NIST Interagency Report
NIST IR 8515

Forensic Science Environmental Scan 2023

NIST Forensic Science Program

Henry Swofford, Ph.D.
Special Programs Office
Laboratory Programs
National Institute of Standards and Technology

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8515

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

NIST Interagency Report
NIST IR 8515

Forensic Science Environmental Scan 2023

NIST Forensic Science Program

Henry Swofford, Ph.D.
Special Programs Office
Laboratory Programs
National Institute of Standards and Technology

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8515

March 2024

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

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

NIST IR 8515
March 2024

Disclaimers

Certain equipment, instruments, software, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement of any product or service by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

The opinions, recommendations, findings, and conclusions in this publication are those of the author and do not necessarily reflect the views or policies of NIST or the United States Government.

Publication History
Approved by the NIST Editorial Review Board on 2024-03-01

How to Cite this NIST Technical Series Publication
Swofford, HJ (2024) Forensic Science Environmental Scan 2023. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency Report (IR) NIST IR 8515. https://doi.org/10.6028/NIST.IR.8515

NIST Author ORCID iDs
Henry Swofford: 0009-0003-3114-5958

Abstract

This environmental scan provides a broad view of the forensic science environment (as of 2023) and characterizes high-level historical and current issues and trends that might impact near and long-term decisions by forensic science research and standards programs, including the NIST Forensic Science Program. A historical perspective is provided followed by a more detailed summary of various issues relating to the state of forensic science across five different landscapes: governance, economic, societal, scientific and technological, and legal and regulatory. Each landscape includes a discussion of several issues that are potentially relevant to the administration of forensic science research and standards programs in the United States. The relevance of each issue, and manner in which it might impact those programs, are discussed under the heading "Considerations for Forensic Science Research and Standards Programs" within each subsection and outline potential approaches and priorities that those programs might consider as part of their longer-term strategy moving forward.

Keywords

Forensic Science; Environmental Scan; Strategic Priorities.

Table of Contents

Preface

NIST's mission is to promote innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life. Fundamental to economic security and quality of life is public safety and the effectiveness of the criminal justice system, measured not only in terms of prevalence (e.g., crime detection, deterrence, prevention), but also in terms of the public's perception and trust that the system is fair and just [1, 2]. Supporting accurate, reliable, and trustworthy forensic science is therefore critical to NIST's overall mission; this support is accomplished through the NIST Forensic Science Program [3]. As the NIST Forensic Science Program considers its long-term vision and strategic priorities, it must consider its surrounding environment so that it can be organizationally and programmatically responsive to evolving community and societal needs.

Environmental scans provide a structured framework and systematic means of characterizing relevant issues and trends. This document was developed in 2023 to inform strategic planning efforts by the NIST Forensic Science Program. This document is not meant to be comprehensive or overly detailed to any particular issue or trend, or specific to any single forensic science discipline, method, or practice.

In developing this scan, a variety of published and publicly accessible sources were used. Special effort was made to preserve the original language of the source material and include in-text citations where relevant and practical rather than provide paraphrased statements. All sources are referenced throughout the document. Additional information and considerations relevant to the discussion but peripheral to the core issues are included in the footnotes.

This scan characterizes the forensic science environment across five different but interrelated landscapes: governance, economic, societal, scientific and technological, and legal and regulatory. Each landscape overview is intended to stand alone and includes a discussion of several issues that are potentially relevant to the administration of forensic science research and standards programs in the United States. Some issues and trends, however, are relevant to multiple landscapes and therefore might appear to be redundant. Although some redundancies exist throughout this document, the frame of reference differs as a function of the landscape. The relevance of each issue, and manner in which it might impact forensic science research and standards programs, are discussed under the heading "Considerations for Forensic Science Research and Standards Programs" within each subsection and outline potential approaches and priorities that those programs might consider as part of their longer-term strategy moving forward. Section 7 provides a consolidated list of those considerations across all landscapes. Finally, this scan was developed through the lens of the NIST Forensic Science Program; however, it is not intended to be specific to any particular forensic science research or standards program.

Acknowledgements

This environmental scan was made possible through the contributions and support of several individuals. Special recognition is given to John Butler, Vincent J. Desiderio, John Paul Jones, II, Sandra Koch, Robert Ramotowski, Erica Romsos, and Katherine Sharpless for their detailed reviews, thoughtful comments, and improvements to the quality of this document and to S. Shyam-Sunder for his guidance in formulating it.

1. Introduction

Throughout much of the 20th century, forensic science became a cornerstone of criminal investigations and prosecutions—often viewed as the "silent witness" [4]. "[E]evidence from a variety of forensic sciences was routinely admitted in state and federal courts with very little scrutiny" and "generally without limitation" [5 p1]. Soon after the turn of the 21st century, increased attention was given to forensic science disciplines by media and television, bringing forensic science into popular culture [6]. This led to a phenomenon dubbed the "CSI Effect," describing increased demands for forensic evidence in criminal trials and exaggerated expectations of its capabilities and accuracy, often portrayed with "unfailing certainty" and "no mistakes" being made [6 pp48-49]. Although forensic scientists and litigators often sought to temper the public's expectations by contrasting the fictionalized portrayal of forensic science with challenges faced in the real-world (e.g., the swiftness in which cases are solved, unfettered access to financial resources, and use of sophisticated technology), issues concerning the validity, reliability, and overall scientific rigor of forensic science practices remained unquestioned [5, 7, 8].

While various forensic science disciplines were believed to have sufficient foundations based on their decades of use and "assurances from the forensic science community that the techniques were accurate, effective, and broadly accepted as valid" [5 p1], throughout the 1990s and early 2000s, the federal government and scientific community focused heavily on furthering the development and maturation of DNA identification methods [5, 6, 9, 10]. In the early 2000s, federal funding for advancing DNA identification methods exceeded $200 million per year [11]. While that level of support was important for strengthening the DNA discipline, the forensic science community, acting through the Consortium of Forensic Science Organizations (CFSO),[1] took steps to draw Congressional attention to resource needs for traditional forensic science disciplines other than DNA [12].

In 2005, Congress responded by authorizing the National Academy of Sciences (NAS), through the Science, State, Justice, Commerce, and Related Agencies Appropriations Act of 2006, to assemble the National Research Council (NRC) Committee on Identifying the Needs of the Forensic Science Community to look into these issues [6]. As part of this review, the committee performed an evaluation of the underlying methods and practices of several forensic science disciplines [6].

In February 2009, the NRC released their report entitled *Strengthening Forensic Science in the United States: A Path Forward* [6]. Although the NRC recognized the value forensic science provides to our criminal justice system, they also noted "[t]he law's greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is *science* in any given forensic science discipline" and underscored the importance of ensuring forensic science practices were based on scientifically valid methods

[1] The CFSO represents seven professional organizations and over 21,000 forensic science practitioners throughout the United States [268].

and testimony [6 p87]. What the committee found was in contrast to what the public previously believed, stating that "substantive information and testimony based on faulty forensic science analyses may have contributed to wrongful convictions of innocent people," and "[t]his fact has demonstrated the potential danger of giving undue weight to evidence and testimony derived from imperfect testing and analysis" [6 p4].

The NRC report raised several issues regarding the health of forensic science, noting "[l]ittle rigorous systematic research has been done to validate the basic premises and techniques in a number of forensic science disciplines" [6 p189]. Ultimately, the NRC report highlighted a "fragmented" system with "disparities" across the community [6]. Challenges included "lacking … resources (money, staff, training, equipment)," "inconsistent practices," "absence of … effective oversight," and a "dearth" of well-established scientific foundations underpinning many forensic science methods [6]. In light of this, the NRC made several recommendations to strengthen the foundations of forensic science through a stronger emphasis in basic and applied research and promoting greater standardization and consistency in how forensic science is practiced across the nation [6].

In 2015, six years after the release of the NRC report, the President's Council of Advisors on Science and Technology (PCAST) was charged to "consider whether there are additional steps that could usefully be taken on the scientific side to strengthen the forensic-science disciplines and ensure the validity of forensic evidence used in the Nation's legal system" [13 p1]. In September 2016, the PCAST published their report entitled *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* [13]. In that report, the PCAST concluded that "there are two important gaps: (1) the need for clarity about the scientific standards for the validity and reliability of forensic methods and (2) the need to evaluate specific forensic methods to determine whether they have been scientifically established to be valid and reliable" [13 p1]. The PCAST also made several recommendations, many of which were directed specifically at NIST, "as a science agency which has no stake in the outcome," due to its expertise in science and metrology and its independence and neutrality with respect to the criminal justice system [13 p128].

The PCAST report took a more assertive stance on issues concerning the scientific validity and reliability of results from many forensic science disciplines that, for decades, had been admitted in many courts, and how issues concerning scientific validity ought to correspond to legal admissibility [13].[2] The PCAST report, which questioned the validity of certain forensic science practices, stimulated several responses shortly after its publication from professional organizations representing various communities within the criminal justice

[2] "*Foundational validity* for a forensic-science method requires that it be shown, based on empirical studies, to be *repeatable, reproducible, and accurate,* at levels that have been measured and are appropriate to the intended application. Foundational validity, then, means that a method can, *in principle*, be reliable. It is the *scientific* concept we mean to correspond to the *legal* requirement, in Rule 702(c), of 'reliable principles and methods.' … *Validity as applied* means that the method has been reliably applied *in practice*. It is the *scientific* concept we mean to correspond to the *legal* requirement, in Rule 702(d), that an expert "has reliably applied the principles and methods to the facts of the case." [13 pp4-5, 40-43].

system (e.g., [14–19]).[3] Responses from defense attorneys were welcoming of the report—for example, the National Association of Criminal Defense Lawyers (NACDL) stated the report "offers further evidence of the pervasive use of flawed analysis erroneously presented as grounded in science" [14], and the Innocence Project claimed the report "provided a blueprint for fixing one of the most critical problems plaguing the criminal justice system" [15]. Responses from prosecutors, on the other hand, largely objected to the findings and recommendations in the report—for example, the National District Attorneys Association (NDAA) stated "the NDAA takes issue with, and has substantial concern about, the logic of the [PCAST] report and the manner in which it portrays several forensic disciplines," citing "the pervasive bias and lack of independence apparent throughout the report" [16], and the U.S. Department of Justice (DOJ) issued a statement arguing the claims made by the PCAST were "erroneous" [20].

The NRC and PCAST reports were largely viewed as critical of the state of forensic science, and they have drawn attention from policy makers, researchers, and others within the criminal justice system, stimulating widespread conversations on issues concerning forensic science practices and charting a path forward for forensic science research and science-based criminal justice reforms [21]. While the principal recommendation of the NRC report, to create the National Institute of Forensic Science (NIFS) as a new independent federal agency having governance and oversight of forensic science in the United States [6], has yet to materialize, NIST has been directed by Congress and succeeding Administrations to play a role in shaping the future of forensic science with a focus on strengthening the scientific basis of standards and practices that is informed by rigorous research. Notable efforts at NIST include:

- Partnership with DOJ to establish and co-chair the National Commission on Forensic Science (NCFS), a federal advisory committee to the United States Attorney General on forensic science policies and practices [22];[4]
- Establishment and administration of the Organization of Scientific Area Committees for Forensic Science (OSAC), an organization of 800-plus volunteer members and affiliates across 22 discipline-specific forensic science committees responsible for facilitating the development of forensic science standards to promote consistency and standardization of forensic science practices throughout the United States [23];
- Expansion of its forensic science research program, providing foundational and applied research, development, testing, and evaluation support to the forensic science community across a wide range of topics and disciplines including: biometrics, digital evidence, drugs and toxicology, firearms and toolmarks, evidential statistics, forensic genetics, quality assurance, and trace evidence [3];
- Establishment and administration of the Center for Statistics and Applications in Forensic Evidence (CSAFE), a consortium of academic institutions responsible for developing a stronger statistical and scientific foundation for the interpretation of forensic evidence and elevating the literacy of forensic science practitioners, legal

[3] The White House Office of Science and Technology Policy (OSTP) listed additional responses to the PCAST report by different organizations within the forensic science community [109].
[4] The NCFS charter ended in 2017 and was not renewed by the DOJ.

professionals, and others within the criminal justice system related to probability and statistics [24]; and
- Creation of a foundation studies program, providing technical merit evaluations of the empirical evidence that underpins the scientific reliability of forensic science methods and practices [25].

NIST has invested heavily in strengthening forensic science over the last decade. Despite considerable progress, however, much remains ahead. As NIST assesses the progress and looks to consider its long-term vision and strategic priorities for strengthening the nation's use of forensic science, it must do so in the context of community needs and priorities, its unique measurement science and standards mission, as well as an assessment of the current environment, described in this document across several interrelated landscapes (governance, economic, societal, scientific and technological, and legal and regulatory), that characterize the state of forensic science today.

2. Governance Landscape

Issues and trends relating to the administration of forensic science activities and evidence, and how that might impact the forensic science community, are evaluated in the governance landscape. Looking through a governance lens, there are three key issues that are relevant and potentially impactful to forensic science research and standards programs: (i) administrative controls, (ii) jurisdictional authorities, and (iii) judicial dispositions. All of these issues have been prevalent for decades but did not receive widespread attention until they were highlighted by the 2009 NRC report [6]. Despite being raised nearly 15 years ago, they continue to define the governance landscape of forensic science in the United States today.

2.1. Administrative Controls

In the United States, most forensic science activities are carried out under the administrative controls of state and local law enforcement entities and, in some instances, prosecutorial offices [6 pp183-185]. The 2009 NRC report brought attention to this issue, suggesting that such administrative controls can create "[c]ultural pressures caused by the different missions of scientific laboratories vis-à-vis law enforcement agencies" [6 p185], leading to competitions in budgetary priorities and resource allocations and introducing subtle biases impacting how the evidence might be evaluated, creating a risk that forensic practitioners might "sacrifice appropriate methodology for the sake of expediency" [6 p24]. Over a decade later, calls for criminal justice reform echoed the concerns first raised by the NRC and argued that such administrative controls mean forensic science budgets, priorities, and policies and procedures are ultimately under the control of individuals without appropriate qualifications to operate forensic science laboratories [26]. For example, one commentator claimed:

> *Although seemingly mundane, lab policies define requirements for job qualifications, evidence collection, analytical procedures and equipment. They also determine how results are interpreted, reported and presented in court. While forensic scientists conduct the analysis, the lab administration ultimately controls the scope and influence that their work has on the criminal justice system. The issue is not that police administrators are manipulating the system to their benefit but rather that they are making unqualified decisions.* [26]

Whether these subtle factors are material or not is a separate issue; nevertheless, the mere *perception* of bias could affect the public's trust in the reliability and neutrality of forensic science evidence [27]. Nearly 15 years after the NRC first raised the issue, the American Society of Crime Laboratory Directors (ASCLD) recognized this perception and "acknowledge[d] unconscious biases exists" and that laboratories are working to "enhance public trust and confidence in both the parent agency and the forensic laboratory" [27].

2.1.1. Considerations for Forensic Science Research and Standards Programs

The existence of a public perception of bias is relevant context that should be considered by forensic science research and standards programs when carrying out their missions. It is helpful for forensic science research and standards programs to be detached from litigation to maintain their reputation of impartiality and neutrality; however, that alone might not be enough to promote trust in the validity and reliability of forensic science. This leads to two major considerations for forensic science research and standards programs (Box 2.1.1).

Box 2.1.1. Forensic science research and standards programs should:

(a) Ensure transparency and balance across diverse perspectives when administering programmatic activities so that the outputs align to the needs of all members of the forensic science community and are not unduly influenced (real or perceived) by any individual groups (researchers, practitioners, investigators, prosecutors, defense attorneys, and judges).

(b) Take a leading role and affirmative stance on matters relating to good scientific principles and practices—even in the face of adversarial viewpoints.

2.2. Jurisdictional Authorities

In 2009, the NRC noted "[m]ost forensic science methods, programs, and evidence are within the regulatory province of state and local law enforcement entities or are covered by statutes and rules governing state judicial proceedings" and are therefore handled by state and local jurisdictions [6 p13]. This has two significant consequences: (i) forensic science activities are often anchored by state and local jurisdictional boundaries—in both how they are resourced as well as the constituencies they serve—and (ii) there is no centralized authority governing forensic science activities or universal policies requiring, for example, standardization, conformity assessment, quality management, accreditation, and certification across the nation [6].[5] Instead, forensic science falls under the jurisdictional authority of *hundreds to thousands* of different entities [6].[6] This has led to "great disparities among existing forensic science operations in federal, state, and local law enforcement jurisdictions and agencies … [which] is true with respect to funding, access to analytical instrumentation, the availability of skilled and well-trained personnel, certification, accreditation, and oversight" [6 pp5-6].

As noted by the NRC, the majority of forensic science entities are "sorely lacking in the resources (money, staff, training, and equipment) necessary to promote and maintain strong forensic science laboratory systems … [and a]s a result, the depth, reliability, and overall quality of substantive information arising from the forensic examination of evidence available to the legal system vary substantially across the country" [6 p6]. Consequently, this also means that people might experience differences in the quality and effectiveness of the criminal justice system simply because of the jurisdiction in which they live, which has led some to express concern about the extent to which they might trust forensic science results (e.g., [28 p23]).

2.2.1. Considerations for Forensic Science Research and Standards Programs

Although issues concerning jurisdictional authority of forensic science practices are outside the scope of forensic science research and standards programs, the differences that are created as a result are relevant to those programs and lead to two major considerations (Box 2.2.1).

