscopy software. The implementation plan included three verifications, the first of which focused on the equipment, the second on using the virtual comparison microscopy software to conduct comparisons of cartridge cases, and the third on the use of the virtual comparison microscopy software to sample or group cartridge cases without conducting a complete comparison examination. The purpose of the equipment verification was to demonstrate that the equipment performs as expected when it captures 3D images of cartridge cases and to determine limitations and sources of variability.

To examine variability, 10 consecutive scans of the sinusoidal standard were acquired with two different gels to determine if the gels contributed to variations in the measurements. Next, VA DFS had two examiners acquire the scans of the standard using different gels and at two different times of the day for five days. The purpose of this approach was to determine whether the user interaction results in variation or whether environmental conditions and minor fluctuations in laboratory temperature influenced the measurements. The decision to repeat the test over five consecutive days was to demonstrate stability of the equipment. The instrument records two measurements (Rsm and Ra) for the sinusoidal reference standard, where Ra is the mean of a set of individual measurements of the peaks and valleys of a surface and Rsm is the mean peak width. An ANOVA was used to compare the sets of measurements to determine if a statistically significant variance existed between the groups. VA DFS found that the gel contributed to differences, but the operator and the environment did not. The differences were very slight and did not impact the visual resolution of the image and do not impact the images viewed for comparison purposes, therefore the equipment was determined to be acceptable for use.

The next verification was to determine the limitations of rendering conclusions from 3D images. Two types of samples including cartridge cases from previously completed proficiency tests and cartridge cases obtained from consecutively manufactured slides were used. These cartridge cases had previously been examined using traditional light microscopy techniques. Consecutively manufactured slides have the greatest potential to produce samples with the highest degree of similarity representing close nonmatches and a challenge to the examiner. While the purpose of this verification was to determine the limitations related to comparison conclusions in samples that have a high degree of individual characteristics that are very similar, VA DFS also learned of limitations related to the use of the system and the workflow. For this study, a technician collected the scans but did not orient the cartridge case in the holder or evaluate the images for quality or clarity before providing them to examiners for comparison. This approach resulted in some images that were out of focus, and the examiner had to spend time orienting the images in the software. Each examiner was provided 10 sample sets from previously completed proficiency tests and five sample sets from the Hi-Point consecutively manufactured slide study.[27] A sample set consisted of two known test fires and one unknown cartridge case. 40 sets of images from the proficiency test were created and

20 sets of images from the Hi-Point tests.[27] Fourteen qualified examiners conducted the examinations.

The examiners used the Cadre Virtual Comparison Microscopy software to conduct the comparison and document a conclusion of identification, elimination, or inconclusive for each set. Each set, except for set 34, was completed by at least two examiners. Of these, 13 sets contained conclusions which differed between elimination and inconclusive or identification and inconclusive. There were no sets which had a difference of conclusion between elimination and identification, meaning there were no false identifications or eliminations or sets where one examiner identified the items as having been fired in the same firearm and another examiner eliminated them as being fired in the same firearm. For one set, the five examiners that compared the items recorded inconclusive results when the ground truth conclusion was identification. The examiners recorded that the images were out of focus or too blurry to be of value for comparison.

Set 33 highlights how the examiners used the tools in the software to document their conclusions, which is valuable in understanding how different conclusions can be reached for the same image. In this set, only one examiner felt the need to add notes. This may be because the examiners are required to document the reason for inconclusive, but not necessarily an elimination. The examiners annotated the images marking the areas used to reach the conclusion. The software is flexible to allow for different ways to display the areas being considered when conducting a comparison and rendering a conclusion. Having this level of documentation is useful when examiners come to different conclusions because it allows reviewers to verify that they're looking at the same images, same quality, and same focus, as well as determine which areas they saw agreement or lack of correspondence. This gives technical reviewers insight into how the examiner arrived at their conclusion. This verification demonstrated that examiners consistently reach accurate comparison conclusions with 3D images, but it also highlighted limitations of the equipment and the workflow. VA DFS had intended to set the system up with minimal user interaction as it had with the robots in the DNA and toxicology laboratory. However, VA DFS discovered that although the system is designed for batch capture with minimal interaction, user interaction is sometimes needed to resolve any technical issues and an examiner needs to review the images for clarity before sending them on for evaluation and comparison. The scanner image is limited to the breech face and firing pin, not the entire head stamp area. Since the software only captures a partial area of the cartridge case, the verification study emphasized the need to be cognizant of sample placement to avoid sample switching and the need to have an overall image of the tray set up to verify correct sample placement.

The third verification study involved determining the limitations of 3D images to group cartridge cases for triage or sampling approach for NIBIN entry. The intent was to save resources by conducting a limited exam and selecting the best items for each group for NIBIN entry and search. This verification was conducted with fewer examiners to save resources and the previous studies had demonstrated the technology was suitable for comparison purposes. Two sets consisting of ten cartridge cases each were created from firearms that exhibited similar characteristics. Set one contained cartridge cases fired from a Ruger Model P95DC and from a Ruger Model P95. The second set contained images from a Glock Model 35 or a Glock Model 23. The five examiners were instructed to view the images and group the ones that could possibly have been fired from the same firearm. Class and individual characteristics were visible in the images, and both could be used to group the items.

As a result of the verification studies, VA DFS implemented the Cadre 3D scanner and the virtual comparison microscopy software for use in the administrative NIBIN sampling plan procedure. This procedure is intended to minimize examiner time and provide expedited investigative leads to the submitting law enforcement agencies. This approach is used for cases involving more than one cartridge case and the purpose of submitting the evidence is to determine if the evidence is linked to a previously submitted firearm or another crime scene. In the past, an examiner would conduct comparison exams on all cartridge cases. Those comparison results were then verified by another examiner, and one cartridge case from each group representing a different firearm was entered into NIBIN, wasting resources that could have been used on other cases. With Cadre, VA DFS can have a technician open, mark, and scan the cartridge case, freeing up examiner time. The firearms examiner conducts the comparisons on screen and groups images based on the visible characteristics. One cartridge case from each group is entered in and searched in NIBIN. As additional resources and funding become available, VA DFS plans to implement the use of 3D images for conclusive comparisons, not just triage or grouping cases in the future. The use of 3D images in firearms comparisons is a significant shift from traditional LCM and will require time, and an understanding of limitations to fully implement. There will likely be instances when traditional LCM is appropriate and other scenarios when 3D VCM would be more suitable; therefore, there will likely be a hybrid approach in the future.

DAY 3

FORENSIC INTELLIGENCE

Day 3 of the Firearm Policy and Practice Forum focused on the use of forensic intelligence for the disruption and prevention of gun crime. This included a keynote presentation by Thomas Chittum on the revolution of crime gun intelligence. This was followed by a panel moderated by Commander Stephanie Stoiloff of the Forensic Science Services Bureau of the Miami-Dade Police Department and included the following presentations:

- An overview of the OJP Forensic Intelligence (FOR-INT) Initiatives by Dr. Jonathan McGrath from NIJ and Dr. Basia Lopez of the Bureau of Justice Assistance (BJA)
- Incorporating forensic datasets in criminal investigations by Yaneisy Delgado from the Forensic Science Services Bureau of Miami-Dade Police Department
- NIBIN investigation overview by Lieutenant Branko Stojsavijevic and Ryan Orlovsky from the Milwaukee Police Department
- Using forensic data to solve and prevent crime by Dr. Michael Garvey of the Philadelphia Police Department Office of Forensic Science and Dr. Tara Garvey of the Philadelphia Police Department Intelligence Bureau
- Evidence screening in support of NIBIN by Jessica Ellefritz of the Phoenix Police Department Gun Crime Intelligence Center

Gun Crime Intelligence Revolution

IBIS is a crucial tool used by firearm examiners. IBIS is a computer-based system for comparing digital images of cartridge cases found at crime scenes against digital images of fired casings found at other crime scenes and test fired casings from weapons. IBIS is linked to NIBIN,[65] a national network consisting of linked ballistic imaging systems administered by the Bureau of Alcohol, Tobacco, Firearms and Explosives.[66] The network provides forensic laboratories and law enforcement agencies with a platform to upload and compare digital images of ballistic evidence to a national database to generate investigative leads. NIBIN's value comes from the fact that individual guns are sometimes used to commit multiple violent crimes. For example, a study conducted in New Jersey showed that in instances where there are two shooting events linked by ballistics through NIBIN, 50% of the time, a third shooting event using the same firearm occurred within 90 days. If NIBIN can link a common gun to multiple crime scenes, law enforcement can use that intelligence tactically to link multiple crimes to a common offender and strategically to identify patterns of gun crime and criminal networks by their use of common guns.[67] The firearm examiner is still responsible for making the final determination on identifications, but IBIS can link evidence.[68]

Despite its value, NIBIN is underutilized nationally. Of the laboratories that responded to the 2014 Census of Publicly Funded Forensic Crime Laboratories (CPFFCL), only 39% used the NIBIN database. Of the laboratories that performed firearm/toolmark functions, 66% used the NIBIN database in 2014.[39] Furthermore, despite the availability of the X3P data standard that could allow integration of 3D technologies into NIBIN, NIBIN is designed to only work with Ultra Electronics Forensic Technology Inc.'s (Ultra FTI) IBIS and uses a proprietary matching algorithm. To date the ATF has been unable to adopt the X3P data standard. The 2009 NAS report states that firearms examination is subjective, but X3P would provide a quantifiable way to produce ballistic source identification with an error rate, however X3P is currently interoperable with NIBIN.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

NIJ Forensic Intelligence (FOR-INT) Initiative: Using Forensics for Crime Disruption and Prevention

The National Institute of Justice (NIJ) is developing a forensics intelligence (FOR-INT) framework to assist state, local, and tribal law enforcement and forensic laboratories integrate forensic crime laboratory data into the criminal intelligence and analysis process to advance the disruption and prevention of crime through actionable intelligence. Forensic intelligence comprises the collection, organization, interpretation, and sharing of forensic case data in support of criminal investigations and intelligence procedures. Current practices use forensic data to solve and prosecute individual crimes. However, there are rich datasets amassed within state and local forensic laboratories. The FOR-INT approach uses these datasets to link and track crime patterns specific to local jurisdictions and ultimately reduce crime in those areas. Currently, forensic data are not commonly integrated into crime and intelligence analysis at the state and local levels in the United States. Switzerland and Australia appear to have stronger regional local efforts in comparison to the United States. The end goal of forensic intelligence is to increase the probability of linking cases and understanding crime in a holistic manner by integrating that forensic data into the intelligence processes.

NIJ's strategic research plan on policing emphasizes the importance of examining role and other related aspects of forensic science in investigative processes. Thus, the FOR-INT initiative closely aligns with this plan. Forensic specialists preserve, process, and send evidence collected at the crime scene to the forensic laboratory where the evidence is then tested, interpreted, and eventually reported with the purpose of informing the criminal case in court. Crime analysts focus on gathering, collecting, and analyzing existing data to find series, pattern, and analyze trends, identifying crime hotspots. They also look at research and analysis of long-term problems based on current and historical datasets. At the same time, intelligence analysts develop and link local intelligence based on the information they gather, which may include data necessary for producing actionable intelligence to link and track crime patterns and individuals within local jurisdictions. The goal of this analysis is to reduce crime in those areas.

Forensic intelligence data are used in the strategic, operational, and tactical levels of policing. This includes early identification of suspects and absolving innocent persons from suspicion. Intelligence-led policing (ILP) is a policing model built around the assessment and management of risk.[69, 70] Intelligence officers serve as guides to operations, rather than operations guiding intelligence.[71, 72] An example of an operational intelligence analysis is when law enforcement uses collected drug profiles to find connections between cases that were not suspected, which gives strategic insight into the magnitude and volume of the drug market, number of players in the region, and the relative risks of products to public. Tactical analysis results in on-the-go solutions such as directing patrol officers to the areas where drug dealers operate.

In December 2019, the Promising Practices in Forensic Intelligence report was issued as a work product of the Criminal Justice Coordinating Council, which is chaired by the Bureau of Justice Assistance, in partnership with the Department of Homeland Security's Office of Intelligence and Analysis and the Global Justice Information Sharing Initiative.[73] This report focused on developing a roadmap to provide law enforcement, intelligence functions and fusion centers with promising practices and recommendations on how to develop or enhance relationships between forensic laboratories and intelligence units to build out agency intelligence efforts. This report provides examples of ways to leverage laboratory data, laboratory results, and the analysis to augment intelligence operations. It documents specific examples on how laboratories and intelligence can work together to exchange information and provides a checklist for laboratory intelligence units to use that helps strengthen the relationships and collaboration. The report focuses on the need to create a joint understanding of laboratory and law enforcement requirements by learning about the laboratories' capabilities and services, the types of forensic data laboratories can provide. This will help improve evidence submission processes and data sharing agreements, and inform training on how laboratory data can be used for intelligence purposes.