[5] Although lacking centralization at the national level, progress toward this objective have been made at the state level (e.g., Texas Forensic Science Commission [269] and New York State Commission on Forensic Science [270], among others [264, 271]). A more detailed discussion of forensic science boards and commissions is provided in the legal and regulatory landscape.

[6] According to the Bureau of Justice Statistics (BJS), there are approximately 400 different publicly funded crime laboratories and 18,000 different law enforcement entities operating in the United States [37, 38, 272].

Box 2.2.1. Forensic science research and standards programs should:

(a) Ensure program outputs are both applicable and accessible to all forensic service providers despite differences in resources and account for those resource limitations when carrying out their missions.

(b) Provide coordinated mechanisms to promote sound scientific practices that define how forensic science activities should be practiced by forensic service providers throughout the nation and consider ways to help enable and promote mechanisms to ensure forensic service providers conform to those practices.

It is important for forensic science research and standards programs to consider the resource constraints and differences that exist throughout the forensic science community and administer programmatic activities with a practical perspective. No matter how resource-strapped forensic service providers might be, they should have the opportunities to leverage the outputs from forensic science research and standards programs to strengthen their practices. This will require forensic science research and standards programs to consider strategies to account for those resource limitations when carrying out their missions so that the outcomes of the programs can have meaningful impact for all forensic service providers across the nation.

Although forensic science research and standards programs do not have regulatory mandates or authorities, they can serve as a credible source of information on standards and best practices, conformity assessment schemes, quality management protocols, and accreditation and certification systems for jurisdictional authorities to adopt and use.[7] While promoting good science-based standards and practices is critical, that alone is not enough to engender trust in the validity and reliability of forensic science evidence relied upon by the criminal justice system. To be most effective, there must be a way to ensure forensic service providers conform to those practices.[8]

---

[7] The development and implementation of standards, guidelines, and recommendations relating to how forensic science practices ought to be carried out is important to facilitate standardization. However, standards, guidelines, and recommendations can be developed and promoted by various entities. The mere existence of standards, guidelines, and recommendations alone is not enough to achieve meaningful standardization across the forensic science community. To that end, there must be a recognized and legitimate source for promoting *which* standards, guidelines, and recommendations achieve the desired level of quality to ensure good scientific practices.

[8] Conformity assessment is traditionally carried out by accrediting bodies independent of the forensic service provider. Non-government accrediting bodies provide fee-based conformity assessment services to forensic service providers that voluntarily choose to use those services to demonstrate conformance to national or international standards, such as ISO/IEC 17025:2017 and ISO/IEC 17020:2012; however, those assessments do not necessarily require conformance to discipline-specific forensic science standards such as those recognized by the OSAC Registry [169]. Instead, forensic service providers provide self-declarations of conformance to discipline-specific standards and guidelines [273].

2.3. Judicial Dispositions

In the United States, judicial dispositions relating to forensic science evidence have been a point of issue for decades [6, 29]. While evidentiary rules and admissibility criteria differ across jurisdictions, many apply either the *Frye* [30] or *Daubert* [31] standard [6, 29].[9] Although *Daubert* places greater emphasis on the role of the judiciary to serve as the gatekeeper for ensuring reliable scientific evidence, the NRC claimed that "courts have been utterly ineffective" and "ill-equipped" to accomplish this [6 p53]. The NRC continues, noting "[j]udicial dispositions of *Daubert*-type questions in criminal cases have been criticized by some lawyers and scholars who thought that the Supreme Court's decision would have been applied more rigorously … [and i]f one focuses solely on reported federal appellate decisions, the picture is not appealing to those who have preferred a more rigorous application of *Daubert*" [6 pp95-96]. The NRC continued, "[f]ederal appellate courts have not with any consistency or clarity imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving *Daubert* questions" [6 p96].

In their report, the NRC highlighted a general tendency for courts to more often rule in favor of prosecutors' requests to admit expert testimony despite defense challenges, and differences in judicial dispositions when comparing criminal cases versus civil cases [6]. Specifically, the NRC noted:

> *[T]rial judges exercise great discretion in deciding whether to admit or exclude expert testimony … [and] the vast majority of the reported opinions in criminal cases indicate that trial judges rarely exclude or restrict expert testimony offered by prosecutors; most reported opinions also indicate that appellate courts routinely deny appeals contesting trial court decisions admitting forensic evidence against criminal defendants. … The situation appears to be very different in civil cases. Plaintiffs and defendants, equally, are more likely to have access to expert witnesses in civil cases, while prosecutors usually have an advantage over most defendants in offering expert testimony in criminal cases. And, ironically, the appellate courts appear to be more willing to second-guess trial court judgements on the admissibility of purported scientific evidence in civil cases than in criminal cases.* [6 p11]

More recently, some judges have also reflected on this issue, noting "inconsistent results are not atypical," and that "it accords with a more general perception that courts, both state and federal, remain reluctant to exclude even those kinds of forensic science whose accuracy has been severely questioned over the past few decades" [29 p3]. They continued:

> *Why is this? Perhaps it is because many judges are reluctant to keep from juries evidence that they still "feel" is probative. Or perhaps it is because most judges lack a scientific background … and do not feel comfortable independently assessing the*

[9] Federal courts apply the *Daubert* Standard but state courts differ whether they apply *Frye* or *Daubert*. Under *Frye*, the judge must determine whether the scientific technique upon which the testimony is based is "generally accepted" as reliable in the relevant scientific community [30]. Under *Daubert*, the judge must determine whether the testimony is based on scientifically valid reasoning which can properly be applied to the facts at issue [31].

> *reliability of scientific evidence. … Or perhaps it is because some judges are former prosecutors who regularly introduced such evidence in their earlier careers.* [29 p3]

A similar sentiment was also expressed elsewhere [28]. When evaluating judicial perspectives on various issues relating to forensic science, one judge responded to a question about admissibility of expert testimony, stating:

> *I think that some judges don't like to exclude. They'd rather let it in than exclude and let it go to the jury. If there's an arguable basis for the jury to have accepted something, civil or criminal, then they [tend to] let it go to the jury. And that's a relatively safe place for them to be. If they exclude, they're subject to a reversal for an erroneous exclusion.* [28 p19]

Likewise, when discussing "what types of pressures might judges find themselves under" that could impact judicial dispositions, another judge "point[ed] to political pressures, professional incentives, and biases to their own prosecutorial experiences," [28 p19] stating:

> *I will speculate. I should tell you, though, the statistics are quite striking. Daubert challenges succeed in civil cases frequently. They succeed in criminal cases almost never. And that shows, I think, that there is a double standard operating. So, why is that? One factor is that in most states trial judges are elected, and if they have to face re-election on the basis they are "soft on crime" because by God, I wouldn't even allow fingerprint evidence in, they're in trouble to be re-elected or even to be renominated by the party of their choice. So, election is an element, but I think a more subtle element is going on in most of these cases. The stakes are so much higher and judges, having seen the other evidence in the case, may think "yea, he's probably guilty, but you never know what a jury is going to do. If I keep out this evidence, maybe there won't be a conviction, and I really think it would be unfair to the prosecutor not to at least be able to present this evidence to the jury and they can take it for what it's worth." I think that is a wrongful attitude. I think I'd say a dereliction of duty and really ignores what Daubert is all about or even Frye for that matter. But, I do think that's a common traditional attitude: "I don't want to be responsible for this guy being acquitted, when, what I've heard so far, he's probably guilty." And of course, forensic evidence carries great weight. It has an aura of neutrality that you don't have from testimony of accomplices, for example. So, I think judges are reluctant to keep it out. I'll mention a third factor, which is that most criminal court judges are former prosecutors. Relatively few are former defense lawyers. So, there's also, "oh yeah, of course. I always let this in, I used to do it myself. This is just routine. I recognize this."* [28 p19]

The potential for pressures and other factors, such as judicial selection mechanisms, to impact judicial decisions in criminal cases has also been raised by others in recent years [32, 33]. In the United States, judges are either elected or appointed, depending on the jurisdiction and the type of judicial seat being occupied [32]. Most criminal cases are adjudicated at the state level, for which judicial selection is most often through elections [32].[10] Although in most states judges might not necessarily participate in traditional campaigns, judicial elections have focused on candidates' prior records on issues related to criminal justice, with negative ads citing "soft on crime" and, in some instances, attacking candidates for prior work as a defense attorney providing representation to criminal defendants [32]. Positive ads have cited "tough on crime" and some candidates have taken stances promising not to allow "technicalities to overturn convictions" [32 pp3-6]. Over the years, some have claimed that "the selection of state court judges has become increasingly politicized, polarized, and dominated by special interests—particularly in the 39 states that use elections as part of their system for choosing judges [and t]hese trends put new pressures on state court judges" [33].

Calls for science-based entities to take more active roles in providing independent assessments of forensic science to support judicial decision-making have been made [6, 13, 34, 35]. The PCAST claimed "[t]here is an urgent need for ongoing evaluation of the foundational validity of important methods, to provide guidance to the courts, the DOJ, and the forensic science community" [13 p124]. Likewise, the National Commission on Forensic Science called for an "entity with the capacity to conduct independent scientific evaluations of the technical merit of test methods and practices used in forensic science disciplines" [35] and legal scholars have suggested that the "the justice system may be institutionally incapable of applying *Daubert* in criminal cases because it does not have access to independent scientific expertise on an ongoing basis" [34]. Similar calls have been made by judges themselves. For example, when responding to questions concerning the regulation and oversight of forensic science methods, one judge responded:

> *I think there is a real need for an Institute of Forensic Science staffed by a high-level scientists who could tell us with the neutrality that we deserve, this is good forensic science, this is bad forensic science, this is possible forensic science but it has to be improved and here's how to go about improving it. … I don't think the legal system, ultimately, is well positioned to regulate forensic science. Judges know beans about science. Lawyers know beans about science. The natural thing when you have that kind of problem is to turn it over to the people who do know about science, the scientists.* [28 p14]

2.3.1. Considerations for Forensic Science Research and Standards Programs

Admissibility decisions relating to forensic science evidence are outside the scope of forensic science research and standards programs; however, science-based issues that might affect those decisions, and ensuring the criminal justice system has access to that information, are relevant. This leads to two major considerations for forensic science research and standards programs (Box 2.3.1).

[10] "Nearly all felony convictions—94 percent—occur in state courts, including 99 percent of rape cases and 98 percent of murder cases. The arbiters of these cases, state court judges, are mainly elected. Nationwide, 87 percent of state judges face elections, which occur in 39 states" [32 p1].

Box 2.3.1. Forensic science research and standards programs should:

(a) Take a neutral but affirmative stance on science-based issues affecting the use of forensic science in the criminal justice system, such as standards, conformity assessment, quality management systems, and accreditation.

(b) Ensure more guidance is made available and accessible to members of the forensic science and criminal justice communities on issues relating to scientific validity and reliability and how they can be properly assessed using science-based standards, guidelines, and scientific information concerning forensic science practice.

3. Economic Landscape

Issues and trends relating to the availability and access to resources to support the demands for forensic science services, and how that might impact the forensic science community, are evaluated in the economic landscape. In their 2009 report, the NRC noted the forensic science community "is underresourced in many ways" [6 p6] and "a number of factors have combined in the past few decades to place increasing demands on an already overtaxed, inconsistent, and underresourced forensic science infrastructure … [which] have not only stressed the system's capacity, but also have raised serious questions and concerns about the validity and reliability of some forensic methods and techniques and how forensic evidence is reported to juries and courts" [6 p39]. Particularly at the state and local levels, the NRC found that the community is "sorely lacking in the resources (money, staff, training, and equipment) necessary to promote and maintain strong forensic science laboratory systems" [6 p77]. Expounding on this, the NRC clarified:

> *By using the term "underresourced," the committee means to imply all of its dimensions. Existing data suggests that forensic laboratories are underresourced and understaffed, which contributes to a backlog in cases and likely makes it difficult for laboratories to do as much as they could to inform investigations, provide strong evidence for prosecutions, and avoid errors that could lead to imperfect justice. But underresourced also means that the tools of forensic science are not as strong as they could be. … This underresourcing limits the ability of the many hard-working and conscientious people in the forensic science community to do their best work. … [Compounding these issues, f]orensic science research is not well supported, and … [r]elative to other areas of science, the forensic science disciplines have extremely limited opportunities for research funding.* [6 pp77-78]

Looking through an economic lens, there are two key issues that are relevant and potentially impactful to forensic science research and standards programs: (i) operating budgets and (ii) research and development.[11] Both issues can affect the growth and sustainability of a high-quality forensic science enterprise, and their impacts have been underscored by the NRC report in 2009 [6]. While improvements have been made in the years following the NRC report, these issues continue to define the economic landscape of forensic science in the United States today.

3.1. Operating Budgets

In 2009, the Bureau of Justice Statistics (BJS) estimated the combined annual operating budgets for the nation's 400+ publicly funded crime labs[12] totaled approximately $1.6 billion [36]. This increased to approximately $1.7 billion by 2014 [37] and approximately $2 billion by 2020 [38]. This represented an increase in operating budgets of approximately 6% between 2009 and 2014 and approximately 17% between 2014 and 2020 [36–38]. According to the Bureau of Labor Statistics (BLS), consumer inflation over that same period (January 2009 – December 2020) was estimated to be approximately 23% [39], suggesting that the increases in operating budgets have generally kept pace with consumer inflation but the budgetary challenges noted by the NRC in 2009 [6] have persisted. Further illustrating these challenges, the BJS data reflect the total number of requests received *exceeded* the number of requests completed for yearend 2009, 2014, and 2020 overall and for the following disciplines: controlled substances, DNA databasing, forensic biology, firearms/toolmarks, and trace evidence analyses [38]. In yearend 2020, digital evidence[13] and latent print analyses also experienced a deficit in the number of requests completed versus the number received [38]. Additionally, between 2014 and 2020, the number of backlogged requests[14] increased by approximately 25% (from 570,100 to 710,900) overall,[15] impacted by increases for firearms/toolmarks (up 97%: from 51,100 to 101,000), DNA databasing (up 87%: from 64,8000 to 121,000), controlled substances (up 22%: from 213,700 to 260,600), forensic biology (up 17%: from 107,800 to 126,100), and toxicology analysis (up 16%: from 40,000 to 46,400) [38].[16] These data reflect a forensic science infrastructure that continues to face challenges keeping up with throughput demands, causing many labs (38% in 2014 [37 p4] and 47% in 2020 [38 p7]) to

[11] Two other issues that are worth noting but not discussed in this landscape assessment since they are not necessarily directly tied to accessibility of resources to support the day-to-day demands for forensic services include (1) the costs associated with litigation and damages caused by issues in the examination and interpretation of forensic evidence that could contribute to wrongful convictions (these errors often cost taxpayers millions of dollars in litigation and damages [43]) and (2) the niche commercial market for many forensic science disciplines that limit broad-scale commercial innovation [52].

[12] The Bureau of Justice Statistics (BJS) defines a "crime lab" as "a scientific laboratory (with at least one full-time natural scientist) that examines physical evidence in criminal matters and provides reports and opinion testimony with respect to such physical evidence in courts of law" [274]. This definition does not include law enforcement agencies that employ personnel that perform forensic services outside of a formal laboratory setting, such as identification units.

[13] Although digital evidence services are considerably lower than other "traditional" requests, a 2022 landscape assessment noted that most digital evidence services are conducted outside of traditional publicly funded crime laboratory settings and estimate that there are approximately 11,000 digital evidence service providers in the United States, many of which are not represented by the BJS data [131 p23].

[14] The BJS considered a case backlogged if "it had not been completed and reported to the submitting agency within 30 days of submission" [38].

[15] The BJS notes, however, that some crime labs responded to the COVID-19 pandemic by suspending operations during 2020, which partly accounted for the increase in backlogged requests from yearend 2014 to yearend 2020 [38].

[16] These increases in backlogged requests between yearend 2014 and 2020 are in addition to the increases experienced between yearend 2009 and 2014 for which firearms/toolmarks increased by 6% (from 48,300 to 51,100), controlled substances increased by 53% (from 139,200 to 231,700), forensic biology increased by 4% (from 103,500 to 107,800), and toxicology analysis increased by 45% (from 27,600 to 40,000) [38].

outsource requests for services and to seek funding from other sources, such as grants[17] and fee-for-service business models [38].[18] Despite some labs turning to fee-for-service models, some within the criminal justice community do not consider them desirable [40].[19]

In 2009, the NRC not only noted resource limitations affecting the forensic science community, but also the disparities in resources between different jurisdictions, recognizing "federal programs are often much better programmed and staffed" [6 pp5-6]. This observation is also reflected in the BJS data [38]. In 2020, the overall average budget per request was $620 [38]. "State labs had the lowest budget per request at about $550, while federal labs had the highest at about $900 per request" [38 p12]. County labs and municipal labs were at about $670 and $730 per request, respectively [38 p12]. These data suggest that most of the strain on the forensic science infrastructure is borne by state and local jurisdictions when looking at the system as a whole [38].