NIJ also published an article in the NIJ journal on using forensic intelligence to combat serial and organized violent crimes.[69] The article gives an overview of the variety of laboratory disciplines that can be used in forensic intelligence—including firearms and toolmarks, DNA, seized drug and toxicology—and ways to develop a forensic intelligence program. The article also explored some of the existing activities at the local, state, and federal levels, such as crime gun intelligence centers supported by ATF and Bureau of Justice Assistance. The Drug Enforcement Administration (DEA) also has a drug signature program to link seizures by looking at the chemical characteristics of both the controlled substances (through isotope analysis and other distinct chemical characterization tools) and the cutting agents, adulterants, and other aspects of those drug materials to link up cases. Customs and Border Protection does forensic pollen analysis, which is used to develop pollen profiles to understand where an item or person of interest may have come from around the world and how it was transferred to the US border and entered the country. New Jersey is a good example of a state-level program that uses intelligence from diverse information sets including drug seizures; overdose data; and data from emergency rooms, first responders, and the prescription drug monitoring program to combat the opioid and current drug crises. Switzerland and Australia stand as international examples. Switzerland has had success with investigating crimes involving fraudulent identification documentation, as well as looking at burglaries and using forensic intelligence. Australia has had similar successes looking at illegal drug crimes and property crimes.

The FOR-INT initiative held two meetings of relevant federal agencies in January 2020 and April 2021. These meetings were followed by seven site visits that were conducted between fall 2020 and spring 2021. In Denver and Philadelphia, FOR-INT explored the general forensic disciplines used for intelligence purposes. Cuyahoga County, Cleveland, Ohio, and the surrounding areas, provided an examination of the forensic intelligence activities related to sexual assault kits. FOR-INT participants visited New Jersey to examine the drug monitoring initiative. In Milwaukee, FOR-INT explored crime gun intelligence center activities and social network analysis. Switzerland and Australia provided additional international examples. Switzerland provided examples at the local level, while Australia provided them at a regional level. NIJ is currently focused on developing and publishing a measurable forensic intelligence framework for implementation at state and local law enforcement agencies in the United States to inform prevention and disruption approaches using forensic laboratory data.

The 2016 paper from Bruenisholz and colleagues discusses guiding principles for developing a forensic intelligence approach.[74] Many of these principles have been echoed and exemplified in different ways from the site visits. NIJ has identified five guiding principal categories for the implementation of a forensic intelligence program. The first concerns collaboration and strengthening communications among the partners. The second is the importance of interoperable data and workflows, to ensure that the laboratory workflows and processes complement those of the law enforcement agencies. The third principle is that data integrity and understanding the interpretation and limitations of preliminary data and confirmatory data are crucial. The fourth is that forensic intelligence programs should be committed to the fair and impartial administration of justice and the use of intelligence both to identify suspects and to absolve innocent persons from suspicion. The fifth principle that gaining and maintaining support from all stakeholders will create a sustainable program.

To go along with the guiding principles, NIJ also identified five pillars that are important for implementation. These include organization, process, technology, capabilities and resources, and information sharing. Getting leadership buy-in and support, strategic management, and outreach to the appropriate stakeholder organizations is incredibly important. Some programs have embedded intelligence analysts in laboratory operations, but programs should to evaluate the best placement of staff and resources within the current organization charts and develop an understanding of where the people and resources are best suited to develop those effective enforcement strategies. Programs should also consider current and best practices, and consider modifying standard operating procedures where needed, conducting iterative internal impact assessments of the most effective use of data.

Using appropriate technology ensures that digital data are stored, shared, and communicated effectively. This can lead to timely detection of serial crimes, prioritization of forensic analysis, and linking unsolved crimes or even cold cases.

In terms of capabilities and resources, having the appropriate staff in the right places and ensuring that resources are allocated appropriately maximizes the effectiveness of the operation. Agency partners may wish to cross-train or create details or rotations within laboratories or the intelligence units. Use of capabilities and resources informs strategic planning, research, resource allocation, and real-time surveillance strategies. The goal is to match the staffing and research capacity to meet that demand.

Information sharing and working with partners to develop the appropriate policies can include the development of appropriate memorandum of understanding, agreements to improve links, and understanding trends through this data sharing. There is also potential to improve local and federal partnerships by exchanging information and connecting forensic laboratory stakeholders with the other intelligence and investigative units.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Incorporating Forensic Datasets in Criminal Investigations

In 2016 the Miami-Dade Police Department (MDPD), began a pilot project with one high gun-crime district where ShotSpotter was being implemented. The plan for the pilot was to maintain a 24-hour turnaround time after submission. In 2017, MDPD implemented two crime hotspots and a new initiative, the Violent Crime Task Force, which focused on deterring crime in hotspot areas. In 2019, MDPD expanded ShotSpotter to include three gun-crime hotspots. This expansion included hiring two new technicians for NIBIN entry and one intelligence analyst (IA) at the laboratory. Traditionally, an IA will mine data from various law enforcement sources as well as open sources, working directly with law enforcement personnel in different bureaus. MDPD had never had an IA look at the forensic data holistically. With an IA embedded in the laboratory creating actionable intelligence, MDPD realized that the crime laboratory produced a lot of information, but investigators were not understanding the laboratory reports, and therefore the forensic leads were not being used. In other words, the crime laboratory had not been yielding the actionable intelligence that it was capable of generating.

An IA embedded in the laboratory can act as the translator between scientists, investigators, and prosecutors. The IA presents a comprehensive investigative correlation of events based on NIBIN associations and other forensic data. The key to success was to enter all casings or representative casings submitted as evidence into NIBIN as soon as possible. All DNA profiles that meet CODIS guidelines also needed to be entered, and these forensic results needed to be cross-referenced between all the different disciplines to ensure that a complete picture was being drawn by the crime gun reports. The preliminary vision was that the IA would sit in the Forensic Services Bureau and examine contact and noncontact gunshot detection cases due to ShotSpotter, as well as focus on some of the violent crime cases. MDPD realized later that the IA approach was more to create an overall general picture of all the evidence and investigative leads combined.

The report format followed the Philadelphia CGIC Crime Gun Event (CGE) model but was expanded into "spider webs." A SharePoint site was created so law enforcement personnel could benefit from these findings by searching these reports. Each report is broken in three parts. The summary portion includes a victim's or subject's biographical data, date and time of the crime's occurrence, originating agency case number, and laboratory case number. The originating agency case number is important because often detectives do not have a link between their case number and the laboratory LIMS number. While they cannot search their system for a LIMS number, they can search their system for the originating agency case number, which is the one that allows them to look and find the report and see what evidence was available. It is also a way for an IA to be able to read the incident fully to know what happened, who the subject is, who the victim is, and so forth. MDPD also ran a criminal history and arrest records for victim and subjects. The lead investigator's name is included in the summary section and contact information if it is a case from outside the jurisdiction just to make it easier for the detectives. Social media information, if available, any known associates of any of the subjects listed, or any gang affiliations are also listed in the summary section.

The second part of the report is the spiderweb analysis, which is a visual presentation of the information provided in the summary. The report includes any photos of the subject, victim, or vehicle because sometimes a vehicle or the subject would be described in a NIBIN link or a NIBIN case. The same vehicle might have been the victim's, or it might have been the subject being described. The third part of the report is the mapping analysis. This analysis includes a geographical overview of the cases discussed, with an interactive map link to view the map more clearly. Giving investigators this view allows them to track how guns are moving, and the interactive map gives officers the ability to take a closer look at what buildings are close by and who hangs out in that area to help them visualize the area, rather than just giving them numbers and metrics.

Some of the programs that an IA uses to gather intel are the NIBIN Enforcement Support Systems (NESS), generated by ATF, and the Law Enforcement Information Exchange system (LinX), maintained by the US Naval Criminal Investigative Service (NCIS). LinX allows the IA to read outside agencies' originating reports in real time. At MDPD, the IA also includes in these reports any arrest intake photos, incident reports, civil court documents, open-source information, criminal history databases, license plate readers, facial recognition software, and information from more than 30 databases and other law enforcement programs that are used to give the officer a full picture of what has occurred in the case. At MDPD, the implementation of STARLIMS allowed the IA to cross-reference the disciplines more easily. Previously if the IA wanted to see if CODIS hits in a case, they would have to go to the biology section the laboratory, locate the physical folder, read the report to find the information they were looking for, for example to see if any DNA was picked up from the swab of a gun, or from clothing left on a crime scene. However, with STARLIMS, once the report is available the IA can access it immediately. MDPD recently had a case where a firearm was recovered in a stolen vehicle, and while they were unable to recover latent prints on the firearm or any DNA profiles, they were able to lift the prints from inside the stolen vehicle to identify potential suspects that were using the vehicle and the firearms at the same time. An IA will also use open-source intelligence, which are data available to the general public and include social media platforms such as Twitter, TikTok, Snapchat, Facebook, Instagram, and others as well as internet searches using the subject's names to access obituaries to establish family relationships and genealogy, open court records, marriage certificate, news media, and any other information that is available to the public.

The information provided in the CGE report is based on information pulled from a variety of sources and consolidated by the IA. Since the IA never handles or tests the evidence and the written report is only provided for investigative purposes, the IA does not have to testify in court as the written report is only provided for investigative purposes. The IA may include potential subjects in a report if mentioned in previous cases, as well as intended victims if they might be a potential subject in another case. For example, if cases A, B, and C are all connected and case B has video footage of a subject, it is possible they are the same subject in case A. Even if the detectives are unable to make an arrest, case B may have enough evidence to arrest the subject, and possibly recover the firearm and establish a link to the other crimes. Intended victims are included because they may potentially be the one common factor in multiple shootings (e.g., the gun may belong to their gang rival or could be the gun that the intended victim is using in self-defense). The CGE report may include the criminal history of other victims and subjects involved and a color-coded indication of gang affiliations or narcotic activity.

The CGE report gives an officer a complete overview of the information in the NIBIN LinX report and the data available for him to follow up. The subject's data are also tied to police communication codes such as 52, 55 since the report is for law enforcement. If a subject has past involvement in justice system, the IA may include a 52 to indicate a past issue with narcotics, or a 55 to indicate a past weapons violation. The mapping analysis section is also color coded to allow investigators to differentiate between different guns and their movement. Thus, the investigator can see if a suspect tends to stay in the same region or whether the areas are close enough to each other to indicate a territorial battle between gangs. The mapping analysis sections also include an interactive link that allows the investigator to zoom in on a house and see the satellite images of the surroundings to help jog the investigator's memory. This information could be beneficial to helping visualize the area, recollecting what gangs are active in this area, recalling other violence occurring in the area, and knowing what groups tend to congregate in the area. The CGE reports are available on an internal SharePoint to enable investigators to easily pull past reports to link suspects in current cases with past events or to help with future investigations in cases where a suspect may reoffend upon release from custody. The investigative intelligence is also shared with neighboring jurisdictions.

Milwaukee Police Department NIBIN Investigation Overview

The Milwaukee Gun Crime Intelligence Center (GCIC) was established in 2014 by the Milwaukee Police Department (MPD). NIBIN laboratory and CGIC laboratory are assigned to MPD's fusion division, which is part of the Criminal Investigations Bureau. The CGIC is an interagency collaborative initiative that focuses on the collection, management, and analysis of crime gun data from systems such as NIBIN and eTrace. MPD has had NIBIN since 2013 and is investigating the addition of a second machine due to NIBIN's success in helping to solve crime. The population of the city of Milwaukee is just under 600,000 in the city and approximately 1.5 million people living within the Greater Milwaukee area, which has an area of approximately 96 square miles. There are seven districts in the city of Milwaukee. Milwaukee has historically been one of the most segregated cities in the country with the area having the highest segregation index based on the census results for 2013–2017.[75] The highest levels of segregation exist in the northern metro areas, which are predominately Black; the near south side, which is predominately Hispanic; and the far south side, which is predominately white.

There are approximately 1,600 sworn law enforcement officers working for the MPD who work closely with federal agencies, including the ATF and FBI, as needed.

Currently, MPD is organized into three bureaus, including the seven-district Patrol Bureau, the Criminal Investigations Bureau, and the Administrative Bureau. The Criminal Investigations Bureau, which consists of seven divisions including homicide, violent crimes, sensitive crimes, general crimes, forensics, fusion, and special investigations. The violent crimes division investigates shootings, stabbings, and other violence cases that fall short of homicide. Sensitive crimes include sexual assaults and crimes against children. The general crimes division investigates all other cases. The forensics division captures the fingerprints and crime scene processing for any evidence. The fusion division houses the NIBIN laboratory. Special investigations consists of the fugitive apprehension unit and other units that help out with daily enforcement and the direct patrol missions throughout the city. MPD ended 2021 with 197 homicides, up from 89 in 2019 and 173 in 2020. MPD has also seen a rise in nonfatal shootings from 452 in 2019 to 793 in 2021. CGIC's mission is to prevent gun violence through consistent production of timely, precise, and actionable intelligence concerning gun crimes to identify armed, violent offenders for investigation and targeted enforcement.

There are two ways MPD processes evidence for NIBIN—by the gun or the casings. MPD processes every handgun or long gun rifle that is recovered by officers and has been linked to a crime as well as all casings that are recovered at the scenes. Every gun leaves identifiable marks on the bullet or the casing, which are then used to compare with other crimes and other recovered casings. That information is used to investigate who could be responsible for the crime. When an incident occurs, the MPD recovers casings either through a ShotSpotter notification or directly from incidents such as shootings or homicides. In the case of a shooting or homicide, the casings are sent directly to the NIBIN

Laboratory for entry into the NIBIN database to start the investigation. In other instances, such as when casings are recovered from a ShotSpotter notification where there are no suspects or victims present at the scene, the casings will be entered into inventory by the officers at their respective districts and will be retrieved by property control officers who will bring them to the NIBIN laboratory to be processed by the NIBIN technicians and entered into BrassTrax.

To generate investigative leads, images of the cases are acquired using the IBIS BrassTrax and run against the database using an algorithm to score and rank potential matches to other casings in the system. The system provides the NIBIN technician with a ranked list of potential matches to be reviewed by the firearms examiner. Once the examiner determines whether there is a match, a case number is generated and assigned for follow-up investigation. MPD uses four color-coded designations for the level of the NIBIN cases. Red denotes the most violent crimes and situations in which the identified suspect is a part of the known criminal enterprise or a serial shooter. Black is assigned to the investigator that took the original case to determine what follow up may be needed. Blue is assigned to district-level personnel to monitor specific locations or individuals and direct patrol missions for specific locations where these incidents may be occurring. A green NIBIN case means that there have been no investigative leads generated. NIBIN cases are also assigned one of three levels (low, medium, or high), based on solvability and priority.

In 2018, the MPD responded to a ShotSpotter alert where two 9mm casings were recovered. No other evidence was recovered from the scene. In 2019 MPD responded to a call to an incident that involved shots fired into an occupied vehicle. The victims in this incident had been waiting at a red light when the suspect pulls up and gets out of a black Acura sedan with tinted windows and discharged multiple rounds into the victim's vehicle. After the shots were fired, the suspect returned to the black Acura and fled the scene. Four 9mm casings were recovered at the scene however, officers were unable to locate a video of the incident and had a very limited description of the suspect. The casings were entered into NIBIN but there was no connection to the 2018 case. Later that same day deputies from the Milwaukee County Sheriff's Office (MCSO) conducted a traffic stop on a black Acura TL for excessive window tint. The driver fled the vehicle but was apprehended and the deputies recovered a Taurus 9mm pistol from the vehicle. The pistol was taken to the laboratory for test fire and MPD was able to correlate and match the test fires to the casings left at the scenes of other cases.

In January 2020, officers responded to a double shooting in which a passing vehicle stopped and a suspect confronted a group that had hit the car with a snowball. The suspect got out of the vehicle and shot two of the group. The officers were able to get a description of the suspect and the vehicle and recover nine 9mm casings and a surveillance video of the vehicle. Approximately a week later, the MPD officers attempted to stop a beige Nissan Altima, but the driver fled the scene leading to a high-speed pursuit and ultimately the driver crashed the vehicle. The driver failed the field sobriety test and was in possession of a

Glock 9 mm pistol. An eTrace came back to the suspect, who had purchased the firearm on December 23, 2019.

Crime guns that are recovered are placed in inventory and sent to the forensic unit to process for DNA and fingerprint evidence. NIBN technicians then collect the firearms for test fire and collect casings for triage and entry into the NIBIN database to compare with open cases and look for any matches. All of details regarding the firearms are also entered into the NIBIN case management system for future follow-up. MPD holds a meeting once a week to review any NIBIN cases that have leads. The meetings include the detective bureau, ATF agents, and officers from each district. Each case is reviewed and follow-up assignments are made based on the location where the incident occurred. If a case has a possible suspect that is a prolific offender, MPD will use ATF Task Force Officers to help with the case. Analysts from the fusion division may also put together information to help with investigative leads and prosecution during trial to show the connection between the suspect gun or casing and link to multiple offenses.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Using Forensic Data to Solve and Prevent Crime

The Technical Intelligence Program in the Philadelphia Police Department

There are 80 fusion centers across the United States that were developed after 9/11 to increase terrorism prevention and information-sharing between federal, state, and local agencies. The National Fusion Center Association provides a network to connect these centers for the purpose of sharing intelligence, best practices, and challenges. Every state has at least one fusion center, and some major urban areas also have regional fusion centers. The Delaware Valley Intelligence Center (DVIC) is the regional fusion center for the Philadelphia Metropolitan Area. DVIC is also part of a larger intelligence bureau within the Philadelphia Police Department (PPD), which utilizes the intelligence cycle from evidence collection to analysis to dissemination of investigative leads for all criminal activity, allowing operations and intelligence to drive each other. DVIC servs as the focal point within the state and local government for the gathering and analysis and sharing criminal threat-related information between federal and state, local, tribal and territorial, and private-sector partners. The concept of linking intelligence and operations has applications beyond counterterrorism. While not all-inclusive, the concept can also be applied to counterproliferation, counternarcotics, counterintelligence, and cybersecurity. Counternarcotics is important because drug trafficking can destabilize a region, resulting in socioeconomic problems and political consequences. Counterintelligence and cyberattacks can include gathering counterintelligence against the government, economic espionage (e.g., stealing technology), and cybercrime. These are examples of modern foreign intelligence but these examples are similar to what law enforcement experiences at the local level.

The concept of intelligence and operations driving each other can be applied to law enforcement. Homeland Security threats and terrorism are similar to the increasing violence in our cities. The current levels of gun violence deserve an equal level of coordinating intelligence response to protect local neighborhoods. Counternarcotic operations are used to combat the distribution of narcotics and the opioid epidemic. Counterproliferation may seem more of a global operation, as counterproliferation of nuclear material and other weapons of mass destruction are not typically seen at the local level. However, at the local level, counterproliferation focuses on crime guns. At the local level, cyberintelligence is used in the exploitation of electronic devices to gather investigative leads and to defend against the cyberattacks that routinely affect businesses and local governments. Therefore, law enforcement can learn from established practices within the global intelligence community.

The intelligence cycle begins with a plan that sets the requirements for the information and data that should be collected. Information is then collected, whether it be from human sources, open sources, or scientific and technical data. Next the data and information are processed and organized, and the analyst uses these to produce actionable intelligence. The intelligence is disseminated for use in operations, which in turn often generates new requirements for collection or intelligence sources. The Central Intelligence

Agency is an example of an intelligence organization and user of the intelligence cycle. The agency is designed so that its components complement one another and enable the intelligence cycle. Using this proven intelligence model, the PPD was able to identify similar synergistic relationships in their organization. The frameworks are similar on the federal and local levels. In the PPD, the Office of Forensic Science provides the scientific and technical solutions. The Intelligence Bureau provides analysis and finished intel products or provides the computer support for these processes and patrol operations, and the Detective Bureau is the operational force of the PPD.

A key to producing actionable intelligence is using multiple types of information. The main types of intelligence used in law enforcement are human intelligence (HUMINT), forensic and technical intelligence (TECHINT), and open-source intelligence (OSINT). HUMINT is intelligence produced through the collection of information from human sources. TECHINT includes intelligence that is derived from scientific data and analysis of physical, biological, digital, or other forms of evidence. OSINT comprises intelligence from publicly available information including social media, newspapers, internet searches, videos, and commercial databases. When considering technical intelligence sources for law enforcement and public safety, data and conclusions from the forensic laboratory are the obvious sources and many sources of data that can be obtained from forensic analyses that can be used in the production of intelligence. NIBIN correlation data, eTrace data, firearms trends, and other data associated with a firearm shooting scene or suspected shooter can be used to layer data to strengthen the intelligence product. These discussions are great in theory or if an agency has unlimited resources, but local law enforcement is often faced with significant resource challenges.

The PPD had to refocus efforts toward an intelligence-driven use of forensic data. In some cases, this meant modifying workflows and prioritization schemes to support intelligence-led policing. The forensic analysis procedures are often court-driven and focus on investigative support when requested. Additionally, reporting is usually limited to the specific submitter or investigator for that submission. This process can limit the usefulness of information. A modified version of the process that emphasizes intelligence allows the data to be used for investigative purposes while preparing for court. An example of this is the NIBIN program in which the workflow was modified to screen all ballistic evidence in a timely manner to generate leads that could be used by the investigators and the intel analysts. Full forensic examinations are reserved for those cases that require a higher level of analysis or a formal laboratory report record. This allows a higher number of cases to be triaged through NIBIN, producing more leads. This forensic data can be combined with other sources of forensic data to produce a technical intelligence package, which is then further enhanced with other forms of intel. The cornerstone of crime intelligence is often technical intelligence from eTrace, NIBIN, cell phone data, DNA, prints, and so forth. However, these concepts apply to other types of investigations as well. Drug analysis data and stamp logo data from the chemistry unit can also be used to produce drug intelligence products. These include products such as drug trend analysis, officer

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

safety bulletins, and emerging trends. These products have implications for counter-narcotics operations and public health responses.

Philadelphia is experiencing an increase in gun violence and homicides. In 2020, there were 499 homicides. As of December 2021, there were 530 homicides and over 2,000 nonfatal shootings. From an evidence perspective, the PPD has also seen an increase in evidence collection related to gun violence. Before 2019, the PPD received approximately 3,600 crime guns each year as evidence. In 2019, that number increased to 4,258 and in 2020 to 4,989. By December 2021, PPD had received over 5,000 crime guns with the potential for 6,000 by the end of the year. Approximately 9%–10% of the recovered crime guns are privately made firearms (PMFs). In addition to the crime gun evidence, the Office of Forensic Science (OFS) Firearms Identification Unit (FIU) receives between 35,000 and 50,000 items of ballistic evidence, the majority of which are fired cartridge cases recovered at crime scenes. PPD implemented a comprehensive crime gun intelligence program starting with a pilot project in 2012 that grew each year until it was citywide by 2016. The Crime Gun Intelligence Center (CGIC) is a collaborative effort between the ATF and the PPD OFS, Intelligence Bureau, and Detective Bureau. The ATF interacts directly with the PPD bureaus to gather and disseminate information. This program relies on comprehensive collection and submission of all firearms and ballistic evidence, as well as a timely analysis of the data resulting from workflow changes made at the FIU. These changes allowed production of timely NIBIN results that could be used to provide an initial lead report by the CGIC, followed by more in-depth analytic report produced by the Intelligence Bureau. These reports provide additional intelligence, such as gang associations, social media information, and camera footage, as well as an assessment, recommendations, and intelligence gaps. Additionally, NIBIN leads and related intel reports are a common discussion point during PPD CompStat (Computer Statistics) and shooting review meetings.

The data from firearms reports can also be used to provide emerging firearms trends. One example is regarding PMFs or "ghost guns." In Philadelphia, there has been an increase in PMFs. Five years ago, the PPD recovered fewer than 10 PMFs. In 2019, PPD recovered 95 PMFs, and in 2020 that number grew to 250 PMFs, making up just over 5% of the crime guns recovered that year. By December 2021, PPD had already recovered 530 crime guns, accounting for an almost 10% increase over 2020. In relation to crime gun intelligence and other technical intelligence operations, PPD uses latent prints as an additional source of intelligence in crime gun intelligence, overlaying NIBIN data with any related latent print hits that may identify potential shooters and gang affiliations that can be used to solve current cases and help deploy resources to prevent a future shooting if there is a risk of retaliatory shootings.

While this information is traditionally used in specific investigation it has the potential for larger linkage analysis when used for the inclusion or exclusion of specific subjects. Focusing on crime intelligence, the current policy of the PPD and OFS is to swab every crime gun for DNA. Although the resources are not available to process all these samples, PPD and OFS can decide when to use this capability on a case-by-case basis. When analyzed, these data can be used by the IA to generate a more detailed link analysis. In addition to CODIS, the PPD has established the Delaware Valley Investigative DNA database (DVID), which is a regional DNA database. The database contains data from rapid DNA analysis, samples from convicted offenders, arrestees, volunteered consent samples, and other crime scene evidence of data that are not in CODIS. This program allows the PPD to leverage the resources of both CODIS searches and regional investigative leads. Recently, the PPD has had success with recovering DNA from fired cartridge cases that provided actionable intelligence to operations. Again, there is a capacity issue based on the quantity of Fired Cartridge Cases (FCCs) recovered each year. However, the OFS is working on automated workflows to increase this level of technical intelligence.

Exploitation of electronic devices provides another layer of intelligence that can be used to support crime intelligence. The data obtained from this exploitation can be analyzed in systems such as PenLink, which provides the capability to make connections between unknown people, places, and things. Camera evidence is another type of evidence that is essential to technical intelligence and can be layered with NIBIN data. As an example, NIBIN data can provide a connection between a homicide and a nonvictim shooting. Camera footage from a doorbell camera from a nonvictim shooting, which otherwise may not have been could provide a link to the homicide, depending on the hits' association in time and geography. The key is to layer different types of forensic data with other forms of intelligence to produce actionable intelligence for operations. It is also critical that the intelligence products be user-friendly so that investigators are more likely to follow the leads. When building a technical intelligence program, agencies should develop a collaboration between forensics, intelligence, and investigations through cross-training; develop methods to share data; integrate multiple types of TECHINT, HUMINT, and OSINT whenever possible; and develop a coordinated review process to maintain technical accuracy of the products.