Other sources have revealed similar findings. According to Project FORESIGHT,[20] between 2013 and early 2020 (prior to the COVID-19 pandemic), trends were noted relating to average annual growth in costs, turnaround times, and backlogs [41]. On average, across all participating laboratories and nine disciplines (blood alcohol [BAC], crime scene investigation, digital evidence, DNA casework, fingerprints, fire analysis, toxicology [antemortem, excluding BAC], toxicology [postmortem, excluding BAC], and trace evidence), Project FORESIGHT data indicated a 10% average annual growth in costs, a 15% average annual growth in turnaround times, and a 152% average annual growth in backlogs [41 pp55-58].[21] Project FORESIGHT analysts noted that "[p]art of the additional strain on resources could be attributed to the attention placed on unsubmitted sexual assault kits (SAKs) and the drive to test the 200,000 to 400,000 outstanding SAKs that had yet to be submitted for laboratory analysis … [as well as] the growing opioid crisis" [41 p55].

In 2019, in response to a mandate of the Justice for All Reauthorization Act of 2016, NIJ published a needs assessment of forensic laboratories, which included an examination of personnel, workload, backlog, and equipment needs for both public crime laboratories and medical examiner and coroners' offices [42]. While the BJS and Project FORESIGHT data provide macroeconomic quantitative trends, the NIJ report provides a qualitative summary of key laboratory needs and challenges [42]. In their report, NIJ noted the need for:

[17] In 2020, approximately 72% of crime labs in the United States received funding from federal grants (96% of state labs, 75% of county labs, and 64% of municipal labs) and approximately 44% of crime labs received funding from state or local grants [38 p13].
[18] "In 2014, approximately 4 in 10 crime labs charged the submitting agencies for fees for completing the forensic services" [37 p5]. In 2020, "[a]bout 34% received fees from services performed" [38 p13].
[19] E.g., In 2008, Idaho State Police surveyed members of the criminal justice community on funding alternatives for forensic services. In this survey, they found high disapproval rates for charging local agencies for forensic services (68% of respondents disapprove of charging for a portion of the total cost and 95% of respondents disapprove of charging for the entire cost). Instead, 95% of respondents agree that additional funding should be appropriated to underwrite the costs of additional equipment and personnel for forensics [40].
[20] Project FORESIGHT is a tool that provides laboratory managers with actionable insights into the operational performance of their laboratory. Participation is voluntary and the participating laboratories represent local, regional, state, and national agencies. Faculty from the West Virginia University John Chambers College of Business and Economics analyze the submitted data to identify trends across laboratories and performance of individual laboratories [275].
[21] Crime scene investigation represented a significant increase in average annual growth in backlog (628.5%). By removing crime scene investigation from the disciplines, the average annual growth values for the remaining eight disciplines are adjusted to 9% increase in costs, 15% in turnaround times, and 92% increase in backlogs.

- “[s]ufficient and consistent funding and strategic planning to process increasing amounts of forensic evidence and to address fluctuations in evidence submissions driven by the supply of and demand for forensic services” [42 p3], citing “the estimated funding gap for the 2017 level of forensic casework was $640 million” [42 pp3, 39, 178-179], the “average area of investigation has seen a 60% increase in turnaround time” [42 p179], and “the average backlog across all areas of forensic science has grown nearly 250%” [42 p179];
- “[s]ufficient and consistent funding for forensic practitioner training” [42 p3], citing the cost to train a new analyst for one year to be “in excess of $100,000” [42 p27] and “[f]unding for training is typically only 0.5% of forensic laboratories’ operating budgets” [42 p5]; and
- “[c]ontinued efforts to strengthen quality assurance measures, limit preventable nonconformities, and maintain a healthy workforce in the forensic sciences” [42 p4], citing the need for “[a]dditional support and resources for agencies seeking accreditation” [42 p65].

The findings from the 2019 NIJ needs assessment reinforce the observations from macroeconomic trends and indicate a continued need for access to resources to support basic operations in forensic science laboratories [42]. Resources for capital expenditures, such as updating equipment and instrumentation, validating and implementing new analytical methods (including resources to support personnel training, competency and proficiency testing relating to those methods), and improving aging laboratory facilities and infrastructure are often delayed in favor of personnel and consumables to support immediate throughput needs [42 p59]. The potential consequences of this are missed opportunities for greater throughput and lower costs in the long term and limited abilities to strengthen the scientific rigor and capacity of many forensic science methods [43]. These findings from NIJ were underscored in recent testimony by a past president of the ASCLD and chair of the CFSO, to the Presidential Commission on Law Enforcement and the Administration of Justice on April 14, 2020 [44, 45].

In the years since the 2019 NIJ needs assessment report was published, the world experienced the COVID-19 pandemic. Although the impact of the pandemic on the forensic science community has not been fully characterized, it required forensic service providers to re-think how they approach their work and maintain personnel safety (e.g., [46]). Requirements for and shortages of personal protective equipment (PPE), personnel shortages due to possible or confirmed COVID-19 exposures, increases in costs for basic supplies and PPE, and spikes in crimes and evidence submissions created challenging circumstances for many labs [46–48]. Additionally, consumer inflation increased by 19% between January 2020 and November 2023 [39]. Although resources were already limited,[22] the issues were further exacerbated by warnings for future outlooks [49]. For example, without clear indication of additional resources to support the existing operational infrastructure, the BLS projected the need for forensic science practitioners to grow by an additional 11% between 2021 and 2031, which is “much

[22] E.g., in 2017, NIJ estimated the funding gap to be approximately $640 million [42 pp3, 39, 178-179].

faster than the average for all occupations" (projected at 5% growth) [49]. These projections are due in large part "to process [state and local governments'] high caseloads … [and keep pace with] scientific and technological advances [that] are expected to increase the availability, reliability, and usefulness of objective forensic information used as evidence in trials" [49].

Not only do these resource strains impact current capabilities, but they also hinder laboratories from improving and implementing new methods or technologies [28, 42]. For example, when asked to describe the greatest challenge facing the forensic science community on the use of emerging technologies, such as the use of computational methods (algorithms), multiple laboratory directors answered "resources" [28]. Expounding on this, one laboratory director stated:

> *Resources [are the greatest challenge]. To stay on top of how quick things are developing, it's taking more and more resources. We all have backlogs and we're focusing on those. To take people off of [casework] to train them, then get these new things up to speed and implement them and then change people's minds [takes resources]. ... How can we do a job in a technological field without the resources to bring in these new things? Not only are algorithms coming, they're already here. It's allowing us to do a better and better job. But it takes resources to do that.* [28 pSA-VI 21]

Another laboratory director elaborated:

> *Resources [are the greatest challenge]. Because software itself is expensive, even more so though, is the training and implementation arc of getting people to accept it and understand it, to be able to use it and use it correctly. That's an expensive effort. And let's face it, labs are underwater already. ... Trying to get a group that is underwater, desperately overwhelmed, that can't catch their breath between [cases], to have enough bandwidth to even be able to accept a new tool and not see it as just, "oh my God, you have one more thing." That's going to take time. And, even we don't have bandwidth in there [despite being a relatively well-funded laboratory compared to others]. ... That's what's going to face all of these algorithms. ... It's not that people don't see the advantage of them or see the potential benefit, but how do we get from here to there when everybody is madly trying to decide which horrible, awful crime they're going to put first and which horrible, awful crime goes second. So that's what's under that trivial answer of resources. Then, you also think of all the rest of the infrastructure that goes with being able to effectively use these algorithms — the compute, the storage, the data management — where do we put all of these results? How do we store all of these results? How do we maintain that output, which has probably got some proprietary aspect of the outputs in such a fashion that 20 years from now I can still access those results and be able to explain it? Again, it comes back to a resource issue of all of the infrastructure that goes around the use of that algorithm.* [28 pSA-VI 21]

3.1.1. Considerations for Forensic Science Research and Standards Programs

The forensic science community is facing economic challenges. The increasing demands, growing backlogs, and rising costs all point to a forensic science community with limitations on both operational resources and their ability to keep pace with day-to-day needs. This can impact the extent to which the forensic science community can participate in research and standards development activities, or implement the outcomes produced by research and standards programs. This leads to two major considerations for forensic science research and standards programs (Box 3.1.1).

Box 3.1.1. Forensic science research and standards programs should:

(a) Prioritize efforts that are most relevant, practical, and impactful to addressing the day-to-day challenges faced by the broader forensic science community.

(b) Strengthen partnerships and collaborations among government, academia, professional organizations, and private industry to support forensic service providers in efforts to translate and implement new or improved analytical methods and technologies into practice.

It is important for forensic science research and standards programs to prioritize efforts that align to the most common challenges faced by most forensic service providers. This will maximize the impact of those efforts. Identifying those priorities will require the programs to gain a comprehensive understanding of the wide range of workflows, analytical methods, and operational constraints that characterize current practices, and tailor their programmatic activities toward addressing those issues through supporting research and standards development and implementation. Executing against those priorities will also require the programs to be agile in how they distribute and invest their resources, balancing between flexible shorter-term commitments to address immediate needs and stable longer-term investments to address broad-based developmental challenges.

Partnerships provide a foundation for meaningful collaborations and enable sharing of resources, information, training materials, protocols, equipment, and data. Collaborative approaches for instrument and equipment validations have been called for and proposed as potential strategies to help lower barriers to the implementation of standards, methods, and technologies and accelerate community-wide adoption [42, 45, 50–52].

3.2. Research and Development

In 2009, the NRC concluded:

Forensic science research is not well supported, and there is no unified strategy for developing a forensic science research plan across federal agencies. Relative to other areas of science, the forensic disciplines have extremely limited opportunities for research funding. [6 p78]

The NRC expounded on this, noting that "[n]early all forensic science research funds are channeled through DOJ … [and] NIJ provides the bulk of funds for research" [6 p71].[23] Over the years, however, NIJ has faced decreasing budgets to support research and development in forensic science [53]. The NRC highlighted that "[t]otal expenditures for forensic research were $78 million in FY 2002, but they decreased to $33 million by FY 2009" [6 p71]. Although the President announced a five-year, $1 billion initiative to improve the use of DNA in the criminal justice system in 2003, between 2003 and 2008, the NRC noted that much of the money appropriated for this purpose was for backlog reduction and only $26 million was made available in grants for new research on forensic tools and techniques [6 p71].

In the years following the NRC report, between FY 2011 and FY 2020, NIJ awarded an average of $20.8 million in competitive grants for "research and development in forensic science for criminal justice purposes," with the capacity to fund only $16.9 million in FY 2020 [53 p5]. In FY 2021 and FY 2022, these values were further reduced, leaving NIJ with the funding capacity to award only $14.5 million [54] and $11.7 million [55], respectively. In their 2021 impact report, NIJ notes that "[d]ollar amounts fluctuate annually because there is no dedicated funding source for the Research and Development in Forensic Science for Criminal Justice Purposes program" [53 p5]. Instead, "NIJ must draw from its base funding or transfer funding from Bureau of Justice Assistance forensic science laboratory capacity programs" [53 p5]. Between FY 2011 and FY 2022, the most NIJ was able to fund in a single year occurred in FY 2015 with the capacity to fund $29.7 million [53 p5]; however, in that same year another NRC report concluded that NIJ "funding has been inadequate to meet the needs facing the forensic science field" [56 p60].

Although funding for research and development from NIJ have fluctuated over the last decade and funding available for competitive grants have decreased [53], attention was directed toward promoting research and development through public-private partnerships and promoting research and development from other sources. In 2011, NIJ established a cooperative agreement with RTI International, investing approximately $4 million per year[24] to create a Forensic Technology Center of Excellence [57]. In 2013, NSF published a "Dear Colleague" letter encouraging researchers to submit proposals that address fundamental questions to advance knowledge and education in the forensic sciences [58]. In 2014, NSF, in

[23] The NRC noted: "The [Federal Bureau of Investigation (FBI)] Laboratory also receives roughly $33 million per year for its own research"; however, much of that is applied for internal research and development efforts prioritized by its own organizational needs. Some of that research is published in scientific journals for broader applicability across the community [6 p73]. Outside of the federal government, state and local forensic service providers largely lack any dedicated resources for forensic science research. According to the BJS, in 2014, approximately 14% of crime laboratories had resources dedicated for forensic science research; however, the majority of those resources were provided by NIJ [238 p6]. Between 2014 and 2020, there has been little change with approximately 13% of crime laboratories having resources dedicated for forensic science research in 2020 [38 p14].

[24] Between 2011 and 2022, NIJ invested approximately $42 million across three multi-year awards (2011-DN-BX-K564, 2016-MU-BX-K110, 15PNIJ-21-GK-02192-MUMU) [276].

collaboration with NIJ, invited proposals for the creation of new, multi-disciplinary industry/university cooperative research centers for funding [59].[25] Also in 2014, NIST began receiving appropriated funding for its forensic science research and development program and established four research and development focus areas (Forensic Genetics, Ballistics and Associated Toolmarks, Digital and Identification Forensics, and Statistics) [3].[26] In 2015, NIST established a cooperative agreement with Iowa State University, investing approximately $4 million per year[27] to create a Forensic Science Center of Excellence to provide a stronger statistical foundation to forensic evidence evaluation and interpretation [24]. By 2022, NIST had added four additional research and development focus areas (Drugs & Toxins, Trace Evidence, Biometrics, and Quality Assurance) [3].

In 2016, NIJ commissioned the publication of a landscape study of federal investments in forensic science research and development, summarizing general agency interests, potential funding opportunities, and, where possible, published research needs to assist interested researchers [60], which was further updated in 2021 [61]. While the landscape study does not reflect actual funding or amounts available dedicated specifically for forensic science research, it provides a consolidated list of entities that have or could contribute to funding opportunities for forensic science research and development [60, 61]. Despite those opportunities, the report found that NIJ remains as the "primary federal funder of forensic science research across all forensic science disciplines" and most other funding opportunities are "relevant to specific disciplines or agency missions" [61 p1].

Looking across the forensic science research landscape, this leaves three major entities focused on forensic science research and development geared toward addressing the needs of the broader forensic science community: NIJ, NSF, and NIST [13].[28] While the PCAST recognized that "initial steps" have been taken to "help bridge the significant gaps between the forensic practitioner and academic research communities," they noted that "the total level of Federal funding by NIJ, NSF, and NIST to the academic community for fundamental research in forensic science is extremely small" [13 p127]. Ultimately, the PCAST reinforced the findings from the NRC, claiming "[s]ubstantially larger funding will be needed to develop a robust research community and to support the development and evaluation of promising new technologies" and "Federal R&D efforts in forensic science, both intramural and extramural, need to be better coordinated … to ensure that funding is directed to the highest priorities and that work is of high quality" [13 pp127-128].

[25] In 2014, NIJ invested $1 million toward the development of the Center for Advanced Research in Forensic Science (CARFS) (award 2014-DNR-4918) and in 2021, NIJ invested an additional $0.16 million (award DJO-NIJ-21-RO-0002) [276]. The CARFS is an NSF Industry-University Cooperative Research Center (IUCRC), which is designed to strengthen cooperative research engagements between industry, academia, and government [277].

[26] Prior to 2014, forensic science research and development conducted by NIST was primarily supported through a cooperative agreement between NIJ and NIST Office of Law Enforcement Standards (OLES).

[27] Between 2015 and 2020, NIST invested approximately $20 million for a five-year award [278, 279]. In 2020, NIST renewed the award for an additional five years, investing an additional $20 million [280].

[28] In addition to NIJ, NSF, and NIST, the PCAST also recognized engagement from the Defense Forensic Science Center (DFSC) within the Department of Defense [13 pp36-38]. Between 2011 and 2015, the PCAST noted that DFSC received approximately $9.2 million to fund extramural research and development efforts; however, they also noted that DFSC, like NIJ, does not have a line item in the budget dedicated to forensic science research [13 p38].

3.2.1. Considerations for Forensic Science Research and Standards Programs

Challenges associated with the research and development infrastructure needed to sustain the needs of the forensic science community—both in terms of resources and strategy—have led to questions concerning the validity and reliability of forensic science practices within the criminal justice system [6, 13].[29] Although the forensic science community has witnessed an increased emphasis on the need for research, and the federal government has provided some resources toward that objective, resources remain limited [13]. This situation leads to three major considerations for forensic science research and standards programs (Box 3.2.1).

Box 3.2.1. Forensic science research and standards programs should:

(a) Assess the strengths and limitations of their existing research and development infrastructure, including their competencies, capabilities, and available resources to identify the most meaningful ways they can contribute to the broader forensic science research and development ecosystem (e.g., balancing their efforts across a spectrum of basic and applied research, short-term and long-term priorities, intramural and extramural funding, and support for method development and translation, including commercialization).

(b) Coordinate across different research and standards programs to establish a shared strategy for addressing the broader research and development needs of the forensic science community and, where relevant, align their respective resources toward those shared priorities and objectives.

(c) Ensure that investments in forensic science research and development include support for translation and implementation of their outputs so that they can be impactful and yield the greatest practical returns for the forensic science community and criminal justice system.

[29] Although the focus of this assessment is on the United States, similar concerns have been raised elsewhere, e.g., in a 2019 report, the United Kingdom Science and Technology Committee, appointed by the House of Lords, found: “Research and development in forensic science is under-resourced and uncoordinated. This has resulted in serious concerns about the scientific validity of some forensic science fields and the capability to provide evaluative interpretation of forensic science evidence. … Unless these failings are recognised and changes made, public trust in forensic science evidence will continue to be lost and confidence in the justice system will be threatened” [281 p4]. Further, the Committee noted “over the last 10 years only £56 million had been spent on 150 studies relating to forensic science. This accounted for a ‘relatively small percentage’ of their overall expenditure in that time, with the ‘annual expenditure of [UK Research and Innovation (UKRI)] over that 10-year period [being] roughly £6 billion.’ The percentage is less than 0.1%. … [However, t]he list of projects UKRI referred to in supplementary written evidence included under the category of forensic science many projects which, on analysis, did not address forensic science research questions, had little forensic science content[,] or which referred to forensic science as one of many possible applications of the research” [281 p43].