Evidence Screening in Support of NIBIN

Phoenix is the fifth largest city in the United States with 518 square miles. The Phoenix Police Department (PHXPD) has approximately 2,700 sworn personnel and processes about 430 firearms per month (5,200 per year) and 180 bags of cartridge cases per month (2,200 per year) in our NIBIN program. The NIBIN program is set up outside of the crime laboratory and is instead embedded within investigative unit as part of the violent crimes bureau and includes both civilian and sworn personnel who are responsible for all NIBIN processing. This includes all of the crime gun intelligence, investigation of NIBIN leads, test firing of firearms, swabbing of firearms, follow up of the investigations, and the actual arrests of individuals identified by NIBIN leads. To process evidence through NIBIN, PHXPD has a close partnership with the crime laboratory (PHXPDCL) to make sure that the policies and procedures fall in line with the crime laboratory and their policies and procedures as well as the County Attorney's Office for the purpose of prosecution.

When Phoenix established the CGIC, they wanted to it to benefit investigations. Also, although investigations are a priority, Phoenix also wanted to look at the big picture to disrupt the cycle of firearms-related violence by focusing on investigating violent individuals rather than just focusing on individual incidents of violence. A study conducted by the US Sentencing Commission found that, of the 3,446 firearms offenders they followed, 68.1% were rearrested for a new crime during the eight-year follow up period and that the firearms offenders reoffend sooner than nonfirearms offenders.[76] In Phoenix, the top 2% of violent criminals are the repeat offenders that commit a large majority of the firearms-related crimes. Feedback from the investigators indicated PHXPD was using valuable resources to arrest the same individuals repeatedly. When establishing the CGIC Phoenix, wanted to focus on reducing firearms-related crime and saving lives but also to keep those violent repeat offenders off the street.

In examining the end goal of the CGIC, Phoenix worked with the county attorney's office and prosecutors to establish a vertical prosecution model and a system of investigating firearms-related crimes in that attempts to identify links between different incidents early in the investigations. This would allow prosecutors to submit charges for multiple incidents at once or be able to request that the judge allow a bond. PHXPD recognized the only way to put that structure into place with incidents linked together was to have a process in place to swab firearms evidence for DNA and enter it into NIBIN within the two business days. The ATF minimum required operating standards (MROS) require that all firearms evidence is processed in NIBIN within two business days of the time that the evidence is received.

At that time, PHXPD was sending firearms for prints and DNA in almost every investigation. To meet the ATF MROS, PHXPD needed a new prescreening process that would support NIBIN entry but would also fit in with the needs of the prosecutors, the court, and investigations. PHXPD collaborated with the county attorney's office and the PHXPDCL, bringing all three groups to the table with data and statistics to enable an open conversation about the challenges they faced and possible solutions.

Due to the volume of requests the PHXPDCL receives for firearms evidence in standard weapons cases, PHXPDCL did not have the capacity to be able to process latent prints, DNA, and firearms requests and conduct a full forensic examination unless the case was already scheduled for trial. The group developed a prescreening method that allowed evidence to be entered into NIBIN ahead of examining the latent prints or analyzing DNA. Focused heavily on want versus need, the group realized that although the agencies would appreciate having prints and DNA on everything, that was not always necessary. For example, all PHXPD officers have body cameras and if there are three officers with body cameras showing that an individual pulled a firearm out of the waistband of a suspect, then latent prints and DNA are not needed to identify the suspect. The prosecutor and courts can use the body camera footage and written reports from the officers who state they pulled the gun out of a waistband to prove possession of the firearm.

Figure 41. Cartridge casing evidence collection containers used by PHXPD

Image Courtesy of Jessica Ellefritz presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

These conversations led to establishing an evidence screening method to facilitate the multidisciplinary processing of firearms, DNA, and latent print examinations. Using this method allows Phoenix to enter evidence into NIBIN to establish links at the beginning of an investigation so that the prosecutor can use these links during the initial court appearance. There may only be probable cause on three of five cases linked in NIBIN, but the prosecutor can show an escalation of violence and repeat offenses to ask that bond be denied and the suspects held in jail until the case goes to trial. To achieve this goal, the first change Phoenix made was to impound all evidence casings in plastic snap top vials. Training the officers and crime scene specialists to impound the casings individually with the head stamp facing up allows the technicians to triage and prescreen all casings without handling them. As Figure 41 shows, PHXPD can hold individual evidence containers under the microscopes to prescreen the casings, identify how many firearms are involved in the incident, triage which sample is best for NIBIN entry, and then identify the samples that will be put into NIBIN. This allows the investigators to impound all casings from one scene and package them together. The individual vials can also be numbered to indicate where a casing was located at the scene, allowing recreation of the scene based on location guns without touching the casings.

PHXPD created a tier system for triage that prioritizes cases including homicides, officer-involved shootings, and other high-profile incidents. In those first-tier cases, PHXPD will swab the casing ahead of time to preserve DNA before NIBIN entry. If there are multiple casings at a scene, PHXPD will only swab one casing that is removed from the vial for entry. All the other casings are kept in the vials and are available for further prints and DNA to allow examiners to pull prints and DNA off the casings in the future. For any case that falls outside of the top priority tier, single casings will be swabbed every time, recognizing that that the one casing is their only opportunity to potentially get DNA. If lower tier cases involve multiple casings that are verified during prescreening to be from the same gun, PHXPD enters one without swabbing, leaving the other casings available for prints and DNA. This allows the CGIC to expedite the process by using triage to prioritize obtaining DNA in high-profile cases and ensure that lower profile cases have the option to analyze prints and DNA if needed later in the investigation or for trial

Next, the CGIC needed to address a tiered process for examining guns. First-tier guns are processed on the same day, whereas second-tier guns typically come to CGIC the following day. The top tier again includes homicides, officer-involved shootings, and anything else considered high profile. In those situations, either a detective or one of the civilian personnel will go out to the scene, while crime scene investigators handle the scene, take photos, and handle evidence collection. At the conclusion of their collection, investigators pass custody of the guns and the casings to one of the CGIC sworn detectives or civilian examiners, who bring the evidence directly back to the office for examination. In most cases, the CGIC can have casings and guns test fired, swabbed, and entered into NIBIN with results and leads before the scene of a homicide has been cleared. PHXPD also typically does this in weapon cases if an individual is in custody. This way, the CGIC can provide the detective with any lead information while they have that individual in an interview room, allowing them to adjust and modify their interviews with the individual based on the preliminary NIBIN results. In these instances, the detectives can indicate whether they need prints or not. In homicide cases, if the detective determines that they need prints, evidence will always go to the laboratory first and undergo the print process in the PHXPDCL and then be routed to the CGIC. If the detective does not request prints, then the CGIC will proceed with their NIBIN process. The CGIC uses bleach and clean butcher paper and only swab the textured areas of the firearm that they have to touch to do a test fire. These areas include the grip, the trigger, and the front site (because that is typically where the DNA congregates), as well as the rear textured area of the slide. All the smooth surfaces, any accessories, the magazine, any live ammo, and any holsters are still available for further prints and DNA. In addition to not touching any part of the gun that is not used to perform a test fire, the CGIC also use shoe covers over the port of the shooting tank so the gun will not touch the shooting tank at the time of the test fire. These additional steps prevent contamination and allow the CGIC to test fire the firearm while the crime scene is still live or while the suspect is still in custody, allowing intelligence to be incorporated into the interviews and investigation in real time.

Second-tier guns, which are those impounded by patrol, are submitted to the property room. The following day, the CGIC will collect the firearms from the property room and process them. In those situations, the CGIC determines whether a gun was found on a person or not, and the patrol officers include that information in the description of the impound. The CGIC will document the gun and can proceed with test firing it using their clean techniques. In case evidence comes back later that indicates that a weapon was used in a homicide, the gun is still available for prints and DNA. If the gun is found in any other location than on a person, even if it is in a backpack on their back, the CGIC consider that not in the person's possession (e.g., the subject could claim that they did not know there was a gun in the backpack or that the backpack did not belong to them). In these situations, the CGIC will swab the firearm before test firing it to preserve the DNA in case it is needed later. The CGIC marks the areas of the firearm that are swabbed on a sketch in case the firearm needs go to the PHXPDCL for DNA or prints after an investigation. This allows the laboratory to see the exact areas of the firearm that were previously swabbed.

Figure 42 shows how the analyzed areas are documented using diagrams of the of different types of firearms, highlighting the areas of the firearm that have been swabbed. Police assistants that are assigned to the CGIC place the firearm on a clean surface and swab the textured areas while wearing gloves. The swabs are placed directly into the envelopes and the envelopes are marked with the areas the swabs came from. The swabs are placed into the same evidence packaging as the firearm to go back to the laboratory if additional testing is needed. All test fires are performed on site at police headquarters. There are five civilians who are responsible for swabbing, test firing, and entering firearms and casings into NIBIN. There are also seven detectives that handle the investigations. As NIBIN leads are generated, they are assigned to a detective who works as a liaison with individual case agents from different incidents, regardless of whether they are Phoenix-based incidents or incidents originating in a neighboring agency's jurisdiction. The CGIC detectives compile all the different reports from each contributing agency and package them together for submission to the county attorney's office. Civilians conduct 90% of the NIBIN processing, but the CGIC detectives are also cross-trained to perform NIBIN entries in case they need to go to a scene, bring evidence back, test fire the firearm, enter it into NIBIN, and get it correlated while the investigation is active.

Figure 42. Diagrams of where the firearm was swabbed for DNA by the CGIU

Image Courtesy of Jessica Ellefritz presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

Changing the procedures to process evidence outside of the crime laboratory environment, changing the shell casings storage vials, and the changing the process for swabbing and test firing process firearms has had a great impact on the CGIC's ability to process evidence. In November 2017, PHXPD was processing about 21% of the eligible items coming into custody, with an average turnaround time of 125 days to enter an item of evidence into NIBIN. Of that 21%, only about 2.5% of it was entered within two business days. Because of the turnaround time and the small volume of evidence processed, PHXPD was only averaging about 20 leads a month, which was not very much for the fifth largest city in America. As of November 2021, PHXPD processed and entered 100 percent of the eligible evidence into NIBIN and approximately 98 percent was entered within two business days, with an average turnaround time of one day. PHXPD now averages 190 leads per month, with approximately 2,300 hundred leads for 2021, compared to 204 at the beginning of the program. PHXPD now has a formal policy that makes NIBIN processing mandatory. In 2017 PHXPD took as long as 255 days to process firearms evidence, but immediately after the policy went into place, the turnaround time dropped to between one to four days. The impact of the new process also resulted in an increase in the volume of NIBIN leads and the ability to push this intelligence out to investigators in real time.

In summary, the processing that occurs at the CGIC, outside of the crime laboratory, is preliminary and used for intelligence purposes. By providing investigators with preliminary NIBIN leads, the CGIC gives investigators information earlier in their investigations and makes the additional connections that they need to build probable cause. Without it, the evidence is sitting on a shelf waiting for the crime laboratory to process it, and they do not have the additional information from other linked cases. Investigators do not need confirmed leads to be able to move forward into investigating additional links and additional probable cause. The crime laboratory remains the experts in firearms analysis and the firearms examiner will confirm any leads and testify in court regarding to the leads and whether the gun was actually connected to the scene.

DAY 4

LEGAL ASPECTS

Day 4 of the Firearm Policy and Practice Forum focused on legal aspects, the admissibility of firearm and toolmarks evidence, and courtroom testimony. This included a keynote presentation by Erich Smith on the validation study of the accuracy, repeatability, and reproducibility of firearm comparison: the impact of the FBI firearms black box study. This was followed by a panel that was moderated by Raymond Valerio, JD, from the Queens New York District Attorney's Office and included the following presentations:

- An overview of firearm and toolmarks admissibility decisions after the PCAST report by Amie Ely, JD, of the National Association of Attorneys General
- Firearm and toolmark evidence admissibility challenges by Raymond Valerio, JD

Validation Study of the Accuracy, Repeatability and Reproducibility of Firearms Comparison: Impact of the FBI Firearms Black Box Study

The FBI Firearms Black Box study hypothesis was that trained/qualified firearms examiners can accurately determine source conclusions (repeatability) when applying the AFTE theory of identification and reproduce the same result(s) when later encountering the same comparison.[77] This study complements those conducted previously, adding important additional features. A previous study conducted on the accuracy of firearms examiners was generally viewed favorably by the President's Council of Advisors on Science and Technology (PCAST), but PCAST advised that additional, similarly designed "black box" investigations were required to establish foundational validity.[6] The present study implemented a fully randomized, open set, and double-blind design involving challenging comparisons of fired bullets and cartridge cases. To maintain double-blind conditions, the experimental study was conducted under contract by scientific staff at another organization, which sent randomized specimens to participating examiners and performed statistical analysis. Following a call for participants, volunteer active examiners were provided with 15 comparison sets of two known and one unknown cartridge cases fired from a collection of Beretta® and Jimenez® firearms and 15 comparisons sets of two known and one unknown bullets fired from Beretta® and Ruger® firearms. To minimize reproduction of marks, the ammunition selected for testing was Wolf Polyformance® 9mm Luger (9x19mm), with acrylic polymer-coated, steel cartridge cases and lead core, copper-plated, steel-jacketed bullets.