4. Societal Landscape

Issues and trends relating to public perception of forensic science within the context of the criminal justice system, and how that might impact the forensic science community, are evaluated in the societal landscape. Public perception can vary across several different dimensions; however, the one that underlies all of these is the extent to which society has trust and confidence[30] that the criminal justice system is fair, and the information relied upon is accurate. Looking through a societal lens, there are two key issues that are most relevant and potentially impactful to forensic science research and standards programs: (i) trust and confidence in the institution[31] and (ii) trust and confidence in the methods.[32]

4.1. Trust and Confidence in the Institution

Trust and confidence in the criminal justice system have been prevalent in national conversations for years [62]. Since 1993, the Gallup Organization has assessed levels of trust and confidence Americans have in different societal institutions, including the criminal justice system [63]. What they found is that between 1993 and 2022, less than one-fourth of Americans surveyed expressed “a great deal” or “quite a lot” of confidence in the criminal justice system, with 2022 marking the lowest point at 14% (ranging between a high of 34% in 2004 and a low of 14% in 2022, and sharply declining from 24% in 2020) [63]. Rather, between 1993 and 2022, approximately one-third of Americans surveyed expressed “none” or “very little” confidence in the criminal justice system, with 2022 marking the second worst perception on record at 46% (ranging between 49% in 1994 and 23% in 2004, sharply increasing from 36% in 2020) [63].

Although these data tell us how the public views the criminal justice system overall, information about forensic science specifically is limited. Further, the institution of forensic science is not bounded by the core functional components of the criminal justice system. Recognizing that forensic science activities are often performed by individuals operating within government, scientific, or law enforcement institutions, trust and confidence in forensic science as an institution might better be characterized as a mosaic of public perception across each of those individual institutional elements.

4.1.1. Government

Most forensic science activities in the United States are performed by government entities [49]. Likewise, most forensic practitioners analyze evidence in support of requests from other government, law enforcement, and prosecutorial agencies for use within government-led investigations or criminal prosecutions [6]. Consequently, characterizations of

[30] The terms “trust” and “confidence” refer to similar concepts for purposes of this environmental assessment—some sources refer to “trust” or “confidence,” whereas other sources refer to “trust and confidence.”
[31] I.e., that the institution is fair and the people that represent it will “do the right thing.”
[32] I.e., that the methods used for analyzing and reporting the evidence are accurate and reliable.

public trust and confidence in forensic science as an institution should include consideration of public trust and confidence in government. Although trust and confidence in government is a very broad issue, and not every dimension is relevant or impactful to forensic science, it is relevant to consider the issue at a macro level to assess general trends.

National and international surveys assessing various dimensions of public trust and confidence in government have been conducted year after year, and, with few exceptions, all point to trust in government being on a steady decline since the 1960s to near historic lows in 2022 [64]. According to data from the Pew Research Center, the percent of individuals surveyed within the United States who say they trust the government "to do what is right just about always/most of the time" ranged from 73% to 77% in the early part of the 1960s [64 p4]. Then, between the late 1960s and throughout the 1970s, trust in the federal government declined significantly [64 p4]. By 1980, 27% of individuals surveyed expressed trust in the federal government [64 p4]. Between 1980 and 2000, survey results fluctuated between a high of 49% in 1985 and a low of 17% in 1994 [64 p4]. In 2001, public trust in government nearly doubled immediately following the September 11th terrorist attacks (from 31% just before the attacks to 60% immediately after the attacks) [64 p4]. However, that surge was short-lived. By 2007, public trust in government returned to its pre-September 11th levels, only to decline further during the Great Recession to 17% by 2008 and 10% by 2011 [64 p4]. Between 2011 and 2022, on average, public trust in government increased marginally and held steady between 15% and 20% [64 p4]. Although there have been minor fluctuations since the September 11th terrorist attacks, the Pew Research Center noted that "[l]ow public trust in [the] federal government has persisted for nearly two decades" [64 p4].

The Pew Research Center notes confidence in career government employees to be much higher than trust in government overall; however, that too is on a decline [64 p9]. Between 2018 to 2022, confidence in career government employees at federal agencies declined by 9 percentage-points (from 52% in 2018 to 43% in 2022) [64 p9]. Another point of interest is that "[m]any say [the] federal government unfairly benefits some people over others," citing that 60% of individuals polled believe such a statement describes the government "extremely well" or "very well" [64 p14]. Related, less than 10% believe the government is (i) "[a]ble to address new problems as they come up," (ii) "[r]esponds to the needs of ordinary Americans," or are (iii) "[c]areful with taxpayer money" [64 p14]. Americans are united, however, in their views that the government does a good job in its handling of natural disasters, terrorism, food and medicine safety, and workplace standards, with the Pew Research Center citing approximately two-thirds of individuals in agreement with each of those claims [64 pp28-29].

Although these data from the Pew Research Center reflect key trends related to public perception of the federal government, they also indicate Americans have a more favorable view of their local and state governments compared to federal government [64]. The Pew Research Center noted that among those who shared an opinion about their local and state governments, "[a]bout two-thirds (66%) say they have a favorable view of their local government, compared with 54% who have a favorable view of their state government and just

32% who have a favorable view of the federal government"; however, between May 2018 and May 2022, favorable views of state and local governments were are also on the decline, decreasing by approximately 5% each [64 p44].

These statistics from the Pew Research Center are reinforced by other sources, including a 2022 Gallup poll assessing "trust in government" [65] and the 2022 Edelman Trust Barometer [66]. The Edelman Trust Barometer, which provides a measure of public trust on an international scale across four different societal institutions, including government, non-government organizations, business, and media, finds that society is now in a general state of "distrust," with trust in government nearly tied to that of media, with only 52% and 50% of respondents indicating trust, respectively [66 p5]. Compared to the results from May 2020, where trust in government was highest compared to the other three institutions (at 64%), trust in government declined by 12% in the two-year span covering the COVID-19 pandemic [66 p5]. Looking specifically at the United States, however, trust in government is lower at 39%, representing a decline of 3 percentage-points from the year prior [66 p42]. The Edelman Trust Barometer suggests that "distrust" is the new default, finding on average across all countries, 59% of individuals expressed a tendency to distrust until they see evidence that something is trustworthy versus the alternative tendency to trust until they see evidence that something is untrustworthy [66 p19]. This finding represents a global societal issue, with most individuals responding in this way in 24 countries [66 p19]. Further compounding the issue, on a two-dimensional axis comparing competency and ethics, government was rated negatively in both dimensions, worse than other societal institutions (non-government organizations and businesses) [66 p32]. The Edelman Trust Barometer finds that "information quality" is "now the most powerful trust builder across institutions" [66 p36] and concluded "restoring trust is key to societal stability," claiming "[e]very institution must provide trustworthy information [and c]lear, consistent, fact-based information is critical to breaking the cycle of distrust" [66 p37].

Although these trends relating to public trust and confidence in government reflect a challenging outlook overall, much of that appears to be impacted by trust and confidence in politicians rather than career government employees [64]. Nevertheless, trust in government as an institution "to do the right thing" can have widespread implications affecting how the public might perceive, and the extent to which they might trust, the work and testimony of government employees in the context of the criminal justice system.[33]

4.1.2. Science

Forensic science is broadly defined as "[t]he application of scientific principles and techniques to matters of criminal justice[,] especially as relating to the collection, examination, and analysis of physical evidence" [67]. In their 2009 report, the NRC recognized that "the term

[33] Despite that, one study suggests that jurors perceive expert witnesses in criminal trials who are employed by government institutions have greater credibility based on the (false) assumption that there are licensure requirements and those individuals are held to higher expectations than expert witnesses not employed by government institutions [282].

'forensic science' encompasses a broad range of disciplines, each with their own distinct practices, … [and s]ome activities require the skills and analytical expertise of individuals trained as scientists (e.g., chemists or biologists); other activities are conducted by scientists as well as by individuals trained in law enforcement (e.g., crime scene investigators, blood spatter analysts, crime reconstruction specialists), medicine (e.g., forensic pathologists), or laboratory methods (e.g., technologists)" [6 p38]. Further, the American Academy of Forensic Sciences asserts that "[a] forensic scientist is first a scientist" [68]. Thus, when characterizing public trust and confidence in forensic science as an institution, it must also include consideration of public trust and confidence in science generally.

As public trust and confidence in science change over time, they can have implications for policy and public decision-making processes. As a result, understanding public trust and confidence in science has been a point of interest among societal leaders and thus, has been assessed in various ways through public surveys and polls [66, 69–71]. Over the last several decades, trust in science as an institution has remained relatively stable, with public trust and confidence in science being generally favorable [69]. According to the General Social Survey (GSS), administered by the National Opinion Research Center (NORC) at the University of Chicago, which has monitored societal changes in various social institutions since 1972, on average over the last several decades (between 1972 and 2018), approximately 40% of individuals surveyed express a "great deal of confidence" and approximately 45% express "some confidence" in scientific leaders, yielding a combined result of approximately 85% [69]. In their 2013 final report, NORC noted that "[p]ublic confidence in the scientific community stands out as among the most stable of about a dozen institutions rated in the GSS since the mid-1970s" [72]. In recent years, however, the Pew Research Center finds that trust and confidence has taken a downward turn [70]. According to their data, in 2019, 35% of Americans expressed a "great deal of confidence" and 51% expressed "a fair amount of confidence" in scientific leaders "to act in the best interest of the public" (86% combined), whereas only 13% expressed "not too much [or] no confidence at all" [70 p4]. However, just two years later, in December 2021 those values dropped to 29% expressing a "great deal of confidence" and 48% expressing "a fair amount of confidence" (77% combined) and those who expressed "not too much [or] no confidence at all" increased to 22% [70 p4]. Although these data reflect a downward trend during the COVID-19 pandemic, they still indicate that most Americans (approximately three-fourths) have trust and confidence in science [70 p4]. While these data from the Pew Research Center focus heavily on public perception among Americans, the 2022 Edelman Trust Barometer reflect a similar sentiment on a global scale, showing 75% of individuals expressing trust in scientists—ranking science above all other institutions evaluated [66 p14].

Although public trust and confidence in science is generally high, there are some in the population who do not have such a favorable view and the scientific community has sought to better understand why and what factors might influence those perspectives. In 2015, the NAS held a roundtable entitled "Trust and Confidence at the Interfaces of the Life Sciences and Society: Does the Public Trust Science?," which explored these issues in greater depth and sought to "highlight research on the elements of trust and how to build, mend, or maintain trust; and examine best practices in the context of scientist engagement with lay audiences around social issues" [73 p3]. What they found is that trust is a complicated topic, and there is "no universally accepted definition of trust" [73 p5]. Instead, trust is "a complex landscape in which personal characteristics like culture, religion, values, and personal histories—when combined with science's own shortcomings like inconsistent findings and conflict of interest—can promote lack of trust in both scientists and the scientific enterprise" [73 p18]. One workshop participant expanded on this, noting:

> *[N]o one is questioning the scientific method, they are questioning whether scientists adhere to it. All the things causing concern about science—such as the problem in being able to reproduce results in medical and other research—risk the good standing of the scientific enterprise.* [73 p19]

Workshop participants also noted, however, that "[s]cience does not exist in isolation[; rather, p]ublic discussions of new discoveries, applications, or concerns about science take place in a sociopolitical context" [73 p21]. For example, in the context of public policy, one workshop participant noted that as "debates about an issue become more polarized, science is increasingly used as the 'ultimate trump card' [and p]oliticians bring science and scientists into the policy arena in an effort to say 'the science is on my side'" [73 p21]. Participants suggest that this can have a damaging effect, with one noting "if the motive of the scientist is [perceived to be] anything other than seeking knowledge … we run the risk of saying that the scientist is just another form of advocate who is selectively using evidence in order to engage in a persuasive campaign" [73 p21]. The workshop concluded with the following note:

> *Scientists should … specify the limits of what is not known according to scientific norms. Scientists also have to trust that a public audience is capable of understanding the science. … [T]oo often the public experiences the voice of science as disdainful—a voice that says, "We have consensus; now accept it." That is an appeal to authority, which … does not go over well in the modern world. The modern world reacts poorly to authority, but is responsive to a voice that lets the audience participate in the process of understanding and draw its own conclusions. … Scientists who engage rather than only disseminate can become more effective in the public arena.* [73 p38]

In 2017, the NAS held another roundtable, entitled "Examining the Mistrust in Science," which sought to "explore trends in public opinion of science, examine potential sources of mistrust, and consider ways that cross-sector collaboration between government, universities, and industry may improve public trust in science and scientific institutions in the future" [74]. In this roundtable, workshop participants echoed earlier findings, claiming "[t]here seems to be an erosion of the standing and understanding of science and engineering among the public[, and p]eople seem much more inclined to reject facts and evidence today than in the recent past" [74 p1]. These observations are represented in the data as well—workshop participants referenced data from the Pew Research Center and noted that when looking historically at trends related to public opinions of science, "there is no single story about what people think about science across all issues" [74 p2]. Instead, "[s]ome science issues are strongly driven by politics and ideology, some are strongly connected with religious beliefs, and others with gender" [74 p2]. Ultimately, different viewpoints were discussed in attempts to characterize the factors that have contributed to this feeling of "mistrust" and there seems to be little consensus on how to best address it, particularly as science continues to be a factor of debates affecting public policy [74].

A few years after the NAS convened the second roundtable on issues of public "mistrust of science," the COVID-19 pandemic reinforced their findings. In 2021, science journalist Ryan Cross examined the historical patterns of trust in science and the impact of the pandemic [75]. Not only did the pandemic place science and scientists at the forefront of public policy with global influence, but it had immediate and material consequences on public health, welfare, and economic security [75]. Although there was a brief period of public unity at the beginning of the pandemic, as scientists became more and more influential in shaping public policy, public perception and viewpoints became more polarized [75]. As Cross noted:

> *The pandemic has placed science under the public microscope. Never before has a scientific issue so immediately and directly affected the lives of everyone at the same time. And never before have scientists created so much new information on one subject so quickly. By early January 2021, PubMed had documented more than 90,000 publications mentioning "COVID-19." … Most of those papers were not flashy enough to make headlines. The ones that did make the spotlight in the news or on social media typically revealed new, surprising, or contradictory results, with implications that put people either at ease or on edge.* [75]

Cross continued:

> *Scientists usually debate and resolve questions … over years, sometimes decades, in conference halls and academic journals. The urgency of the pandemic has warped that scientific process and put its accelerated results on display. … All of the sudden everyone is watching scientists, and they are seeing all the messiness of how science happens.* [75]

Finally, Cross suggested that "[i]t is tempting, but inaccurate, to boil the problem down to one of science education" [75]. Cross continued, noting that people who argue about scientific issues "are not antiscience"; rather, they "just disagree on who is the legitimate expert and who's got the right science" [75].

Although the majority of people surveyed have favorable views in terms of trust and confidence in science, the underlying issues are complex; and while much of the data on this topic focus on science in the context of public policy and ideology, it illustrates an equally challenging backdrop when science is used in the criminal justice system, which is inherently adversarial yet requires immediate and final resolution on complex issues. Such backdrop is particularly relevant for forensic scientists, whose primary responsibility is to apply scientific inferences to disputed matters in the sensitive domain of criminal justice.

4.1.3. Law Enforcement

In 2009, the NRC recognized that "forensic science practice was started in, and has grown out of, the criminal justice and law enforcement systems … [and m]any forensic techniques were developed to aid in the investigatory phase of law enforcement and then were adapted to the role of aiding in prosecution by providing courtroom testimony[; t]hus, forensic practitioners who work in public crime laboratories often are seen as part of the [law enforcement or] prosecution team" [6 p187]. Although many forensic practitioners are civilians, "[s]ome forensic tests might be conducted by a sworn law enforcement officer" [6 p36] and most forensic science activities are carried out under the administrative controls of law enforcement [6 pp183-185, 27]. Consequently, characterization of public trust and confidence in forensic science as an institution must also include consideration of trust and confidence in law enforcement.

Trust in law enforcement has long been a measure of societal health [76]. As trust declines, it can undermine the legitimacy of law enforcement [76]. Without legitimacy, police lose their ability and authority to function effectively and maintain law and order [76, 77]. Consequently, understanding public trust and confidence in law enforcement has been a point of focus among government and societal leaders for decades and has been assessed across various dimensions through public surveys and polls [63, 70, 78].

According to data from the Gallup Organization, between 1993 and 2022 the percentage of individuals within the United States who say they have "a great deal" or "quite a lot" of confidence in police averaged 55% and ranged between 64% (in 2004) and 45% (in 2022), with the data reflecting a general decline over the last decade [63]. Those who responded as having "a great deal of confidence," however, represented the sharpest decline over the last five years, between 2017 to 2022 (decreasing from 31% to 19%, respectively) [63]. Similarly, the percentage of Americans who say they have "very little" confidence in police has steadily increased from 7% in 2005 to a record high of 17% in 2022 [63]. The Pew Research Center finds similar trends, noting a 10 percentage-points decrease (from 30% to 20%) of individuals responding as having "a great deal" of confidence in police to "act in the best interest of the public" between 2018 and 2021 [70]. Similarly, the percentage of individuals responding as having "not too much [or] no confidence at all" increased from 22% in 2018 to 31% in 2021 [70]. Collectively, these data suggest that public trust and confidence in police overall is it at a record low [63, 70].