The firearms and ammunition selected for this study were purposely chosen because of their propensity to produce challenging and ambiguous test samples, creating difficult comparisons for examiners. The firearms, bullets, and cartridge cases used for the study were collected by researchers in the first laboratory and delivered to scientists in the contract laboratory, who then conducted the study and engaged in the generation and distribution of test packets, and collection and analysis of the data.

A total of 173 qualified examiners took part in the study. The participating examiners were asked to follow the provided instructions rather than adhere to their laboratory policies and were instructed not to discuss their results with anyone else in their laboratory. To further maintain the double-blind, "black box" nature of the study, the team associated with communicating with the examiners was not aware of the contents of each comparison set, and the experimental/analysis group was never aware of the examiners' identities.

The total number of comparisons carried out was 20,130, of which 8,640 were tested for accuracy, 5,700 were tested for repeatability, and 5,790 were tested for reproducibility. Definitive false positive error rate estimates that take examiner heterogeneity into account are 0.66% for bullets and 0.93% for cartridge cases. False negative error rate estimates are 2.87% (bullets) and 1.87% (cartridge cases). These estimates are based on data that include comparisons from barrels produced sequentially in time and those separated in the manufacturing process, rounds fired early in the life of a barrel and after many rounds had been fired, and rounds fired from both high and low cost-point firearms. Individual error rates within each of these categories have also been calculated and vary slightly from the overall average in ways that might be expected (e.g., higher error rates are seen for rounds widely separated in firing order than sequentially, lower cost point firearms have a higher false negative error rate than average). As in earlier studies, the majority of errors were produced by a relatively small number of examiners. The numbers found in the current study are generally consistent with the results reported in prior studies, and therefore constitute the foundational validation the PCAST report said was lacking. This study will impact the forensic science community by providing empirical measurements of the accuracy, repeatability, and reproducibility of analyses performed by firearms examiners, for cartridge case and bullet sample sets.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Five Years Later: An Overview of Firearms and Toolmarks Admissibility Decisions After the PCAST Report

The 2016 PCAST report announced new standards to determine whether forensic science disciplines are foundationally valid and whether particular experts in a scientific discipline should be permitted to testify.[6] The report applied these standards to certain forensic disciplines and made recommendations to federal prosecutors, judges, and federal agencies. In applying the standards, the report found that DNA analysis of both single-source and simple mixtures samples and latent fingerprints were foundational valid but found that DNA analysis of complex mixtures, firearms identification, bitemark analysis and forensic odontology, footwear analysis, and hair microscopy all lacked foundational validity. Since the PCAST report, there have been 94 admissibility decisions issued including 40 at the federal level and four courts of appeals including the 1st, 2nd, 7th, and 9th district courts, and 54 at the state level including the supreme courts of Connecticut, Kentucky, Maryland, Nebraska, and New York. The posture of these decisions can be found in Figure 43, but these decisions have been mostly at pretrial and appeal. The criteria used to examine these decisions were that there must be access to the written decision or oral transcripts and the decision must mention or relate to the PCAST report. These decisions covered a range of forensic disciplines and included 57 decisions regarding firearms and toolmarks. Of these 57, 29 were pretrial decisions, 26 were appeal, 1 was post-conviction, and 1 was midtrial.

Figure 43. Posture of post PCAST court decisions

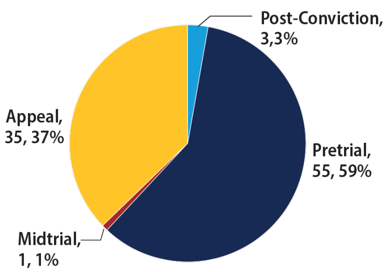


Image Courtesy: Amie Ely presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

The Firearms Appellate Decisions have mostly affirmed trial court decisions (22 decisions). Four of these found either if there was error it was harmless or noted there was error but also noted the error was harmless.[78-81] There were also four decisions that were reversed.

Figure 44 shows a map of the US District appellate courts. In 2017, the 9th circuit affirmed in *United States v. Johnson* (875 F.3d 1265) that the trial court and properly admitted the expert testimony that the test-fired bullet matched the bullet recovered from the crime scene “to a reasonable degree of ballistics certainty.”[82] Also in 2017, the 2nd circuit affirmed in *United States v. Gil* (680 Fed. Appx. 11) that there was no error when the court did not conduct a Daubert hearing and allowed the expert to testify “to a reasonable degree of certainty in the field of ballistics.”[83] In 2020, the 7th circuit also affirmed in *United States v. Brown* (973 F.3d 667) that the “defendants brought the PCAST report to the district courts attention, but the district court shoes not to give it dispositive effect and that choice was withing its set of options.”[84] In both the *Brown and Johnson* cases, the courts allowed the “reasonable degree of certainty” statement, but this is not currently permitted by the United States Department of Justice (DOJ) under the Uniform Language of Testimony and Reports (ULTRs) for testimony in federal courts.

Figure 44. Map of District Circuit Courts of Appeal

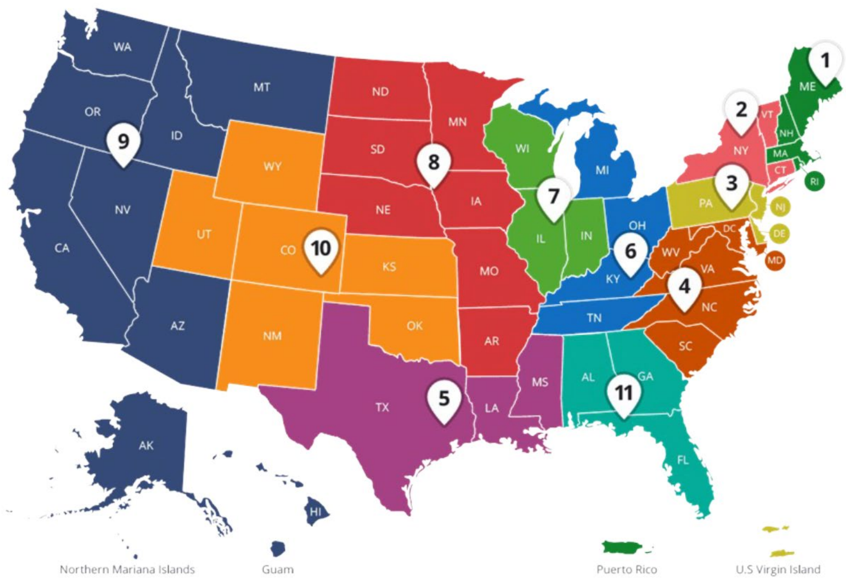


Image Courtesy: Amie Ely presentation at the 2022 Firearms and Toolmarks Policy and Practice Forum.

The three different types of conclusions an examiner can reach under the DOJ ULTRs for the firearms discipline include.

1. **Source identification:** "Source identification" is an examiner's conclusion that two toolmarks originated from the same source. This conclusion is an examiner's opinion that all observed class characteristics agree, and the quality and quantity of corresponding individual characteristics is such that the examiner would not expect to find that same combination of individual characteristics repeated in another source and has found insufficient disagreement of individual characteristics to conclude they originated from different sources.
2. **Source exclusion:** "Source exclusion" is an examiner's conclusion that two toolmarks did not originate from the same source.
3. **Inconclusive:** "Inconclusive" is an examiner's conclusion that all observed class characteristics agree, but there is insufficient quality and/or quantity of corresponding individual characteristics such that the examiner is unable to identify or exclude the two toolmarks as having originated from the same source. The basis for an inconclusive conclusion is an examiner's opinion that there is an insufficient quality and/or quantity of individual characteristics to identify or exclude.

There are some cases at the state level where PCAST has also been a factor. In *Williams v. Commonwealth*, the Kentucky appellate court affirmed that Daubert hearings are generally not required for firearms and toolmarks because the court can take judicial notice of the science of firearm and toolmarks evidence due to how long it has been admitted into the court.[85] In *State v. Wheeler*, the defense had argued at the pre-trial hearing that the government's expert was not qualified to opine that the seven shell casings found at the scene were fired from

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

the same gun. Based on the findings noted in the PCAST report, the trial court allowed extensive cross-examination but ultimately concluded that the witness was qualified to testify and render an opinion on the shell casings. The Nebraska Supreme Court affirmed this conviction noting that the expert's testimony was not prejudicial but rather supported the state's theory that the same gun fired all the shell casings and that the defense had sufficient opportunity to present a counter theory.[86] There were several other post-PCAST cases in state court where the intermediate court of appeals affirmed convictions citing that no Daubert or Frye hearing was needed and the firearms evidence and testimony was admissible, including *State v. Mills* in Missouri, *Williams v. Commonwealth in Kentucky, State v. Hatfield* in Washington, *People v. Rodriguez* in Illinois, and *State v. Allen* in Louisiana.[85, 87-90] There were also several appellate decisions that affirmed convictions after pretrial admissibility hearings, including *State v. Castro DeJesus*, *Williams v. Texas*, *State v. Boss* and *State v. Eaglin*.[91-94] There were also several court decisions that affirmed convictions even though there was plain error where the defense didn't argue to the trial court that the evidence should not have admitted including *State v. Oliver*, *State v. Griffin* and *Williams v. United States*.[95-97]

In *Abruquah v. State of Maryland*, the court remanded the case to have the trial court apply Daubert instead of Frye to review the expert testimony to reconsider its decision to admit toolmark evidence/testimony in light of *Rochkind v. Stevenson* (471 Md. 1, 236 A.3d 630 [Md. 2019]), which adopted the Daubert standard. The trial court concluded the testimony was still admissible under Daubert and the trial court affirmed the conviction.[98] In the *State v. Raynor* trial, the appellate court reviewed of defense claim that trial court erred in denying a Porter hearing to determine reliability of firearm evidence, and that more recently published NAS and NRC reports (*State v. Porter* [241 Conn. 57 (1997)]) found firearm and toolmark methodology unreliable. Appellate Court opined that those new studies do not nullify existing case law, the firearm expert explained both the methodology and limitations during testimony, and the jury had a chance to weigh evidence. Appellate Court stated that firearm evidence admissibility is well established, and affirmed judgment.[99] In two intermediate appellate court decision, the first being the *People v. Azcona*, the appellate court decided that the trial court had "committed multiple errors" related the firearms expert testimony. The court did not say that firearms and toolmarks evidence is inadmissible just that the manner in which it was presented in this case was in error and remanded the case.[100] In *State v. Ghiqliotty*, the New Jersey court of appeal affirmed the decision to have a Frey hearing for the admissibility of BULLETTRAX software, which is a relatively new software, but reversed the decision to require extensive discovery of the algorithms the software used and limitations on the expert testimony.[101]

As previously stated, there have been 29 pretrial/midtrial decisions that include 13 federal pretrial, 1 federal midtrial, and 16 state pretrial decisions. The motions filed included motions by the defense to exclude or limit the government's firearms and toolmarks expert testimony, motions by the government to exclude the defense expert testimony, and also a couple of civil cases. There are no full exclusions based solely on PCAST arguments, but limitations are relatively common. These have included either the expert or the government agreeing to use the ULTR language, or the court orders the expert to abide by certain limitations. There have been three pretrial decision to admit testimony at the federal level with no limitations including *United States v. Romero-Lobato*, *United States v. Chester*, and *United States v. Taylor.*[102-104] There have also been four federal cases where the government has proposed to limit testimony by adhering to the ULTR language, agreed to not use language such as "match," "to the exclusion of all others," reference any level of statistical certainty or to limit testimony to class characteristics.[105-108]

While the DOJ ULTRs contain limitations on an expert's testimony and reporting, these limitations do not make any substantive or material changes to the examiner's opinion of identification, exclusion, or inconclusive. Rather these limitations serve as guardrails to ensure the examiners conclusions remain within prescribed bounds and do not exaggerate their conclusions. The limitations also provide some context about the examiner's conclusion. A conclusion provided during testimony or in a report is ultimately an examiner's opinion and is not based on a statistically derived or verified measurement or on comparison to all other firearms or toolmarks. Therefore, in accordance with the DOJ ULTRs, an examiner shall not:

- Assert that a "source identification" or a "source exclusion" conclusion is based on the "uniqueness" of an item of evidence.
- Use the terms "individualize" or "individualization" when describing a source conclusion.
- Assert that two toolmarks originated from the same source to the exclusion of all other sources.
- Assert that examinations conducted in the forensic firearms/toolmarks discipline are infallible or have a zero-error rate.
- Provide a conclusion that includes a statistic or numerical degree of probability except when based on relevant and appropriate data.
- Cite the number of examinations conducted in the forensic firearms/toolmarks discipline performed in his or her career as a direct measure for the accuracy of a conclusion provided. An examiner may cite the number of examinations conducted in the forensic firearms/toolmarks discipline performed in his or her career for the purpose of establishing, defending, or describing his or her qualifications or experience.
- Assert that two toolmarks originated from the same source with absolute or 100% certainty or use the expressions "reasonable degree of scientific certainty," "reasonable scientific certainty," or similar assertions of reasonable certainty in either reports or testimony unless required to do so by a judge or applicable law.