Related, according to the GSS, public perception of national spending on law enforcement experienced its most significant change between 2018 and 2021 [78]. Since it was first measured in 1985, public perception that national spending on law enforcement was "too much" fluctuated between 5% and 10% until 2009 [78]. Then, between 2010 and 2017, it fluctuated between 10% and 14% [78]. However, between 2018 and 2021, it increased from 10% in 2018 to 23% in 2021 [78]. Similarly, public perception that national spending on law enforcement was "too little" declined from 51% in 2018 to 41% in 2021 [78]. Between 2020 and 2021, however, the Pew Research Center noted a different trend, with more Americans expressing support for spending on police *in their area* (support for increased spending rose from 11% in 2020 to 21% in 2021 and support for decreased spending declined from 12% in 2020 to 6% in 2021) [79]. The Pew Research Center noted, however, that “changing attitudes about police spending in their area have occurred amid rising public concern about violent crime” [79].[34]

While these data reflect the American point of view in aggregate, the Pew Research Center noted “[v]iews on police funding continue to differ widely by race and ethnicity, age and political party” [79] and the Gallup Organization points out that Americans’ confidence in police differs across racial lines, noting in 2020 that “[f]ifty-six percent of White adults and 19% of Black adults say they have "a great deal" or "quite a lot" of confidence in the police [and t]his 37-percentage-point racial gap is the largest found for any of 16 major U.S. institutions rated in Gallup's annual Confidence in Institutions poll” [80].[35] In 2021, the Gallup Organization reported a slight improvement but overall public perception “remains low” [81].

In response, over the last few years, many law enforcement agencies have prioritized efforts to improve transparency and accountability, such as turning to body-worn cameras, to strengthen public perception and rebuild that public trust and confidence [82 p3, 83].

Although much of the data reflecting public opinions of law enforcement are often discussed in the context of concerns over appropriate use of force, accountability for misconduct, fair treatment between police officers and members of the public, and concerns over rising crime across the nation, these issues reflect an overall erosion of confidence and public trust in law enforcement as an institution [63, 70]. These issues can have implications for law enforcement activities in other contexts, such as those involving collection, analysis, interpretation, and testimony of forensic evidence within the criminal justice system by individuals employed by law enforcement entities.

4.1.4. Considerations for Forensic Science Research and Standards Programs

Generally, most Americans do not have high trust and confidence in government [64–66], law enforcement [63, 70], or the criminal justice system [63]. Science is the only institution related to forensic science where most Americans expressed favorable trust and confidence [66, 69–72]; however, this can be impacted when scientists offer differing or contradictory results or views on issues, or when those results or views are presented in a way that might be perceived as biased by external motivators [73, 74].

These statistics provide a reflection of how the public might perceive, and the extent to which they might trust, the work and testimony of forensic practitioners. In recent years, many

[34] Although it is difficult to assess the actual cause for the trends in public perception of spending on law enforcement, the way the questions were framed between the GSS and Pew Research Center surveys could have an impact on how participants responded—the former framing the question on a national level and the latter framing the question on a local level.
[35] The survey in 2020 “was conducted after George Floyd was killed while in police custody in Minneapolis in late May” [80].

have looked to NIST as a viable entity that is external and impartial to the criminal justice system to advise on science-based issues affecting the forensic science community [6, 13]. NIST too has recognized the importance of its role in strengthening the trust of critical institutions that our society relies upon—in its 2020 Environmental Scan, NIST concluded: "In the current climate of distrust and skepticism, especially mistrust of science, there are fewer institutions that remain avowedly objective and unbiased … [and a]s an organization that prides itself on its objectivity and lack of scientific bias, NIST has positioned itself as the federal 'trust' agency … with strong scientific underpinnings that lack political agendas" [84 p12].

In the context of forensic science and the criminal justice system, there are four major considerations for forensic science research and standards programs (Box 4.1.4).

Box 4.1.4. Forensic science research and standards programs should:

(a) Ensure transparency, objectivity, neutrality, and impartiality when administering programmatic activities, such that the public can build and maintain trust that their scientists are credible, and that their outputs are not unduly influenced (real or perceived) by individual interests or agendas.

(b) Take an active and assertive stance on matters relating to good scientific principles or practices—even when issues are controversial, or viewpoints are adversarial.

(c) Speak with a consistent voice on matters related to science and standards and be able to communicate and engage with all members of the criminal justice system and the public on how to recognize sound scientific and technical practices.

(d) Establish ways for forensic service providers to demonstrate their adherence to sound scientific and technical practices and conformance to recognized standards, guidelines, and recommendations through impartial and independent third-party entities rather than relying solely on self-declarations.[36]

4.2. Trust and Confidence in the Methods

Forensic science evidence can impact decisions and outcomes within the criminal justice system [85–87]. It is not only influential on juror decision making [85], but also "many [other] decision points in the justice process," including those made by investigators and prosecutors [86 pS89]. Those decisions are often based on the extent to which the evidence presented is

[36] Conformity assessment is traditionally carried out by third-party accrediting bodies. Non-government accrediting bodies provide fee-based conformity assessment services to forensic service providers that voluntarily choose to use those services to demonstrate conformance to national or international standards, such as ISO/IEC 17025:2017 and ISO/IEC 17020:2012. However, those assessments do not necessarily require conformance to discipline-specific forensic science standards such as those recognized by the OSAC Registry [169]. Instead, forensic service providers provide self-declarations of conformance to discipline-specific standards and guidelines [273].

perceived to be accurate and reliable [85, 87]. Thus, understanding public perceptions of the accuracy and reliability of forensic methods is important to assessing whether those perceptions are properly calibrated and contributing to a fair and effective administration of justice [88].

Up until 2009, the accuracy and reliability of forensic science evidence was largely unquestioned, and public perception was that many traditional forensic science disciplines were on solid scientific footing given their decades of use [5, 7]. By 2009, however, this began to change, particularly within the scientific community, when the NRC claimed that "the interpretation of forensic evidence is not always based on scientific studies to determine its validity" [6 p8], and questioned "whether—and to what extent—there is *science* in any given 'forensic science' discipline" [6 p87]. Following a more thorough and critical evaluation, in their 2009 report, the NRC noted "[l]ittle rigorous systematic research has been done to validate the basic premises and techniques in a number of forensic science disciplines" [6 p189] and advances in DNA technology "have revealed that, in some cases, substantive information and testimony based on faulty forensic science analyses may have contributed to wrongful convictions of innocent people" [6 p4]. In the years following the 2009 NRC report, scientists, scholars, litigators, and popular media continued to raise concern about the limited empirical research underpinning the validity and reliability of some forensic science methods and questioned the extent to which the public *ought* to trust and have confidence in forensic evidence resulting from those methods [13, 89–97].

Over the last decade, various studies have measured public perception of the accuracy and reliability of forensic science methods in different ways [8, 87, 96, 98–100]. In 2022, data from multiple studies published between 2008 and 2019 were combined to assess changes over time and to compare results across disciplines, providing the most recent and comprehensive assessment of public perception of the accuracy and reliability of forensic methods to date [88]. Although the various studies had some methodological differences and not all disciplines were accounted for in each of the studies, from these data the authors drew three broad inferences:

> *First, the public views many practices (even DNA in the most recent study) as quite fallible. … Second, there is considerable miscalibration between public views of reliability and the scientific consensus. Third, there may be early indications that public perceptions of forensic science's reliability are falling.* [88 p273]

As it relates to trends over time, on a scale from 0 (no reliability) to 100 (completely reliable), in 2008, participants rated DNA as approximately 94%; however, in 2012 this value dropped to 89% and in 2019 it dropped to 83% [88]. Similarly, in 2008, participants rated fingerprint evidence as approximately 90%; however, in 2012 this value dropped to 82% and in 2019 it dropped to 78% [88]. Although the data are less complete for other disciplines, some other disciplines experienced similar declines [88].[37] For example, between 2015 and 2019, blood pattern analysis declined from approximately 78% to 64% and firearms and toolmarks

[37] See table 1 and figure 1 in reference [88 p274] for a summary across all studies and disciplines.

declined from approximately 79% to 68% [88]. As it relates to how the public perceives certain disciplines compared to others, the authors noted:

> *While DNA, along with fingerprints, is generally perceived as highly reliable in accordance with the current evidence base, it was not rated very differently [compared] to forensic dentistry/bitemark analysis, which may be cause for concern. … Bitemark analysis has never been validated, likely has a high error rate, and has contributed to many wrongful convictions—and so a public perception that weighs DNA analysis similarly to bitemarks would be poorly calibrated to the evidence. … The studies also revealed other miscalibrations between respondent perceptions and what the evidence base currently suggests—such as lower than warranted views on document analysis, and possibly higher than warranted views on firearm/toolmark analysis.* [88 p275]

From these data, the authors concluded:

> *While some forensic practitioners have historically claimed infallibility in their methods, emerging research not only demonstrates that this is not the case, but also that the public does not uniformly believe this to be the case. … If forensic science is to maintain its reputation or even improve it, we recommend that forensic scientists critically examine their research and practices and consider aligning them with the move toward openness and transparency occurring elsewhere within science.* [88 p281]

4.2.1. Considerations for Forensic Science Research and Standards Programs

Ideally, public trust and confidence in various forensic science disciplines are calibrated to the *actual* measures of accuracy and reliability—"where the public's views of forensic scientific research claims reflect the strength of the evidence behind them" [88 p272]. For public perception of the accuracy and reliability to be properly calibrated, there must be empirical studies demonstrating what those actual measures are, and care must be taken to ensure that those measures are generated through properly designed studies that have been administered or evaluated by a credible source. As to what is considered a "credible" source, the PCAST claimed "[t]o ensure that the scientific judgments are unbiased and independent, such evaluations must clearly be conducted by a science agency with no stake in the outcome" and "agencies that apply forensic [science] methods within the legal system have a clear stake in the outcome of such evaluations" [13 p124].[38] This leads to five major considerations for forensic science research and standards programs (Box 4.2.1).

[38] The PCAST recommended: "This responsibility should be lodged with NIST. … Our intention is not that NIST have a formal regulatory role with respect to forensic science, but rather that NIST's evaluations help inform courts, the DOJ, and the forensic science community" [13 p124]. Similar recommendations were made by the NCFS [35, 283]. In fiscal year 2018, Congress appropriated funding for NIST to begin conducting technical merit evaluations (also referred to as scientific foundation reviews) [153 p22].

Box 4.2.1. Forensic science research and standards programs should:

(a) Ensure scientific evaluations of forensic methods and practices are conducted in a transparent, consistent, and systematic manner.[39]

(b) Approach scientific evaluations of forensic methods and practices in a manner that has a clearly defined scope that is broad enough to produce meaningful information but narrow enough to ensure it can be completed within a timeframe that is acceptable to the forensic science community given the resources available.[40]

(c) Ensure evaluations of forensic methods and practices not only focus on the capabilities and limitations to accurately and reliably answer specific and clearly defined forensic questions, but also include a discussion of whether there are other forensic questions for which the forensic methods or practices could produce accurate and reliable answers.[41]

(d) Have well-defined research plans that include contributions to the development of actual measures of accuracy and reliability of different forensic methods and practices to promote and enable better calibration of public perception.

(e) Prioritize outreach, using various approaches to reach broad and diverse audiences, to inform the public about the results of scientific evaluations and results of research that will contribute to calibrating what the public perception of the accuracy and reliability of forensic science methods *should* be, which is grounded by *actual* measures of accuracy and reliability, to ensure the public is not misled and unduly over- or under-values the evidence presented in the context of a particular case.

5. Scientific and Technological Landscape

Issues and trends relating to the rigor and capacity of the methods, techniques, and practices underlying the collection, examination, and analysis of physical evidence, and how that might impact the forensic science community, are evaluated in the scientific and technological landscape. Issues relating to "science" and "technology" are often framed

[39] The NCFS noted that "[e]valuations of all aspects of technical merit must be respected by all stakeholders if these evaluations are to be utilized by the legal and scientific communities" [283 p3]. Such reviews should be carried out in way that everyone understands how they are done, which criteria inform key decisions, and what measures were used to ensure fair and impartial procedures for conducting the reviews.
[40] An acceptable timeframe should be a point of discussion with community members; however consistent with the original vision of the PCAST calling for annual reports, reviews should not take more than 12 months to carry out [13 p124].
[41] Although some forensic methods or practices might not be reliable for answering a particular forensic question in one context, they might have intrinsic value and be appropriate for answering a different forensic question in another context [284]. "It makes no sense to assess the reliability of any forensic technique in the abstract. A forensic method is only 'reliable' as far as it helps answer the particular questions asked in the context of a particular case. Asking the wrong questions will undoubtedly deliver the wrong answers, even if the best and most fully validated forensic method is applied. Conversely, some forensic methods are perceived by some commentators to have less intrinsic value or even questionable reliability. But these methods might yield the answer to a crucially relevant question." [284].

through the lens of innovation—aimed at producing new or improved methods, techniques, and tools to yield new capabilities with better accuracy, faster turn-around times, lower costs, smaller footprint, or greater accessibility and applicability to diverse materials, environments, and conditions. However, limited resources often constrain what is practically achievable. Thus, investments in innovation must be approached in a deliberate way, ensuring needs are identified and prioritized in ways that align to longer-term visions and strategies.

Indeed, an aim of the NRC Committee on Identifying the Needs of the Forensic Science Community was to "chart an agenda for progress in the forensic science community and its scientific disciplines" [6 pxix]. What they found was "a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods" [6 p8] and "call[ed] for *real science* to support the forensic disciplines" [101 p2]. The 2009 NRC report played an important role in the science and technology landscape of forensic science, causing the community to consider not only what research and development is necessary to fill critical gaps but also how those developments can be successfully translated to practice [6].

Looking through a scientific and technological lens, therefore, there are three key issues that are most relevant and potentially impactful to forensic science research and standards programs: (i) validity and reliability, (ii) research and development, and (iii) translation and implementation.

5.1. Validity and Reliability

In their 2009 report, the NRC provided a general description of the principles of science and interpreting scientific data, including the general process for conducting studies to establish the validity and reliability of analytical methods and various characteristics and performance metrics that must be addressed in those studies [6 pp111-125]. In their description, the NRC often referred to well-established criteria set forth by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) Standard 17025 (General Requirements for the Competence of Testing and Calibration Laboratories) for the validation of novel analytical methods and point to various techniques that can be used to validate a method [6 pp113-116]. However, despite the availability of these criteria through the international standard, the NRC noted "much forensic evidence … is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing to explain the limits of the discipline" [6 pp107-108]. Following their review of the scientific foundation supporting various forensic science disciplines [6 pp127-182], the NRC concluded "[l]ittle rigorous systematic research has been done to validate the basic premises and techniques in a number of forensic science disciplines … [yet t]he committee sees no evident reason why conducting such research is not feasible" [6 p189].[42] Specifically, they viewed the following methods as having sound scientific foundations:

[42] The Committee was most critical of those methods that rely heavily on expert interpretation of observed patterns and claims relating to source attribution (i.e., individualization or identification of a specific source as the donor), such as friction ridge (fingerprints), shoeprints and tire tracks, toolmarks and firearms, handwriting, bitemarks, bloodstain patterns, fire debris, hair, and fiber analyses.

- DNA analysis [6 p133]
- Controlled substances analysis (for known substances) [6 pp135-136]
- Analysis of explosives and fire debris (as it relates to the identification of ignitable liquid residues but not for reconstruction purposes or determinations of whether or not a fire was intentionally set) [6 pp172-173][43]

The NRC viewed the following methods as having potential for sound scientific foundations, but needing more research to establish those foundations and greater standardization in practices to ensure the examinations can be conducted in a repeatable and reproducible manner before they can be considered as valid and reliable:

- Friction ridge analysis [6 pp142-145]
- Shoeprints and tire tracks [6 pp149-150]
- Firearms and toolmarks [6 pp154-155]
- Analysis of hair evidence [6 pp160-161][44]
- Handwriting comparisons [6 pp166-167]
- Bloodstain pattern analysis [6 pp178-179]
- Analysis of fiber evidence [6 pp162-163]
- Analysis of paints and coatings [6 p170]

Finally, the NRC viewed bitemark analysis as lacking “evidence of an existing scientific basis” and concluded more “research is warranted in order to identify the circumstances within which the methods of forensic odontology can provide probative value” [6 pp175-176].

When reflecting on the 2009 NRC findings to the NCFS in 2014, the Hon. Harry Edwards, co-chair of the NRC Committee on Identifying the Needs of the Forensic Sciences Community, expounded on this issue, claiming “[m]y fellow Committee members and I were astonished by the paucity of rigorous scientific research testing the validity and reliability of the forensic disciplines” [101 p2]. Edwards continued, claiming *“[j]udicial review, by itself, will not cure the infirmities of the forensic community,”* arguing these issues can only be solved by the forensic science community embracing a scientific culture [101 p2]. Specifically, Edwards commented:

> *When our Committee issued its Report, I heard a number of very smart people suggest that once lawyers began to introduce the Report in judicial proceedings, judges would limit the admissibility of forensic evidence and issue seminal decisions that would result in dramatic reforms of the forensic disciplines. I did not believe it then, and I do not believe it now. Absent meaningful action by scientists and forensic analysts, the courts*

[43] The NRC noted: “The scientific foundations exist to support the analysis of explosions, because such analysis is based primarily on well-established chemistry. ... [However, e]xperiments should be designed to put arson investigations on a more solid scientific footing” [6 pp172-173].