Under the DOJ ULTR, if an examiner cannot identify or exclude a firearm as the source of a cartridge case or bullet, then the opinion is inconclusive.

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

Admission with Limitations

United States v. Tibbs: the firearm "cannot be excluded" as having fired the recovered cartridge case & "may have fired"

United States v. Adams: General rifling characteristics

United States v. Shipp & *State v. Burton*: "consistent with"

United States v. Davis & *United States v. Medley*: "consistent with"

United States v. Glynn & *United States v. Mouzone*: "more likely than not"

United States v. Williams II: No "unqualified" opinions

People v. A.M & *People v. Ross*: Class characteristics only

There are several decisions in which the federal court ordered the expert to limit testimony. These include *United States v. White*, where the court ordered that the expert "may not testify to any specific degree of certainty as to his conclusion that there is a ballistics match between the firearms seized…and those used in the various shooting incidents."[109] In the *United States v. Medley*, the court ruled that the expert could not testify that cartridge cases found at the scene of the crime were fired from the gun associated with the defendant or express a confidence level as to his opinion.[110] In *United States v. Shipp*, the expert was limited to testifying that bullet fragment and shell casing are "consistent with" being fired from recovered firearm and that the recovered firearm "cannot be excluded" as source of bullet and bullet fragment.[111] In both *United States v. Davis* and *United States v. Adams*, the court ruled that the expert cannot say "match" or that the cartridges were fired from the same firearm, and the expert in Adams was limited to testifying to general rifling class characteristics.[112, 113]

There are also several unpublished pre-trial decisions at the state level in Virginia, Kentucky, Missouri, Massachusetts, and Colorado where the court ruled to admit the firearms and toolmarks expert testimony without limitation.[114-119] There are also numerous decisions at the state level where the court ruled to admit the expert testimony of the firearms and toolmarks examiner with limitations. In the *State v. Terrell* hearing, the court permitted the use of the terms "common origin" when referring to the recovered casing having been fired from the same gun but did not allow the use of the terms "practical impossibility" that another firearm could have fired the casing.[120] In *United States v. Valdez*, the court permitted expert testimony that the bullets and casings were fired from the same firearm but limited testimony from exerting "100% certainty" or to the "exclusion of all others."[121] In *State v. Gibbs*, the court allowed the expert to testify to a "match" but could not say "100% certainty" or "to the exclusion of all others."[122] In *Abruquah v. State*, the court permitted the expert to opine that the bullet came from the gun but could not express "absolute or scientific certainty."[98] In *State v. Burton*, the court limited the expert to testify only that it is "consistent" that the shell casings found at the scene came from the gun in question.[123] In *United States v. Tibbs*, the court limited the expert to testifying that the "recovered firearm cannot be excluded as the source of the cartridge casing."[124] In *State v. Goodwin-Bey*, the expert was limited to testifying that the gun "could not eliminate" as the source.[125] In *People v. Ross*, the expert was limited to testifying only to the class characteristics.[126] There have also been at least one firearms case where the judge excluded a defense expert who was proposed to testify regarding the PCAST report.[127]

Federal Rule of Evidence 702 and the Federal Rule of Criminal Procedure 16

Firearms identification expert testimony has been admissible for over a century, with many states recognizing the reliability of firearms identification expert witnesses and admitted their testimony on multiple occasions in the early 1900s.[128-132] These cases predated the adoption of the Federal Rules of Evidence (FRE) 702 and the landmark US Supreme Court case of *Daubert v. Merrell Dow Pharmaceuticals Inc.* (509 US 579 [1993]). In Daubert, the Court established a set of five factors a judge could use to weigh the reliability of scientific evidence. The Daubert standard considers five factors for evidence admissibility including if the method/technique has been tested, has been subjected to peer-review and publication, has a known error rate, has existing and maintained standards, and is widely accepted within the scientific community. Since Daubert focuses on the reliability and scientific validity of a specific methodology, some courts felt that the Daubert standards should only be applied to scientific techniques that can be tested and did not apply to the admissibility of subjective evidence such as psychology or other nonphysical sciences. The US Supreme Court, however, upheld the district court's decision which then extended Daubert to cover all expert testimony. While the federal courts must adhere to the Daubert standard, each state has a Rule of Evidence that defines the who is qualified to give expert testimony. Most states have adopted the Federal Rules of Evidence (FRE) 702 which allows for the admissibility of evidence that is based on the experts scientific, technical, or other specialized knowledge that will help the trier of fact understand the evidence and does not specify that the evidence undergo Daubert. The adoption of FRE and Daubert did not affect the admissibility of firearms expert testimony.

Under the Daubert standard, the factors that are considered in determining whether the methodology is valid are:

- whether the theory or technique in question can be and has been tested;
- whether it has been subjected to peer review and publication;
- its known or potential error rate;
- the existence and maintenance of standards controlling its operation; and
- whether it has attracted widespread acceptance within a relevant scientific community.

Currently there is discussion surrounding proposed changes to FRE 702, including a small change proposed for bullet (d), which allows for a comment in the notes to the rule that would place additional emphasis on the fact that courts need to be look at this evidence to make sure that it should be admissible. This is important for firearms and toolmarks examiners because when the courts have examined questions of admissibility, it has traditionally been more of a question of the weight of the evidence. The courts have decided to admit the evidence, but the defense counsel will be able to cross examiner regarding crucial points of contention. However now courts may be looking more carefully to determine whether the evidence should be allowed in at all. The fact that the committee is highlighting that courts have this gatekeeping rule may cause courts to look a more carefully at evidence that they've been willing to allow in in the past. The rule also reminds courts that the evidence only needs to help the trier of fact, the jury, to understand the evidence and in this regard, it may increase the admissibility of expert testimony.

Federal Rules of Evidence (FRE) 702:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

- the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- the testimony is based on sufficient facts or data;
- the testimony is the product of reliable principles and methods; and
- the expert has reliably applied the principles and methods to the facts of the case.

There is also a proposed addition to the footnote that would be appended to Federal Rule 702. The amendment is pertinent to the testimony of forensic experts in both criminal and civil cases. The note suggests that forensic experts should avoid assertions of absolute or 100 percent certainty or to a reasonable degree of scientific certainty and that the methodology is subjective and thus

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

potentially subject to error. The note also suggests that, when possible, a judge should receive an estimate of the known or potential rate of error of a methodology based on studies that reflect how often the method produces accurate results. The next note that is particularly important states that expert opinion testimony regarding the weight of feature comparison evidence, evidence that a set of features corresponds between two examined items, must be limited to those inferences that can reasonably be drawn from a reliable application of the principles and methods. This note empowers judges to make sure that the expert does not overstate the evidence and to place some of the limiting decisions in the previously discussed trial court decisions. Finally, the note does not bar testimony that comports with substantive law requiring opinions to a particular degree of certainty. Therefore, if a court chooses to require experts to use that reasonable degree of certainty language, the note will not overturn that decision but rather, as the note states earlier, it is not a good idea to use that language.

The proposed changes to the Federal Rule of Criminal Procedure (a)(1)(G) would require the government to abide by a time to make disclosure and require courts to set that time. That time has to be sufficiently before a trial to provide a fair opportunity for the defendant to meet the government's evidence. The amendment also includes more information about what content needs to be included in the Rule 16 disclosure. It must include a complete statement of all opinions that the government is going to be eliciting from the witness; the basis and reason for those opinions; and the witness's qualification, including a list of all publications authored in the previous 10 years, list of all other cases in which the witness has testified as an expert or by deposition during the previous four years. The witness is generally required to sign the disclosure. Similar requirements are in the note that governs what defense attorneys need to provide to the government. Several states have analogous rules to the rule of criminal procedure 16. Therefore, if the rule is changed on the federal level, some of these states may change their rules as well.

Firearm and Toolmark Evidence Admissibility and Current Challenges

In 2009, the National Academy of Sciences issued a report, questioning several forensic disciplines, including firearms and toolmark identification. While the survey did not make any recommendation on the admissibility of firearms and toolmarks, it did question the subjective nature of the discipline, the lack of standards, lack of a statistical foundation for the establishment of error rates, as well as the ambiguous terminology used in the AFTE Theory of Identification, such as when and examiner finds there is “Sufficient Agreement” to claim a bullet or cartridge case can be identified as having been fired by a specific firearm. The report also called for more empirical studies regarding firearms and toolmarks and noted a lack of an established method citing that the AFTE theory of identification is insufficient and subjective and that publications in the AFTE journal are not peer-reviewed. In response to the 2009 NAS report, firearms examiners and researchers conducted no less than six studies to assess the reliability of their discipline. Most of these studies were published in journals such as the *AFTE Journal* and *Journal of Forensic Sciences* (peer-reviewed). These studies reported a false positive rate of 1.01% or less, supporting the theory that when an examiner made the opinion or conclusion of “identification,” they were correct nearly 99% of the time.[12, 14, 18, 23, 24, 30, 32]

The PCAST report on forensic science was published in 2016. The authors of the PCAST report took issue with the six previously cited studies and questioned the foundational validity of these studies, finding some were “closed set, white box” studies where the answer is always present in the test-set, deductive reasoning could assist the examiner in the study, and the study did not assess the examiner’s abilities, only their methodology. PCAST determined that only open set, black box studies that analyzed the examiner and not the methodology—and where the answer may or may not be present—would suffice to create the “foundational validity” needed for admissibility of firearms expert testimony. The report also took issue with the high inconclusive rates found in the one study they felt was appropriately designed, the 2014 Ames (Baldwin) study.[14] Yet from a forensic science perspective, the Baldwin study was less than perfect. Unlike other firearms studies, it was not peer reviewed nor published. It omitted bullets from the study and only analyzed cartridge cases, so only half of the discipline was analyzed. It also did not employ consecutively manufactured barrels, making comparisons somewhat easier that some of the other studies that did evaluate consecutively manufactured slides.

Firearms examiners responded to this criticism in PCAST by designing and executing multiple black box studies, using the model prescribed by PCAST. The Kiesler Study in 2018, the Jaimie Smith Study in 2020, and the FBI’s black box study with AMES in 2021 (pre-print) answered the PCAST Report’s call for black box studies with open sets.[17, 22] These three black box studies reported consistent false positive rates of less than 1%, again demonstrating that when a firearms examiner makes an opinion or conclusion of “identification,” they are accurate more than 99% of the time even though these studies prohibited the examiner from using their laboratories’ QA/QC processes such as tech review or verification to screen for errors. Most studies also employed consecutively manufactured barrels or slides as part of their design, which create the best-known nonmatch available and inject the specter of subclass characteristics. Consequently, this makes identification much more difficult for the examiner than in casework where QA/QC is used, and different firearms are examined during an investigation signifying that the error rates in case work are likely less than the 1% noted in these studies.

One main concern of critics of firearms and toolmarks examination involves error rate. Some critics of firearm and toolmark identification were unsatisfied with the consistently low error rate of 1% or less and claim that “inconclusive” decisions in some of these studies, which can run as high as 30% of all responses, should not be treated as a nondecision. Instead, inconclusive decisions should be treated as errors, thus making the error rate exponentially higher. However, these critics do not account for the fact that inconclusive results are a feature of firearms and toolmark identification. Nor do they account for the fact that these studies did not permit the use of QA/QC or verification. Yet most studies used consecutively manufactured barrels or slides, which makes the studies more difficult than actual casework. Inconclusive results are a facet of all forensic science and feature in comparison disciplines, including latent prints, trace evidence, and DNA. Inconclusive results enable examiners to achieve an exceptionally low rate of error by providing an inconclusive result. Reducing the number of false positive conclusions reduces the rate of error. Counting an “inconclusive” finding as an error essentially eliminates it as a possible result, leaving an examiner with only “identification” and “elimination” conclusions. This does not accurately reflect how forensic science is conducted and would require examiners to speculate or guess whether some comparisons constitute an identification. Inconclusive findings allow examiners to make more reliable opinions when issuing identifications and exclusions by giving them an option for reporting inconclusive when they do not have enough feature information to reach a definite identification or exclusion.

In FSI Synergy, Drs. Dror and Scurich argue that inconclusive decisions are errors since ground truth is known.[133] They argue that inconclusive decisions in these validation studies allow participants to skip the answer. However, if examiners were reporting inconclusive too often, the sensitivity would be very low. Yet the sensitivity in the Keisler and Baldwin studies are approximately 99 percent.[14, 22] Todd Weller and Dr. Max Morris responded to Dror and Scurich by making the point that “insistence to classify all decisions as correct or error is overly simplistic.”[134] Dror and Scurich responded to Weller and Morris but didn’t address their points.[135] Biedermann and Kotsoglou added on their critique to Dror and Scurich in the same journal.[136] Inconclusive is a legitimate conclusion in firearm and toolmark analysis. Inconclusive decisions are found in many scientific fields, including DNA analysis with mixture interpretation, analytical, and stochastic thresholds. Inconclusive is also exhibited in drug chemistry, pathology, and other fields. Study designers do not review every casing and bullet such as in the Baldwin study. Also, certain firearms and ammunition do not mark consistently, and every study participant gets a different set to examine in these studies. Without examining every set, there is no way to know if one set marked better than another. Moreover, inconclusive decisions are an indication that an examiner is being cautious, and this is reflected in casework. Finally, inconclusive decisions do not lead to wrongful convictions and are therefore not the problem.

PCAST counted inconclusive in the denominator, not as false positive error. They state that false positive rates are based on the proportion of conclusive examinations. The Baldwin study examined 2,178 different source comparisons that included 22 false positives, 1,421 eliminations, and 735 inconclusives, which calculates a false positive rate of 1.01%. Using the PCAST methodology, they remove the 735 inconclusives, and the false positive rate slightly goes up to 1.5%. The Dror/Scurich methodology adds the inconclusives to the numerator, which balloons the false positive rate to 34.8%. The only way to get the false positive error rate above 5% is by adding inconclusives to the numerator, as Dror and Scurich suggest. It is also important to note that none of the validation studies consider technical review which is part of the workflow when conducting case work.

$$\text{Baldwin et al. FPR} = \frac{22}{[22+1421+735]} = 1.01\%$$

$$\text{PCAST FPR} = \frac{22}{[22+1421]} = 1.5\%$$

$$\text{Dror/Scurich FPR} = \frac{22+735}{[22+1421+735]} = 34.8\%$$

https://doi.org/10.3768/rtipress.2022.cp.0014.2204

A Path Forward

This is a critical time for pattern examination forensic science disciplines. This forum highlighted the investments that NIJ, NIST, and the broader scientific community have made to improve the science and support the reliability of the discipline since the PCAST report. Innovated workflows and emerging technologies are allowing investigators to use firearms and toolmark evidence to provide exploratory leads early in the criminal investigation. Allowing law enforcement agencies to stop the cycle of gun crime earlier and strengthen sentencing for those that commit violent crime. Researchers continue to work on integrating metrology into firearms and toolmark examination to enhance the subjective nature of the discipline by adding scoring functionality to better define the weight of the evidence.

Challenges to forensic science are to be expected, and these challenges are not a unique to firearm toolmark evidence. All forensic disciplines, including DNA analysis, experience these challenges. Although there have been dozens of courts that have approved current DNA analysis technology, yet the community still must regularly respond to challenges to the evidence. The challenges make the attorneys better because they must focus on the most salient issues and help the forensic community continue to question and study the state of the science to improve the accuracy and reproducibility of the science through continuous improvement. The community should accept the challenges and answer them with the science.

REFERENCES

1. Grinshteyn E, Hemenway D. Violent death rates in the US compared to those of the other high-income countries, 2015. Prev Med 2019;123:20–6. PubMed

2. Crime in the United States; Federal Bureau of Investigation. 2020. https://crime-data-explorer.app.cloud.gov/pages/downloads

3. Abiedalla Y, DeRuiter J, Clark CR. Product ion tandem mass spectrometric differentiation of regioisomeric side-chain groups in cathinone derivatives. Rapid Commun Mass Spectrom 2016;30(14):1713–21. PubMed

4. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. Strengthening forensic science in the United States: a path forward. Washington (DC): National Academies Press; 2009. Available from: https://www.ojp.gov/pdffiles1/nij/grants/228091.pdf

5. National Commission on Forensic Science. Recommendation to the Attorney General: technical merit evaluation of forensic science methods and practices. 2016. Available from: https://www.justice.gov/archives/ncfs/page/file/905541/download

6. Forensic Science in Criminal Courts. Ensuring Scientific Validity of Feature-Comparison Methods. Executive Office of the President: President's Council of Advisors on Science and Technology; 2016. Available from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf

7. National Research Council. Ballistic imaging. Washington (DC): The National Academies Press; 2008.

8. Butler J, Iyer H, Press R, Taylor M, Vallone P, Willis S. NIST Scientific foundation reviews. NIST Interagency/Internal Report (NISTIR). Gaithersburg (MD): National Institute of Standards and Technology; 2020. https://doi.org/10.6028/NIST.IR.8225

9. Pauw-Vugts PW, A. & Øren, L. & Pfoser, L. FAID2009: Proficiency test and workshop. 2013;45:115–127.

10. Peterson JL, Markham PN. Crime laboratory proficiency testing results, 1978-1991, II: resolving questions of common origin. J Forensic Sci 1995;40(6):1009–29. PubMed

11. Bunch S, Murphy D. A comprehensive validity study for the forensic examination of cartridge cases. AFTE J 2003;35(2):201–3.

12. Fadul TG, Gabriel A, Wilson E, Stoiloff S, Gulati S. Empirical study to improve the scientific foundation of forensic firearm and tool mark identification utilizing consecutively manufactured Glock EBIS barrels. NCJ Number 244232. 2013 Dec. https://www.ojp.gov/pdffiles1/nij/grants/244232.pdf

13. Lyons DJ. The identification of consecutively manufactured extractors. AFTE J 2009;41(3):246–56.

14. Baldwin D, Bajic S. Morris, M. and Zamzow, D. A study of false-positive and false-negative error rates in cartridge case comparisons. NCJ 249874. Defense Forensics Science Center. 2014 Apr. Available from: https://www.ojp.gov/ncjrs/virtual-library/abstracts/study-false-positive-and-false-negative-error-rates-cartridge-case

15. Murphy D. CTS error rates, 1992–2005 firearms/toolmarks. Presented at the 41st Association of Firearm and Tool Mark Examiners (AFTE) Training Seminar, Henderson, NV. 2010 May 5. Available from: https://afte.org/uploads/documents/swggun-cts-1992-2005.pdf

16. DeFrance CS, Van Arsdale MD. Validation study of electrochemical rifling. AFTE J 2003;35(1):35–7.

17. Smith JA. Beretta barrel fired bullet validation study. J Forensic Sci 2021;66(2):547–56. PubMed

18. Fadul T. An empirical study to evaluate the repeatability and uniqueness of striations/impressions imparted on consecutively manufactured Glock EBIS gun barrels. AFTE J 2011;43:37–44.

19. Walters A. Results of the latest ENFSI firearms identification proficiency test (FAID 2015). Denver (CO): Association of Firearm and Toolmark Examiners Annual Training Conference; 2017.

20. Smith E. cartridge case and bullet comparison validation study with firearms submitted in casework. AFTE J 2005;37(2):130–5.

21. Mayland B, Tucker C. Validation of obturation marks in consecutively reamed chambers. AFTE J 2012;44(2):167–9.

22. Keisler MA, Hartman S, Kilmon A, Oberg M, Templeton M. Isolated pairs research study. AFTE J 2018;50(1):56–8.

23. Smith TP, Smith AG, Snipes JB. A validation study of bullet and cartridge case comparisons using samples representative of actual casework. J Forensic Sci 2016;61(4):939–46. PubMed

24. Fadule T, Hernandez GA, Stoiloff S, Gulati S. An empirical study to improve the scientific foundation of forensic firearm and tool mark identification using 10 consecutively manufactured slides. AFTE J 2013;4(4)5:376–93.

25. Lilien R. Firearm forensics black-box studies for examiners and algorithms using measured 3D surface topographies. NIJ Report. 2019 May.

26. Kerkhoff W, Stoel RD, Mattijssen EJ, Berger CE, Didden FW, Kerstholt JH. A part-declared blind testing program in firearms examination. Sci Justice 2018;58(4):258–63. PubMed

27. Cazes MG, Goudeau J. Validation study results from Hi-Point consecutively manufactured slides. AFTE J 2013;45(2):175–7.

28. Mattijssen EJ, Witteman CL, Berger CE, Brand NW, Stoel RD. Validity and reliability of forensic firearm examiners. Forensic Sci Int 2020;307(110112).

29. Thompson SG, Cásarez NB. Solving Daubert's dilemma for the forensic sciences through blind testing. Houst Law Rev 2020;57(3):617.

30. Stroman A. Empirically determined frequency of error in cartridge case examinations using a declared double-blind format. AFTE J 2014;46(2):157–74.

31. Mattijssen EJ, Witteman CL, Berger CE, Zheng XA, Soons JA, Stoel RD. Firearm examination: examiner judgments and computer-based comparisons. J Forensic Sci 2021;66(1):96–111. PubMed

32. Hamby JB, Brundage DJ, Thorpe J. The identification of bullets fired from 10 consecutively rifled 9mm Ruger pistol barrels: a research project involving 507 participants from 20 countries. AFTE J 2009;41(2):99–110.

33. Chumbley LS, Morris MD, Bajic SJ, Zamzow D, Smith E, Monson K et al. Accuracy, repeatability, and reproducibility of firearm comparisons part 1: accuracy (2021) https://arxiv.org/abs/2108.04030

34. Law EF, Morris KB. Evaluating firearm examiner conclusion variability using cartridge case reproductions. J Forensic Sci 2021;66(5):1704–20. PubMed

35. Monkres J, Luckie C, Petraco ND, Milam A. Comparison and statistical analysis of land impressions from consecutively rifled barrels. AFTE J 2013;45(1):3–20.

36. Hamby JE, Norris S, Petraco ND. Evaluation of Glock 9 mm firing pin aperture shear mark individuality based on 1,632 different pistols by traditional pattern matching and IBIS pattern recognition. J Forensic Sci 2016;61(1):170–6. PubMed

37. Gutowski S. Error rates in the identification sciences. Forensic Bull 2005;23:23–9.

38. Wilson-Wilde L, Smith S, Bruenisholz E. The analysis of Australian proficiency test data over a ten-year period. Forensic Sci Policy Manage 2017;8(1–2):55–63. https://doi.org/10.1080/19409044.2017.1352054

39. Bureau of Justice Statistics. Data collection: Census of publicly funded forensic crime laboratories. 2014. Available from: https://bjs.ojp.gov/data-collection/census-publicly-funded-forensic-crime-laboratories

40. Muehlberger CW. Col. Calvin Hooker Goddard 1891–1955. J Crim Law Criminol Police Sci 1955–6; (46):103–4.

41. Senin N, Groppetti R, Garofano L, Fratini P, Pierni M. Three-dimensional surface topography acquisition and analysis for firearm identification. J Forensic Sci 2006;51(2):282–95. PubMed

42. Duez P, Weller T, Brubaker M, Hockensmith RE 2nd, Lilien R. Development and validation of a virtual examination tool for firearm forensics. J Forensic Sci 2018;63(4):1069–84. PubMed

43. Chapnick C, Weller TJ, Duez P, Meschke E, Marshall J, Lilien R. Results of the 3D virtual comparison microscopy error rate (VCMER) study for firearm forensics. J Forensic Sci 2021;66(2):557–70. PubMed

44. Swofford H, Campod C. Implementation of algorithms in pattern & impression evidence: a responsible and practical roadmap. Forensic Sci Int Synergy 2021;3:100142. https://doi.org/10.1016/j.fsisyn.2021.100142

45. American Academy of Forensic Sciences. Standard for topography comparison software for toolmark analysis. First edition. ANSI/ASB Standard 062. 2021. Available from: https://www.aafs.org/asb-standard/standard-topography-comparison-software-toolmark-analysis

46. Chu W, Song J, Vorburger TV, Thompson R, Silver R. Selecting valid correlation areas for automated bullet identification system based on striation detection. J Res Natl Inst Stand Technol 2011;116(3):647–53. PubMed

47. Chen Z, Song J, Soons JA, Thompson RM, Zhao X. Pilot study on deformed bullet correlation. Forensic Sci Int 2020;306:110098. PubMed

48. Zhang H, Gu J, Chen J, Sun F, Wang H. Pilot study of feature-based algorithm for breech face comparison. Forensic Sci Int 2018;286:148–54. PubMed

49. Weller T, Brubaker M, Duez P, Lilien R. Introduction and initial evaluation of a novel three-dimensional imaging and analysis system for firearm forensics. AFTE J 2015;47(4):198–208.

50. Song J, Vorburger T. Proposed bullet signature comparisons using autocorrelation functions. Toronto (Canada): Proceedings of National Conference of Standards Laboratories; 2000. Available from: https://www.nist.gov/publications/proposed-bullet-signature-comparisons-using-autocorrelation-functions

51. Lightstone L. The potential for and persistence of subclass characteristics on the breech faces of SW40VE Smith & Wesson sigma pistols. AFTE J 2010;42(4):308–22.

52. Song J. Proposed “congruent matching cells (CMC)” method for ballistic identification and error rate estimation. AFTE J 2015;47(3):177–85.

53. Johnson M, Cole F, Raj A, Adelson H. Microgeometry capture using an elastomeric sensor. ACM Trans Graph 2011;30:139–44.