[44] The NRC noted: “The committee found no scientific support for the use of hair comparisons for individualization in the absence of nuclear DNA. Microscopy and mtDNA analysis can be used in tandem and may add to one another’s value for classifying a common source, but no studies have been performed specifically to quantify the reliability of their joint use” [6 p161].

> *will continue to admit forensic evidence in criminal trials, without regard to its scientific validity and reliability. Why? Because precedent supports this practice. Yes, there have been a few trial court decisions that have limited the admission of some forensic evidence; but, to date, there has not been a single federal court of appeals decision that has curbed its admissibility. … The simple truth is that the Supreme Court is not going to issue a seminal decision like Brown v. Education to change the culture of the forensic community. The burden falls on the scientific community to get this done.*
>
> *… [E]ven as many judges have recognized that the methods used by [various forensic] experts have not been scientifically verified, they have continued to admit questionable and overdrawn testimony from forensic practitioners on the grounds that such evidence has been relied upon in the justice system for many years. Each ill-informed decision becomes a precedent binding future cases. The courts will not be able to move beyond this misguided precedent until real science is brought to bear in assessing the validity and reliability of forensic disciplines and in establishing quantifiable measures of uncertainty in the conclusions of forensic analyses. Judges, lawyers, and jurors need an honest accounting from scientists. Without it, we will continue to default to past judicial decisions that overestimate the scientific validity of forensic disciplines.* [101 pp2-3]

Following the 2009 NRC report, large scale black-box and white-box testing schemes[45] were conducted to measure the accuracy, repeatability, and reliability of widely practiced forensic science disciplines that rely on expert interpretation, with an initial emphasis on friction ridge (fingerprints) and firearms analyses [102–108]. Then, in 2015, the PCAST was tasked with assessing "whether there are additional steps on the scientific side, beyond those already taken by the Administration in the aftermath of the highly critical 2009 [NRC] report on the state of the forensic sciences, that could help ensure the validity of forensic evidence used in the Nation's legal system" [13 px]. In September 2016, the PCAST issued their report and "concluded that there are two important gaps: (1) the need for clarity about the scientific standards for the validity and reliability of forensic methods and (2) the need to evaluate specific forensic methods to determine whether they have been scientifically established to be valid and reliable" [13 px].[46]

In their 2016 report, the PCAST expounded on the principles outlined by the NRC and took a more prescriptive approach when describing the scientific criteria for validity and reliability, breaking it down into two distinct concepts: (i) "foundational validity" and (ii) "validity as applied" and specified several elements that must be accounted for when making a determination of scientific validity and reliability [13 pp44-46]. In general, the PCAST claimed:

---

[45] Black-box testing refers to studies designed to evaluate the output of a method or process (i.e., analyst's decision) without regard to how those outputs are produced. White-box testing refers to studies designed to evaluate the basis for the output (i.e., analyst's decision) of a method or process.

[46] The PCAST focused on "feature comparison" methods—specifically, methods for comparing DNA samples, bitemarks, latent fingerprints, firearm marks, footwear, and hair.

> *Scientific validity and reliability require that a method has been subjected to empirical testing, under conditions appropriate to its intended use, that provides valid estimates of how often the method reaches an incorrect conclusion. For subjective feature-comparison methods, appropriately designed black-box studies are required, in which many examiners render decisions about many independent tests (typically, involving "questioned" samples and one or more "known" samples) and the error rates are determined. Without appropriate estimates of accuracy, an examiner's statement that two samples are similar—or even indistinguishable—is scientifically meaningless: it has no probative value, and considerable potential for prejudicial impact. Nothing—not training, personal experience nor professional practices—can substitute for adequate empirical demonstration of accuracy.* [13 p46]

Specifically, the PCAST defined the concepts of "foundational validity" and "validity as applied" as follows:

For "foundational validity," the PCAST stated:

> *[The] method has been subjected to empirical testing by multiple groups, under conditions appropriate to its intended use [and t]he studies must (a) demonstrate that the method is repeatable and reproducible and (b) provide valid estimates of the method's accuracy (that is, how often the method reaches an incorrect conclusion) that indicate the method is appropriate to the intended application.* [13 pp5, 65-66]

For "validity as applied," the PCAST stated:

> *The forensic examiner must have been shown to be capable of reliably applying the method and must actually have done so. … Determining whether an examiner has actually reliably applied the method requires that the procedures actually used in the case, the results obtained, and the laboratory notes be made available for scientific review by others. … [Further, t]he practitioner's assertions about the probative value of [the evidence] must be scientifically valid. The expert should report the overall false-positive rate and sensitivity for the method established in the studies of foundational validity and should demonstrate that the samples used in the foundational studies are relevant to the facts of the case. … And the expert should not make claims or implications that go beyond the empirical evidence and the applications of valid statistical principles to that evidence.* [13 pp6, 65-66]

The PCAST further specified the "key criteria for validation studies to establish foundational validity" as follows:

> *Scientific validation studies—intended to assess the validity and reliability of a metrological method for a particular feature-comparison application—must satisfy a number of criteria.*

(1) *The studies must involve a sufficiently large number of examiners and must be based on sufficiently large collections of known and representative samples from relevant populations to reflect the range of features or combinations of features that will occur in the application. In particular, the sample collections should be:*

a. *representative of the quality of evidentiary samples seen in real cases. (For example, if a method is to be used on distorted, partial, latent fingerprints, one must determine the random match probability—that is, the probability that the match occurred by chance—for distorted, partial, latent fingerprints; the random match probability for full scanned fingerprints, or even very high quality latent prints would not be relevant.)*

b. *chosen from populations relevant to real cases. For example, for features in biological samples, the false positive rate should be determined for the overall US population and for major ethnic groups, as is done with DNA analysis.*

c. *large enough to provide appropriate estimates of the error rates.*

(2) *The empirical studies should be conducted so that neither the examiner nor those with whom the examiner interacts have any information about the correct answer.*

(3) *The study design and analysis framework should be specified in advance. In validation studies, it is inappropriate to modify the protocol afterwards based on the results.*

(4) *The empirical studies should be conducted or overseen by individuals or organizations that have no stake in the outcome of the studies.*

(5) *Data, software, and results from validation studies should be available to allow other scientists to review the conclusions.*

(6) *To ensure that conclusions are reproducible and robust, there should be multiple studies by separate groups reaching similar conclusions.* [13 pp52-53]

The PCAST then applied these criteria to several forensic feature comparison methods and offered their view as to whether, and to what extent, these criteria had been satisfied [13 pp67-123].[47] While the PCAST recognized the emergence of a culture shift in the forensic science

[47] The disciplines considered by the PCAST included: DNA analysis of single-source and simple-mixture samples, DNA analysis of complex-mixture samples, bitemarks, latent fingerprints, firearms identification, and footwear analysis. The PCAST was able to make assessments of "foundational validity" but often were not able to express a determination on "validity as applied" since that required evaluation of how a specific method was applied in the context of a specific case. The PCAST did, however, indicate what factors should be accounted for to enable an assessment of "validity as applied" in the context of a specific case.

community toward acknowledging the need for empirical studies to assess the validity and reliability of various forensic methods in the years since the 2009 NRC report [13 pp63-65], and they applauded some research that had been done since then,[48] they note that their "observations and findings … are largely consistent with the conclusions of earlier NRC reports" [13 p69]. Specifically, the PCAST viewed the following methods as having been shown to be "foundationally valid":

- DNA analysis of single-source samples or simple mixtures of two individuals [13 p75]
- DNA analysis of complex mixtures (based on probabilistic genotyping software) [13 p82][49]
- Latent fingerprint analysis [13 p101]

However, the PCAST asserted the following methods had not yet been shown (as of 2016) to be "foundationally valid":

- DNA analysis of complex mixtures (based on combined probability of inclusion-based approaches) [13 p82]
- Bitemark analysis [13 p87][50]
- Firearms analysis [13 p112][51]
- Footwear analysis (identifying characteristics) [13 p117][52]

In their report, the PCAST took a more narrow and explicit stance on their assessment of whether various methods were considered scientifically valid and reliable compared to the NRC [13], which stimulated various responses on these issues [109].[53] In the years following the PCAST report, additional efforts were undertaken to contribute to the research-base by conducting large scale black-box, white-box, and interlaboratory testing schemes assessing the accuracy and reliability of several other disciplines and methods, including footwear examination [110–113], palmar friction ridge comparisons [114], bloodstain pattern analysis [115], firearms examination [116, 117], handwriting comparisons [118], DNA mixture interpretation [119], digital evidence examinations [120], facial identification examinations [121], and trace evidence examinations (e.g., hair [122], tape [123, 124], glass [125–127]).

---

[48] Specifically, the PCAST recognized references [102, 104] as appropriately designed black-box studies for latent fingerprints and firearms identification, respectively.

[49] The PCAST noted: "At present, published evidence supports the foundational validity of analysis, with some programs, of DNA mixtures of 3 individuals in which the minor contributor constitutes at least 20 percent of the intact DNA in the mixture and in which the DNA amount exceeds the minimum required level for the method. The range in which foundational validity has been established is likely to grow as adequate evidence or more complex mixtures is obtained and published" [13 p82].

[50] The PCAST noted "bitemark analysis … is far from meeting such standards. To the contrary, available scientific evidence strongly suggests that examiners cannot consistently agree on whether an injury is a human bitemark and cannot identify the source of bitemark with reasonable accuracy" [13 p87].

[51] The PCAST noted: "firearms analysis currently falls short of the criteria for foundational validity, because there is only a single appropriately designed study to measure validity and estimate reliability. The scientific criteria for foundational validity require more than one such study, to demonstrate reproducibility" [13 p112].

[52] The PCAST noted: "PCAST has not evaluated the foundational validity of footwear analysis to identify class characteristics (for example, shoe size or make)" [13 p117].

[53] E.g., ASCLD argued that the PCAST took too narrow of a view in their assessment of research, dismissing "a wealth of existing research because it does not meet an arbitrary criteria of black box studies with an ideal sample size" [17].

In 2017, the American Association for the Advancement of Science (AAAS) issued reports on fire investigation and latent fingerprint examination, assessing the extent to which there is research underlying the scientific validity and reliability of the methods and charting a specific research agenda to strengthen those foundations moving forward [128, 129]. Specifically, as it relates to fire scene investigation, the AAAS report concluded "[l]ittle is known about the consistency and accuracy of conclusions among experienced investigators when presented with the same data" [128 p7]; however, they concluded fire debris analysis (the examination of samples from fire scenes for traces of ignitable liquid residues) is "based on established scientific principles … [and t]he validity and reliability of fire debris analysis is well established … [but] there is still work to be done" [128 pp48-49]. As it relates to latent fingerprint examination, the AAAS report concluded "[s]cientific research has convincingly established that the ridge patterns on human fingers vary greatly among individuals [and t]hese findings provide a scientific basis for the use of fingerprint comparison to distinguish individuals" [129 p5]. However, the AAAS report also concluded "[w]hile latent print examiners may well be able to exclude the preponderance of the human population as possible sources of a latent print, there is no scientific basis for estimating the number of people who could not be excluded and, consequently, no scientific basis for determining when the pool of possible sources is limited to a single person" [129 p9].

In 2018, NIST began conducting technical merit evaluations (also referred to as scientific foundation reviews) on DNA mixture interpretation, digital evidence, bitemark analysis, and firearms examination, and in 2020, NIST published a general description of the approach that would be taken [25]. In June 2021, the draft report for DNA mixture interpretation was released [130] followed by the final reports for digital evidence in 2022 [131] and bitemarks in 2023 [132]. Among the two published final reports to date, the NIST reports concluded: (i) "that digital evidence examination rests on a firm foundation based in computer science ... [and t]he application of these computer science techniques to digital investigations is sound and only limited by the difficulties of keeping up with the complexity and rapid pace of change in [information technology] [131 pp52-53], and (ii) "[t]he data available does not support the accurate use of bitemark analysis to exclude or not exclude individuals as the source of a bitemark" [132 p24].[54]

In the years following the PCAST report, greater attention were also directed toward the calculation of error rates and how performance measures should be calculated—specifically relating to whether and how "inconclusive" responses ought to be treated when calculating error rates [5]. Although this issue of "inconclusive" responses in binary decision making is not new, discussion surrounding the use of "inconclusive" decisions in the context of forensic science began to manifest in earnest in 2019 when the argument was raised that there is a lack of transparency and accountability on the use of "inconclusive" decisions, and

[54] In May 2023, the American Academy of Forensic Sciences (AAFS) Forensic Odontology Section, American Board of Forensic Odontology (ABFO), and American Society of Forensic Odontology (ASFO) issued a joint response objecting to some of the findings in the NIST Bitemark report, "highlighting areas of agreement, concern, and missed opportunities" [285]. Legal commentators, however, agreed with the findings from the NIST report and pointed to the failures of bitemark analysis and its contributions to wrongful convictions, claiming: "After 300 years of using of bite mark evidence in criminal cases, its legitimacy has been irreparably undermined by the general scientific community ... [and t]he use of such evidence is scientifically and morally unacceptable in our criminal legal system" [286 p5].

recommendations were made for the forensic science community to establish criteria to determine whether and when “inconclusive” decisions are justifiable [133]. Shortly thereafter, there was further debate regarding how generalizable error rate calculations that include “inconclusive” decisions are to the disciplines overall and how “inconclusive” decisions and “participant drop-out” or “missing data” should be handled when assessing the validity and reliability of forensic science methods [134–141].

5.1.1. Considerations for Forensic Science Research and Standards Programs

Over the last two decades, there has been debate around the extent to which several forensic science methods are scientifically supported, and whether they satisfy various criteria and ought to be considered scientifically valid and reliable. The importance of demonstrating that forensic science methods are valid and reliable is not at issue; rather, *what* constitutes an acceptable and mutually agreeable approach to assessing validity and reliability, *who* is best positioned to conduct those assessments, and *how* such determinations ought to be made is less clear. This leads to three major considerations for forensic science research and standards programs (Box 5.1.1).

Box 5.1.1. Forensic science research and standards programs should:

(a) Provide explicit criteria on how scientific validity and reliability ought to be assessed and the basis for such determinations.

(b) Develop guidelines and tools for the forensic science community that describe how research that is intended to measure validity and reliability of forensic science methods should be designed and executed.

(c) Develop guidelines and tools that enable the forensic science community to assess the validity and reliability of a particular method *as applied* in the context of a particular case.

Criteria for assessing scientific validity and reliability should be broadly applicable across all commonly practiced forensic science disciplines and methods and include relevant context to ensure they can be tailored to specific disciplines and methods.[55] Further, these criteria should allow for evaluating scientific literature and distinguishing between research that is optimally designed and executed and research that is informative but less rigorous, applying care not to be overly dismissive or overly inclusive[56] but clear on how to evaluate the weight

[55] A key element of these criteria should be that the materials and/or data upon which studies of validity and reliability are based are accessible for independent review or testing to assess reproducibility of the results. For other considerations to include in these criteria, see [13 pp52-53, 287, 288].

[56] E.g., considerations might include categorizing research based on “levels of evidence,” such as those used within clinical medicine [289].

and merit of published materials.[57] Finally, criteria for assessing validity and reliability and the research underlying such determinations should be specific enough such that they can be applied in a repeatable and reproducible manner, irrespective of who actually conducts such assessments.

Further, it is important for guidelines and tools that describe how research should be designed and executed (for purposes of measuring foundational validity and reliability) to align with and incorporate the criteria set forth for assessing validity and reliability and address validation of novel methods (i.e., developmental validation or foundational validation), verification of previously validated methods (i.e., internal validation), and on-going monitoring of performance of previously implemented methods within the framework of a quality management system.[58] These guidelines should provide clarity on what performance characteristics are necessary and what statistical methods are appropriate,[59] among other parameters relevant to the specific context (e.g., sample selection, minimum sample size, acceptance criteria).[60]

While the validity and reliability of forensic science methods can be demonstrated based on an existing body of research as it relates to the "foundational validity," such determinations cannot necessarily be generalized to the context of a specific case [13]. Thus, the forensic science community will also need to demonstrate that a method that has been determined to be "foundationally valid" was applied to the circumstances of a specific case in a valid and reliable manner (e.g., "validity as applied") [13]. Measures of validity and reliability of a particular method *as applied* in the context of a particular case should be based on clearly defined, standardized methods and using test specimens most representative of the quality and other relevant attributes of the evidentiary material(s) to which the method was applied.[61]

5.2. Research and Development

A common challenge shared across the forensic science community is to "enable the reliable inference of knowledge from uncertain information" [6 p111]. This requires that the forensic science community possess both (i) the *capability* to make those inferences with the highest degree of accuracy and precision possible and (ii) the *capacity* to do so at scale. In the years leading up to the 2009 NRC report, the capability to make valid and reliable inferences was largely assumed [5, 7]. However, when the NRC was instructed by Congress to "assess the present and future resource needs of the forensic science community" and "make recommendations for maximizing the use of forensic technologies and techniques" [6 p5], they challenged that assumption, claiming:

---

[57] E.g., see generally [287]. See also [288 p3] for general "tenets of literature review [that] should be considered in a critical review process that evaluates the merit of an individual article."
[58] E.g., guidelines and tools might include process mapping to facilitate their consistent application.
[59] Including appropriate statistical treatment of issues concerning "inconclusive" decisions and "participant drop-out" or "missing data."
[60] E.g., see [290–292].
[61] E.g., guidelines and tools might include the use of process maps to demonstrate alignment with standard practices and means of representing relevant performance metrics by bracketing the factor space around the conditions represented in the context of the case at hand.