54. Johnson M, Adelson E. Retrographic sensing for the measurement of surface texture and shape. 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009;1070–7. https://doi.org/10.1109/CVPR.2009.5206534

55. Knowles L, Hockey D, Marshall J. The validation of 3D virtual comparison microscopy (VCM) in the comparison of expended cartridge cases. J Forensic Sci 2022;67(2):516–23. PubMed

56. Jones NS, Fornaro E, editors. (2021). 2021 National Institute of Justice Forensic Science Research and Development Symposium. RTI Press Publication No. CP-0013–2104. https://doi.org/10.3768/rtipress.2021.cp.0013.2104

57. Lilien R. Evaluation of 3D virtual comparison microscopy for firearm forensics within the crime lab. National Criminal Justice Reference Service (NCJRS), NCJ 255942. 2020 Dec. Available from: https://www.ojp.gov/ncjrs/virtual-library/abstracts/evaluation-3d-virtual-comparison-microscopy-firearm-forensics

58. Gun violence archive. 2021 Nov 16. Available from: https://www.gunviolencearchive.org/

59. Biasotti AA. A statistical study of the individual characteristics of fired bullets. J Forensic Sci 1959;4(1):34–50.

60. Biasotti AM, Moran BR. Firearms and toolmark identification. In: Faigman DK, Saks M, Sanders J, editors. Modern scientific evidence: the law and science of expert testimony. Volume 4. St. Paul (MN): Thomson Reuters/West; 2008.

61. Bunch SG. Consecutive matching striation criteria: a general critique. J Forensic Sci 2000;45(5):955–62. PubMed

62. Creppage KE, Yohannan J, Williams K, Buchanich JM, Songer TJ, Wisniewski SR et al. The rapid escalation of fentanyl in illicit drug evidence in Allegheny County, Pennsylvania, 2010-2016. Public Health Rep 2018;133(2):142–6. PubMed

63. Song J, Vorburger TV, Chu W, Yen J, Soons JA, Ott DB et al. Estimating error rates for firearm evidence identifications in forensic science. Forensic Sci Int 2018;284:15–32. PubMed

64. Vergeer P, van Es A, de Jongh A, Alberink I, Stoel R. Numerical likelihood ratios outputted by LR systems are often based on extrapolation: when to stop extrapolating? Sci Justice 2016;56(6):482–91. PubMed

65. US Department of Justice Office of the Inspector General Audit Division. The Bureau of Alcohol, Tobacco, Firearms and Explosives' National Integrated Ballistic Information Network Program. Audit report 05-30. 2005 Jun. Available from: https://oig.justice.gov/reports/ATF/a0530/index.htm

66. The Bureau of Alcohol. Tobacco, Firearms and Explosives. Minimum required operating standards for national integrated ballistic information network sites. 2019. Available from: https://www.atf.gov/firearms/docs/undefined/mrosnibinsitesauditstandardspdf/download

67. The National Crime Gun Intelligence Governing Board. Crime gun intelligence—Disrupting the shooting cycle. A best practices guide for implementing successful crime gun intelligence programs: 2018. Available from: https://crimegunintelcenters.org/wp-content/uploads/2018/09/CGI-Manual-Best-Practices-ATF-27-AUG-18.pdf

68. King W, Wells W, Katz C, Maguire E, Frank J. Opening the black box of NIBIN: A descriptive process and outcome evaluation of the use of NIBIN and its effects on criminal investigations. Final report. 2013. Available from: https://www.ojp.gov/pdffiles1/nij/grants/243875.pdf

69. Lopez BE, McGrath JG, Taylor VG. Using forensic intelligence to combat serial and organized violent crimes. 2020 Oct 21. Available from: https://nij.ojp.gov/topics/articles/using-forensic-intelligence-combat-serial-and-organized-violent-crimes

70. Ratcliffe JH. Intelligence-led policing. 2nd ed. London: Routledge; 2016.

71. Criminal Intelligence Coordinating Council (CICC). Promising practices in forensic lab intelligence. 2019. Available from: https://it.ojp.gov/GIST/1211/Promising-Practices-in-Forensic-Lab-Intelligence

72. Ross A. Elements of a forensic intelligence model. Aust J Forensic Sci 2015;47(1):8–15.

73. Global Advisory Committee. Promising practices in forensic lab intelligence. 2019 Oct. Available from: https://bja.ojp.gov/library/publications/promising-practices-forensic-lab-intelligence

74. Bruenisholz E, Prakash S, Ross A, Morelato M, O'Malley T, Raymond MA et al. The intelligent use of forensic data: an introduction to the principles. Forensic Sci Policy Manag 2016;7(1-2):21–9.

75. Brookings. Black-white segregation edges downward since 2000, Census shows. 2018. Available from: https://www.brookings.edu/blog/the-avenue/2018/12/17/black-white-segregation-edges-downward-since-2000-census-shows/

76. US Sentencing Commission. Recidivism among federal firearms offenders. 2019 Jun 27. Available from: https://www.ussc.gov/research/research-reports/recidivism-among-federal-firearms-offenders

77. Monson KL, Smith ED, Bajic SJ. Planning, design and logistics of a decision analysis study: The FBI/Ames study involving forensic firearms examiners. Forensic Sci Int Synerg. 2022 Feb 19;4:100221. https://doi.org/10.1016/j.fsisyn.2022.100221

78. People v. Perez, 2019 Cal. App. Unpub. LEXIS 4165 (Cal. Ct. App. June 20, 2019).

79. Morones v. State, 2017 Tex. App. LEXIS 10964 (Tex. Ct. App. Nov. 22, 2017).

80. Williams v. United States, 210 A.3d 734 (D.C. Ct. App. 2019).

81. Gardner v. United States, 140 A.3d 1172 (D.C. Ct. App. 2016).

82. United States v. Johnson, 875 F.3d 1265. (9th Cir. 2017).

83. United States v. Gil, 680 Fed. Appx. 11. Docket No. 16–524-cr. (2d Cir. 2017).

84. United States v. Brown, 973 F.3d 667 (7th Cir. 2020).

85. Williams v. Commonwealth, No. 2019-CA-000187-MR, 2020 Ky. App. LEXIS Unpub 222 (Ky. Ct. App. Mar. 27, 2020).

86. State v. Wheeler, 956 N.W.2d 708 (Neb. Sup. Ct. 2021)

87. State v. Mills, 623 S.W.3d 717 (Mo. Ct. App. 2021).

88. State v. Hatfield, No. 77512-0-1, 2019 Wash. App. LEXIS 3027 (Wash. Ct. App. Dec. 2, 2019).

89. People v. Rodriguez, 106 N.E.3d 436 (Ill. App. Ct. 2018).

90. State v. Allen, No. 2017 KA 0306, 2017 La. App. Unpub. LEXIS 325 (La. Ct. App. Nov. 1, 2017).

91. State v. Castro DeJesus, 436 P.3d 834 (Wash. Ct. App. 2019).

92. Williams v. Texas, No. AP-77053, 2017 Tex. Crim. App. Unpub. LEXIS 906 (Tex. Ct. Crim. App. Nov. 1, 2017) (unreported opinion).

93. State v. Boss, 577 SW3d 509 (Mo. Ct. App. 2019).

94. State v. Eaglin, 239 So.3d 1001 (La. Ct. App. 2018).

95. State v. Oliver, No. A-5140-16T1, 2020 N.J. Super. Unpub. LEXIS 339 (N.J. Super. Ct. App. Div. Feb. 18, 2020) (per curiam).

96. State v. Griffin, 268 N.C. App. 96 (N.C. Ct. App. Oct. 15, 2019).

97. Williams v. United States, No. CF1-18032-10 (DC Ct. App. Jan. 21, 2016) (unpublished opinion).

98. Abruquah v. State, 471 Md. 249 (Md. 2020) (per curiam).

99. State v. Raynor, 254 A.3d 874 (Conn. 2018).

100. People v. Azcona, 58 Cal.App.5th 504 (Cal. Ct. App. 2020).

101. State v. Ghigliotty, 463 N.J. Super. 355 (N.J. Super. Ct. App. Div. 2020).

102. United States v. Romero-Lobato, 379 F. Supp. 3d 1111 (D. Nev. 2019).

103. United States v. Chester, No. 13 Cr. 774 (N.D. Ill. Oct. 7, 2016) (unpublished order).

104. United States v. Taylor, No. 15–20152 (E.D. Mich. Mar. 24, 2016) (transcript of oral ruling).

105. United States v. Harris, 502 F.Supp.3d 28 (D.D.C. 2020).

106. United States v. Hunt, 464 F.Supp.3d 1252 (W.D. Okla. 2020).

107. United States v. Johnson, No. 16 Cr. 281 (PGG), 2019 US Dist. LEXIS 39590 (SDNY Mar. 11, 2019).

108. Ricks v. Pauch, No. 17–12784, 2020 US Dist. LEXIS 50109 (E.D. Mich. Mar. 23, 2020).

109. United States v. White, No. 17 Cr. 611 (RWS), 2018 US Dist. Lexis 163258 (SDNY Sept. 24, 2018).

110. United States v. Medley, No. 17 Cr. 242 (S.D. Md. Apr. 24, 2018) (unpublished opinion contained in transcript of motion hearing).

111. United States v. Shipp, 422 F.Supp.3d 762 (EDNY 2019).

112. United States v. Davis, No. 4:18-cr-00011, 2019 US Dist. LEXIS 155037 (W.D. Va. Sept. 11, 2019).

113. United States v. Adams, 444 F.Supp.3d 1248 (D. Or. 2020).

114. State v. Acoff, No. 16-F-43 (W. Va. Cir. Ct. Sept. 23, 2016) (unpublished order).

115. Commonwealth v. Brock, No. 18-CR-00069 (Ky Circuit Court Jul. 30, 2021) (unpublished order).

116. State v. Kennedy, No. 1516–CR00258–01 (Mo. Cir. Ct. Mar. 31, 2017) (unpublished order).

117. People v. Purpera, No. 2016 Cr. 7798 (Colo. Dist. Ct. Aug. 12, 2018) (unpublished decision).

118. Commonwealth v. Legore, No. SUCR 2015–10363 (Mass. Super. Ct. Nov. 17, 2016) (unpublished decision).

119. Commonwealth v. Hernandez, No. SUCR 2014–10417, 2015–10384 (Mass. Super. Ct. Dec. 21, 2016) (unpublished decision).

120. State v. Terrell, No. CR170179563, 2019 Conn. Super. LEXIS 827 (Conn. Super. Ct. March 21, 2019).

121. United States v. Valdez, No. 2016 CF1 002267 (D.C. Super. Ct. Jan. 30, 2018) (unpublished order).

122. State v. Gibbs, No. 1819003017, 2019 Del. Super. LEXIS 639 (Del. Super. Ct. Dec. 9, 2019).

123. State v. Burton, No. CR23–150831 (Super. Ct. Conn. Feb. 1, 2017) (oral ruling).

124. United States v. Tibbs, No. 2016–CF1-19431, 2019 D.C. Super. LEXIS 9 (D.C. Super. Ct. Sept. 5, 2019).

125. State v. Goodwin-Bey, No. 1531–CR555–01 (Cir. Ct. Mo. Dec. 16, 2016) (unpublished).

126. People v. Ross, 68 Misc. 3d 899 (N.Y. Sup. Ct. June 30, 2020) & People v. A.M., 2020 NYLJ LEXIS 1401 (N.Y. Sup. Ct. June 30, 2020).

127. Commonwealth v. Mulazim, No. 15 Cr 0592 (Ky. Cir. Ct. May 2018) (unsigned order).

128. Commonwealth v. Best, 62 N.E. 748, 750 (Mass. 1902).

129. State v. Clark, 196 P. 360, 367–368 (Ore. 1921).

130. Laney v. United States, 294 F. 412, 416 (D.C. Cir. 1923).

131. Evans v. Commonwealth, 19 S.W. 2d 1091, 193–1099 (Ky. Ct. App. 1929).

132. State v. Campbell, 239 N.W. 715, 719–725 (Iowa 1931).

133. Dror IE, Scurich N. (Mis)use of scientific measurements in forensic science. Forensic Sci Int Synerg 2020;2:333–8. https://doi.org/10.1016/j.fsisyn.2020.08.006 PubMed

134. Weller TJ, Morris MD. Commentary on: I. Dror, N Scurich "(Mis)use of scientific measurements in forensic science" Forensic Science International: Synergy 2020 http://dx.doi.org/10.1016/j.fsisyn.2020.08.006. Forensic Sci Int Synerg 2020;2:701–2. https://doi.org/10.1016/j.fsisyn.2020.10.004

135. Scurich N, Dror IE. Continued confusion about inconclusives and error rates: Reply to Weller and Morris. Forensic Sci Int Synerg 2020;2:703–4. https://doi.org/10.1016/j.fsisyn.2020.10.005 PubMed

136. Biedermann A, Kotsoglou KN. Forensic science and the principle of excluded middle: "Inconclusive" decisions and the structure of error rate studies. Forensic Sci Int Synerg 2021;3:100147. https://doi.org/10.1016/j.fsisyn.2021.100147

RTI International is an independent, nonprofit research institute dedicated to improving the human condition. We combine scientific rigor and technical expertise in social and laboratory sciences, engineering, and international development to deliver solutions to the critical needs of clients worldwide.

www.rti.org/rtipress RTI Press publication CP-0015-2204