> *The simple reality is that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. This is a serious problem. Although research has been done in some disciplines, there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.* [6 p8]

Following the NRC report, efforts to strengthen the scientific foundations of forensic science became a key priority for the Obama Administration and other entities (e.g., a detailed summary of key actions and activities taken by the Administration between 2009 and 2014 are provided in [142, 143]). In the subsections that follow, a brief overview of these activities and events is presented, followed by a summary of key challenges and needs identified and recommendations for future research and development.[62]

5.2.1. Significant Activities and Events

The timeline below (Figure 1) highlights major activities and events of the federal government and other entities following the 2009 NRC report.[63]

[62] This landscape analysis provides an overview of the specific challenges, needs, and recommendations identified over the years following the 2009 NRC report. This landscape analysis does not, however, address the specific priority of the challenges and needs identified or discuss the enduring relevance of those challenges and needs at present time. Further, this landscape analysis does not attempt to map specific research and development that has been accomplished to date to those specific challenges and needs. Future work will need to be done to consider the relevance and priority of the challenges and needs identified given the research and development that has been accomplished to date.

[63] The National Academies Press provides an additional timeline of major events between 2009 and 2019 including some items not discussed in this landscape [21].

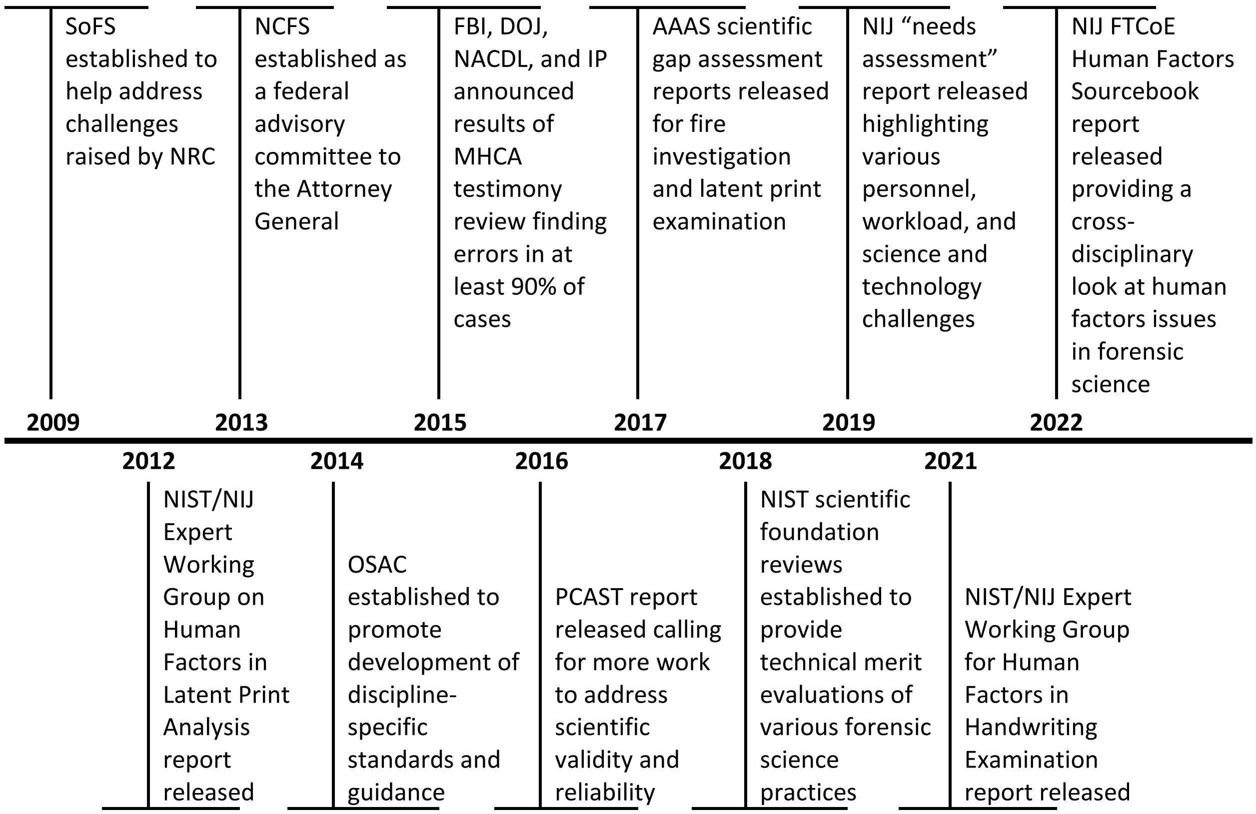

Figure 1. Timeline summarizing major activities and events taken by the federal government and other entities to strengthen forensic science in the United States since 2009.

In July 2009, the White House Office of Science and Technology Policy (OSTP) Committee on Science (CoS) chartered the Subcommittee on Forensic Science (SoFS) to address a wide range of issues relating to how the federal government might address some of the challenges raised by the NRC [144]. In areas relating to science and technology needs, the SoFS made two important contributions: (i) they coordinated with the various Scientific Working Groups (SWGs)[64] to assemble bibliographies of scientific literature purported to support foundational claims for ten different disciplines,[65] and (ii) they issued a report charting the path forward for achieving interoperability of Automated Fingerprint Identification Systems (AFIS) in the United States,[66] outlining several technical and administrative barriers and identifying a series of actions that could be taken to implement the standards needed to achieve interoperability, develop an overarching national connectivity strategy and infrastructure, and support state and local agencies in building connections across jurisdictions [145].

[64] In the late 1980s and 1990s, the FBI Laboratory established Technical Working Groups (TWGs) to issue quality assurance guidelines for various forensic science disciplines. In 1998, references to TWGs were changed to Scientific Working Groups (SWGs). When the OSAC was established in 2014, many SWGs disbanded, to be replaced by OSAC [143].

[65] For an overview of the bibliographies, see [143]. The bibliographies were not evaluated or prioritized on merit as it relates to the extent to which they actually address the foundational claims in the disciplines. They were, however, reviewed by the National Commission on Forensic Science and included in the assessments provided by the PCAST and AAAS reports.

[66] The lack of AFIS interoperability was first raised by the NRC in 2009 as a key technological challenge preventing the forensic science and law enforcement communities from leveraging AFIS technologies to their fullest potential [6 pp269-278].

In 2012, the Expert Working Group on Human Factors in Latent Print Analysis released their report entitled *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* [146]. Convened by NIJ and NIST in December 2008, the Expert Working Group was charged with conducting a scientific assessment of the effects of human factors on forensic latent print analysis and providing guidance to policy makers and government agencies in promoting a national agenda for error reduction in latent print analysis [146]. In their report, the Expert Working Group provided a detailed description of the methods and practices employed by the latent print discipline and identified several challenges and needs facing the discipline as it relates to human factors issues[67]—many of which were represented as recommendations for future research and development [146].[68]

In 2013, the NCFS was established as a federal advisory committee to the United States Attorney General on forensic science policies and practices with the mission to "enhance the practice and improve the reliability of forensic science" [147] by "[s]trengthening the validity and reliability of the forensic sciences; … [i]dentifying and recommending scientific guidance and protocols for evidence seizure, testing, analysis, and reporting by forensic science laboratories and units; and … identifying and assessing other needs of the forensic science communities to strengthen their disciplines and meet the increasing demands generated by the criminal and civil justice systems at all levels of government" [148]. During its tenure, the NCFS adopted a total of 43 work products: 20 Recommendation documents and 23 Views documents[69] on a wide range of issues relating to forensic science research, policy, and practice [149]. In 2017, the NCFS charter ended and was not renewed; however, the NCFS noted that there was much work left to do [149 p6]. When reflecting on the work of the Commission and looking toward the future, the NCFS noted several scientific and technological challenges facing the forensic science community that should be taken on by other entities "working with a multiple stakeholder perspective" [149 p6].[70]

In 2014, OSAC, administered by NIST, was established to address the lack of discipline-specific standards guiding forensic science practices [23]. OSAC comprises more than 800 "volunteer members and affiliates who work in forensic laboratories and other institutions around the country and have expertise in 22 forensic disciplines, as well as scientific research, measurement science, statistics, law and policy" [23]. In addition to the charge of "facilitating the development and promoting the use of high-quality, technically sound standards [that] … define minimum requirements, best practices, standard protocols and other guidance to help ensure that the results of forensic analysis are reliable and reproducible," OSAC can also identify and propose "research needs" to encourage further research and development to address science and technology challenges facing the various disciplines [23]. Since it was

[67] The study of human factors focuses on the interaction between humans and products, decisions, procedures, workspaces, and the overall environment encountered at work and in daily living [146 pvi].
[68] A summary of these recommendations identified by the Expert Working Group is provided in the subsection of this landscape assessment specific for friction ridge examination.
[69] Recommendation documents provided specific acts that the Attorney General could take to further the goals of the NCFS. Views documents represent formal views of the NCFS relating to particular topics.
[70] A summary of these issues identified by the NCFS is provided in the subsection of this landscape assessment that highlights cross-disciplinary challenges, needs, and recommendations.

established in 2014, OSAC has identified over 170 research needs relating to existing gaps and future capabilities affecting forensic science practice [150].[71]

In 2015, the outcome of a review of testimony statements made by FBI experts related to Microscopic Hair Comparison Analysis (MHCA) performed through 1999 was released [151]. The review revealed "[e]rrors in at least 90 percent of cases" [151]. The FBI began the review in 2012 in collaboration with the DOJ, National Association of Criminal Defense Lawyers (NACLD), and the Innocence Project (IP) following exonerations of a few individuals between 2009 and 2012 who were convicted in part on MHCA testimony [151, 152]. In 2018, a more comprehensive assessment was released identifying several systemic causal factors that led to the "examiners provid[ing] statements in reports and testimony exceeding the limits of the science that went unabated for decades through 1999" [152]. Ultimately, these findings led to increased attention on providing specific guidance and more robust monitoring for reports and testimony [152].

In 2016, following a tasking by the Administration in 2015, the PCAST released their report in which they recognized "[t]he Administration has taken important and much needed initial steps by creating mechanisms to discuss policy, develop best practices for practitioners of specific methods, and support scientific research[; however, a]t the same time, work to date has not addressed the 2009 NRC report's call to examine the fundamental scientific validity and reliability of many forensic methods used every day in courts" [13 p39]. In addition to their overarching conclusion that there is a need for on-going evaluation of the validity and reliability of various forensic methods, the PCAST highlighted the need to prioritize several other science and technology challenges—both overarching[72] and within specific disciplines[73]—with a particular emphasis on "transform[ing] subjective methods into objective methods" [13 p125]. The PCAST highlighted specific challenges and needs relating to each feature comparison discipline evaluated except for DNA analysis of non-complex mixtures (single-source or simple mixtures of two individuals) and bitemark analysis [13].[74]

In 2017, the AAAS released their two reports on fire investigation and latent fingerprint examination, respectively [128, 129]. The AAAS reports provide a description of key issues and questions the disciplines seek to address and an assessment of the foundational scientific literature supporting various aspects of the methods and practices within the disciplines.[75]

[71] A summary of overarching issues identified by the OSAC that transcend multiple disciplines is provided in the subsection of this landscape assessment that highlights cross-disciplinary challenges, needs, and recommendations. A summary of specific issues relating to each discipline represented by the OSAC is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.

[72] A summary of these issues identified by the PCAST is provided in the subsection of this landscape assessment that highlights cross-disciplinary challenges, needs, and recommendations.

[73] A summary of specific issues relating to each discipline evaluated by the PCAST is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.

[74] For bitemark analysis, the PCAST noted "PCAST considers the prospects of developing bitemark analysis into a scientifically valid method to be low [and] we advise against devoting significant resources to such efforts" [13 p87].

[75] A summary of specific issues relating to each discipline evaluated by the AAAS reports is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.

In 2018, resources were allocated to NIST to initiate technical merit evaluations (also referred to as scientific foundation reviews) [153]. In 2020, NIST published a general description of the approach that would be taken when conducting the foundation reviews [25]. In June 2021, the draft report for DNA mixture interpretation was released [130] followed by the final reports for digital evidence in 2022 [131] and bitemarks in 2023 [132]. The NIST foundation reviews provide a detailed description of the analytical methods and practices employed in the various disciplines, discussions of key issues and questions the disciplines seek to address through those methods and practices, and assessments of the scientific and technical foundations underlying various aspects of those methods and practices [25].[76]

In 2019, in response to a mandate of the Justice for All Reauthorization Act of 2016, NIJ published a needs assessment of forensic laboratories, captured between 2017 and 2018, entitled *Report to Congress: Needs Assessment of Forensic Laboratories and Medical Examiner/Coroner Offices* [42]. In addition to providing an examination of personnel, workload, backlog, and equipment needs for both public crime laboratories and medical examiner and coroner offices, the NIJ needs assessment also recognized new science and technology challenges facing forensic laboratories due to emerging demands for new types of evidence examinations (e.g., opioids and emerging drug threats, digital and multimedia evidence, sexual assault investigations), the need to understand and mitigate the impacts of human factors on evidence interpretation, and the need to increase efficiencies by scaling existing methods (e.g., backlog reduction) and developing tools and methods for field-based applications [42].[77]

In 2021, the Expert Working Group for Human Factors in Handwriting Examination released their report entitled *Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach* [154]. Convened by NIJ and NIST in 2015, this was the second in a series of expert groups examining human factors in forensic science (following the release of the 2012 expert working group report on latent print analysis [146]). The Expert Working Group was charged with conducting a scientific assessment of the effects of human factors on forensic handwriting examination with the goal of recommending strategies and approaches to improve its practice and reduce the likelihood of errors [154]. In their report, the Expert Working Group provided a detailed description of the methods and practices employed by the handwriting examination discipline and identified several challenges and needs facing the discipline as it relates to human factors issues [154].[78]

In 2022, following a tasking from NIJ, the Forensic Technology Center of Excellence (FTCoE) published a sourcebook for human factors in forensic science to provide a cross-disciplinary look at human factors issues and how they can affect forensic science examinations,

[76] A summary of specific issues relating to each discipline evaluated by NIST reports is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.
[77] A summary of overarching science and technology issues identified by NIJ that transcend multiple disciplines is provided in the subsection of this landscape assessment that highlights cross-disciplinary challenges, needs, and recommendations. Issues relating to particular disciplines is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.
[78] A summary of these recommendations identified by the Expert Working Group is provided in the subsection of this landscape assessment specific for handwriting examination.

interpretations, and communications of results [155]. The sourcebook is published in five separate parts as an open access special issue in *Forensic Science International: Synergy*, with each part a separate chapter covering the following topics:

- the need for research-based tools for personnel selection and assessment in the forensic sciences;
- the benefits of errors during training;
- challenges to reasoning in forensic science decisions;
- describing communication during a forensic investigation using the Pebbles on a Scale metaphor;
- stressors in forensic organizations: risks and solutions [155].

5.2.2. Cross-Disciplinary Challenges, Needs, and Recommendations

The 2009 NRC report provided a large-scale, systematic review of the broad-based science and technology challenges and needs facing the forensic science community [6].[79] The NRC found some existing methods and practices had limited empirical studies assessing the variability of features within the population, lacked quantifiable means for assessing the accuracy and reliability of the methods, and did not have an ability to express the level of confidence one may have in the results [6].[80]

In 2016, another large-scale, systematic review of the broad-based science and technology challenges and needs facing the forensic science community was produced by the PCAST [13]. In addition to identifying several challenges and needs relating to specific disciplines, the PCAST provided several "recommendations on specific actions that could be taken by the Federal Government … to ensure the scientific validity and reliability of forensic feature-comparison methods and promote their more rigorous use in the courtroom" [13 p123]. Specifically, the PCAST called for NIST to undertake:

- "on-going evaluations of the foundational validity of important methods … based on available, published empirical studies" [13 pp124, 128-129];
- "the creation and dissemination of large datasets [relating to latent fingerprint analysis, firearms analysis, and complex DNA mixtures in particular] to support the development and testing of methods by both companies and academic researchers" [13 pp125-126, 129];
- restructuring the OSAC to "ensure independent scientists and statisticians have a greater voice in the standards development process, a requirement for meaningful

[79] In addition to identifying several challenges and needs relating to specific disciplines, the NRC made several observations that were overarching across various disciplines—particularly the pattern evidence disciplines which rely on expert interpretation (e.g., fingerprints, shoeprints and tire tracks, toolmarks and firearms, handwriting, bitemarks, bloodstain patterns, fire debris, hair, and fiber analyses).
[80] E.g., see [6 pp127-182] for a description and summary assessment of each discipline evaluated. A summary of specific issues relating to each discipline evaluated by the NRC is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations. The exception to this is DNA analysis of single-source samples since the NRC considered it to be on sound scientific foundation and did not identify any specific challenges, needs, or recommendations.

scientific validity … [and m]ost importantly, [the development of] a formal committee—a Metrology Resource Committee—at the level of the other three Resource Committees (the Legal Resource Committee, the Human Factors Committee, and the Quality Infrastructure Committee) … [that is] composed of laboratory scientists and statisticians from outside the forensic science community and charged with reviewing each standard and guideline that is recommended for registry approval by the Science [sic] Area Committees before it is sent for final review [by] the Forensic Science Standards Board (FSSB)" [13 pp126, 129-130];
- establishing a forensic science research and development strategy that includes "coordination … across the relevant Federal agencies and laboratories to ensure that funding is directed to the highest priorities" [13 p128] and "address[es] plans and funding needs for (i) major expansion and strengthening of the academic research community working on forensic sciences, including substantially increased funding for both research and training; (ii) studies of foundational validity of forensic feature-comparison methods; (iii) improvement of current forensic methods, including converting subjective methods into objective methods, and development of new forensic methods; (iv) development of forensic feature databases, with adequate privacy protections, that can be used for research; (v) bridging the gap between research scientists and forensic practitioners; and (vi) oversight and regular review of forensic science research" [13 p130].

In 2017, just before the sunset of their charter, the NCFS issued their final report entitled *Reflecting Back—Looking Toward the Future*, which details several additional broad-based challenges and needs continuing to face the forensic science community [149]. Specifically, the NCFS highlighted the need to:

- "address digital forensics" [149 p7];[81]
- "provide guidance on evidence preservation and retention" [149 p8];
- "address source code accessibility and commercial transparency" [149 p8];
- "train forensic science users—law enforcement, lawyers, judges, and the public" on issues relating to scientific validity [149 p9];
- "establish research-based means of effectively and accurately communicating forensic science information with the judicial system and the public" [149 p10];[82]
- "focus on issues with communication and understanding between forensic analysts, investigators, lawyers, judges, juries, and the public" [149 p10].[83]

The 2019 NIJ needs assessment [42] made three broad recommendations, highlighting the need to:

[81] E.g., quality assurance, foundational validity and reliability, and evidence preservation.
[82] E.g., social science research on how particular terminology and/or statistical statements are presented and understood by the fact finder.
[83] E.g., how to address laboratory results that may need further clarification, such as the meaning of a finding of "inconclusive" and the use of presumptive testing as opposed to confirmatory testing in court cases.

- evaluate “existing policies and procedures to increase efficiency and quality of service provision”;
- develop “[h]ighly discriminating, accurate, reliable, cost-effective, and rapid methods for the identification, analysis, and interpretation of physical evidence”;
- establish cooperative relationships between forensic laboratories and research entities to conduct “[p]ractitioner-driven research … to develop evidence-based solutions that can be implemented into the workflows of forensic laboratories and medical examiner and coroner offices” [42 p135].

The 2022 NIJ FTCoE sourcebook on human factors in forensic science highlighted several challenges facing the community through a human factors lens [155]. Although much of the sourcebook discussed the applicability of existing psychology-related research to the forensic science domain and provided recommendations for policy and practice, some recommendations for additional research and development were noted [155]. Specifically, the sourcebook recommended the need to address the following topics:

- development of research-based tools for personnel selection and assessment in forensic science [156];
- understanding the applicability of psychology research findings relating to the concepts of “learning from errors” to forensic science specifically [157];
- understanding the effectiveness and learning benefits of witnessing feedback to errors of others [157];
- understanding the impact of fear of making errors on decision-making in forensic science, including fear of management consequences and fear of emotional consequences [157];
- understanding the effectiveness of “prevention focus” in training, such as whether awareness of riskier decision-making environments (e.g., when an examiner is fatigued) [157].

Finally, as of 2023, several forensic science entities identified and proposed a variety of needs and priorities to encourage further research and development to address various science and technology challenges facing the community.[84] These entities include: OSAC [150], NIJ [158], ASCLD [159], and the CFSO [160]. Collectively, OSAC, NIJ, ASCLD, and the CFSO have identified over 300 recommendations for research across a variety of disciplines and issues. Broadly speaking, those recommendations include studies relating to:

- development and validation of standardized methods, including statistical and computational methods to assess the strength of forensic evidence
- improvements to evidence collection and visualization tools

[84] Some entities have identified broad topics whereas others are more discipline-specific and some are even project-specific. Additionally, some entities have prioritized their research needs whereas others have not. A summary of overarching research needs that transcend multiple disciplines is provided in this subsection that highlights cross-disciplinary challenges, needs, and recommendations. A summary of discipline-specific research needs is provided in the relevant subsection of this landscape assessment that highlights the respective discipline-specific challenges, needs, and recommendations.

- evaluations of the accuracy and reliability of forensic analyses
- development of reference materials and databases
- better understanding of cognitive biases and other factors affecting forensic analyses [150, 158–160]

Several international entities have also formally identified and proposed research needs, including the National Institute of Forensic Science [161], the Netherlands Forensic Institute (NFI) [162], the European Network of Forensic Science Institutes (ENFSI) [163], and the International Forensic Strategic Alliance (IFSA) [164]. Research needs identified by these entities largely echo those identified by U.S. entities and relate to:

- development and validation of new forensic techniques
- evaluation and improvement of existing techniques
- development of tools and communication pathways to enhance the understanding of forensic evidence for non-scientists
- means to promote and enhance the sharing of data and knowledge across jurisdictions [161–164].

Examples include:

- the development of virtual and augmented reality processes for scene investigation and evidence presentation
- the application of artificial intelligence to support automated workflows
- the evaluation and implementation of new technologies to augment subjective interpretations and enhance analytical capabilities [161–164]

The 2021 research and innovation position statement from IFSA captures a consolidated view of the international and cross-organizational perspectives on research and development priorities across the forensic sciences [164], highlighting the need for:

- cross-disciplinary focus on studies to characterize the empirical foundations of current practices, including estimation of error rates, development of reference datasets and materials, and improved reporting mechanisms that reflect limitations, uncertainty, and error rates;
- understanding and characterization of transfer, persistence, and background abundance of impression, trace, and biological materials to support inferences relating to specific activities;
- understanding human factors impacts on forensic inference and decision making, including impact of various types of biases (beyond confirmation and contextual biases), developing new and innovative tools to communicate scientific information to non-scientists, and understanding the use and misuse of scientific data as it relates to investigative and legal decision making;

- data sharing across jurisdictions for interoperability and interchange of information, tools, and systems (e.g., biometrics);
- development of new tools for scene investigation, including methods relating to digital evidence connected at scenes, contactless tools for identifying and recording evidence at scenes, and development of virtual reality and augmented reality processes for scene investigation, evidence presentation and training;
- development and validation of new methods and technologies leveraging interconnectivity, real-time data, and cyber-physical systems, including accessing and analyzing data in remote cloud environments, developing virtual and augmented reality capabilities, developing wearable technologies as novel sensors, detectors, and recording devices, and exploring the usefulness of internet of things (IoT) and other developments in digital technology and data analytics as it applies within and across the justice systems;
- development and validation of applications in artificial intelligence (AI), including the automation of comparisons of forensic data to address subjectivity and to make interpretations objective and subsequently more accurate, such as interpretations of fingerprint, shoeprint, firearms and toolmarks comparisons, chemical profiling, blood pattern analysis, and fire pattern analysis, as well as the use of AI and data analytics to address data analyses related to the extraction, interpretation, and evaluation of information derived from digital devices;
- emerging biological and chemical evidence types, such as proteins or metabolites which can provide information about people and/or their environment for purposes of establishing new capabilities for human identification, interpretation of causes of death, and understanding the distribution and significance of peptides, lipids, and/or metabolites which may be measured in association with an alleged set of activities [164].

5.2.3. Discipline-Specific Challenges, Needs, and Recommendations

The subsections below represent discipline-specific challenges, needs, and recommendations made by various scientific committees since 2009. Formal and independent reviews of research relating to the scientific foundations have not been completed in all disciplines, and the evaluations that have been completed have often focused on specific issues within those disciplines (i.e., evaluations might not be comprehensive).

The absence of formal and independent evaluations of the scientific foundations for specific disciplines does not mean research does not exist or that the scientific foundations are lacking. Research has been conducted over the years across several forensic science disciplines (e.g., [165–168])[85] and guidance has been provided to the forensic science community (e.g., [169]); however, the quality and technical merit of that research and guidance as well as the

[85] In addition to the review papers compiled by the International Criminal Police Organization (INTERPOL), literature reviews and bibliographies have also been compiled by others within different disciplines (e.g., [293, 294]).

extent to which discipline-specific challenges and needs have been adequately addressed is less clear without any formal and independent evaluations.

This landscape analysis does not attempt to (i) evaluate the quality and technical merit of existing research or guidance, (ii) consider the extent to which research and guidance address the challenges and needs identified in previous evaluations, or (iii) provide an assessment of the scientific foundations for any disciplines. Instead, this landscape analysis summarizes the outcomes of formal and independent reviews that have been conducted and provides context around challenges, needs, and recommendations for research and development that have been identified.

The disciplines discussed below are generally distinguished based on the subcommittees represented by the OSAC [23]. Table 1 identifies the disciplines that have been subject to formal and independent reviews of the research supporting their scientific foundations since 2009 by the NRC [6], PCAST [13], AAAS [128, 129], or NIST [130–132, 146, 154] and for which research needs have been identified as a priority by OSAC [150], NIJ [158], ASCLD [159], or CFSO [160].

Table 1. List of forensic science disciplines that have been subject to formal reviews of the research supporting their scientific foundations since 2009 by the NRC, PCAST, AAAS, or NIST and for which research needs have been identified as a priority by OSAC, NIJ, ASCLD, or CFSO. The year indicates the latest year in which the discipline was reviewed by the entity or research needs were prioritized. The hyphen (-) indicates the discipline was not addressed. Footnotes provide additional context about the entry.

		Reviews of Scientific Foundations Completed					**Research Priorities Identified**			
Section	**Discipline**	**NRC**	**PCAST**	**AAAS**	**NIST / NIJ (Human Factors)**	**NIST (Foundation)**	**OSAC**	**NIJ**	**ASCLD**	**CFSO**
5.2.3.1.1	Human Forensic Biology	1996[86]	2016[87]	-	- [88]	2021	2023	2023	2022	2022
5.2.3.1.2	Wildlife Forensic Biology	-	-	-	-	-	2023	-	-	-
5.2.3.2	Chemistry (Controlled Substances and Toxicology)	2009	-	-	-	-	2023	2023	2022	2022

[86] Single source-DNA analysis was previously evaluated by the NRC in 1992 and 1996. In 2009, the NRC recognized the prior evaluation.
[87] PCAST recognized the prior evaluation of single-source DNA analysis and focused on DNA mixture analysis.
[88] NIST is currently working on a human factors report in human forensic DNA analysis, but it has not been published as of the date of this report.

		Reviews of Scientific Foundations Completed					Research Priorities Identified			
Section	Discipline	NRC	PCAST	AAAS	NIST / NIJ (Human Factors)	NIST (Foundation)	OSAC	NIJ	ASCLD	CFSO
5.2.3.3	Trace Materials	-	-	-	-	-	2023	2023	2022	2022
5.2.3.3.1	Analysis of Hair Evidence	2009	- [89]	-	-	-	2023	2023	-	-
5.2.3.3.2	Analysis of Fiber Evidence	2009	-	-	-	-	2023	2023	-	-
5.2.3.3.3	Analysis of Paint and Coatings Evidence	2009	-	-	-	-	2023	-	-	-
5.2.3.3.4	Analysis of Ignitable Liquids / Explosives / Gunshot Residue	2009	-	2017	-	-	2023	2023	2022	2022
5.2.3.4.1	Friction Ridge Examination	2009	2016	2017	2012	-	2023	2023	2022	2022
5.2.3.4.2	Footwear Examination	2009	2016	-	-	- [90]	2023	2023	-	-
5.2.3.4.3	Firearms and Toolmarks Examination	2009	2016	-	-	- [91]	2023	2023	2022	2022
5.2.3.4.4	Forensic Document Examination	2009	-	-	2021	-	2023	2023	2022	2022

[89] The PCAST did not undertake a comprehensive review of the analysis of hair evidence; however, they did comment on supporting documentation provided by the DOJ.
[90] NIST is currently working on a foundation review report in footwear examination, but it has not been published as of the date of this report.
[91] NIST is currently working on a foundation review report in firearms and toolmarks examination, but it has not been published as of the date of this report.

		Reviews of Scientific Foundations Completed					Research Priorities Identified			
Section	**Discipline**	**NRC**	**PCAST**	**AAAS**	**NIST / NIJ (Human Factors)**	**NIST (Foundation)**	**OSAC**	**NIJ**	**ASCLD**	**CFSO**
5.2.3.4.5	Bloodstain Pattern Analysis	2009	-	-	-	-	2023	2023	-	-
5.2.3.5.1	Crime Scene Investigation and Reconstruction	-	-	-	-	-	2023	2023	-	-
5.2.3.5.2	Dogs and Sensors	-	-	-	-	-	2023	2023	-	-
5.2.3.5.3	Fire and Explosion Investigation	2009	-	2017	-	-	2023	2023	-	-
5.2.3.6.1	Forensic Anthropology	-	-	-	-	-	2023	2023	-	-
5.2.3.6.2	Forensic Nursing	-	-	-	-	-	2023	-	-	-
5.2.3.6.3	Forensic Odontology	-	-	-	-	-	2023	-	-	-
5.2.3.6.3.1	Bitemark Analysis	2009	2016	-	-	2023	-	-	-	-
5.2.3.6.4	Medico-Legal Death Investigation	2009	-	-	-	-	2023	-	-	2022
5.2.3.7.1	Digital Evidence	-	-	-	-	2022	2023	-	-	-

		Reviews of Scientific Foundations Completed					Research Priorities Identified			
Section	Discipline	NRC	PCAST	AAAS	NIST / NIJ (Human Factors)	NIST (Foundation)	OSAC	NIJ	ASCLD	CFSO
5.2.3.7.2	Video / Imaging Technology and Analysis	-	-	-	-	-	2023	-	-	-
5.2.3.7.3	Facial and Iris Identification	-	-	-	-	- [92]	2023	-	-	-
5.2.3.7.4	Speaker Recognition	-	-	-	-	-	2023	-	-	-

5.2.3.1. Biology / DNA

Topics covered in this section include human forensic biology and wildlife forensic biology.

5.2.3.1.1 Human Forensic Biology

In 2009, the NRC recognized the foundations of single-source nuclear DNA analysis, stating: “Among existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source” [6 p100]. This assertion comes after two prior NRC evaluations relating to nuclear DNA analysis in forensic science [9, 10] and therefore the NRC did not include DNA analysis methods in its 2009 evaluation.

In their 2016 report, the PCAST noted “probabilistic genotyping software programs clearly represent a major improvement over purely subjective interpretation[; h]owever, they still require careful scrutiny to determine (1) whether the methods are scientifically valid, including defining the limitations on their reliability (that is, the circumstances in which they may yield unreliable results) and (2) whether the software correctly implements the methods” [13 p79]. Further, to support these types of evaluations (and others), the PCAST identified the need for the “creation and dissemination (under appropriate data-use and data-privacy

[92] Although NIST did not conduct a formal and systematic foundation review for either facial identification or iris identification, NIST has conducted research relating to the accuracy of facial identification (e.g., [121]) and conducted a review of relevant literature relating to forensic iris [171].

restrictions) of large collections of hundreds of DNA profiles created from known mixtures—representing widely varying complexity with respect to (1) the number of contributors, (2) the relationships among contributors, (3) the absolute and relative amounts of materials, and (4) the state of preservation of materials—that can be used by independent groups to evaluate and compare the methods" [13 pp82-83].[93]

The 2021 draft NIST report focused primarily on methods for interpreting data from complex DNA mixtures [130].[94] Due to advances in DNA technology providing the ability to detect and analyze very small quantities of DNA, the NIST report identified two overarching issues relating to DNA interpretation that have manifested in recent years: reliability (see generally: [130 pp55-96]) and relevance (see generally: [130 pp97-142]). Ultimately, the NIST report emphasized that "DNA mixtures vary in complexity, and the more complex the sample, the greater the uncertainty surrounding interpretation" [130 p32].[95] Further, the NIST report noted that "[d]ifferent analysts and different laboratories will have different approaches to interpreting the same DNA mixture [and t]his introduces [additional] variability and uncertainty in DNA mixture interpretation" [130 p89].

While the NIST report recognized the foundational reliability of non-complex DNA samples, they noted "[w]ith current laboratory methods, it is impossible to physically separate the DNA within a complex mixture into its constituent parts" [130 p56]. Consequently, "[t]he overall reliability of DNA mixture measurement and interpretation is influenced by many things … [and t]o assess reliability of any system, the factors that impact that system's performance need to be studied and evaluated" [130 pp60, 62].[96] Further, the NIST report noted that "[t]he factor space[97] for DNA mixture interpretation is vast and increases significantly with more contributors[; therefore, i]t is … practically impossible to demonstrate reliability across the full extent of any factor space [and t]his is particularly true without an established and accepted criteria for reliability with complex mixtures involving multiple contributors containing low quantities of DNA template or where there is a high degree of allele overlap among contributors" [130 p82]. Thus, "[d]emonstrating reliability requires that … empirical data [be provided] that is accessible to users of the information for independent assessments of reliability" [130 p82] bracketed around the factors that are most relevant to the context of the specific case.[98] Following a review of the available scientific literature, the NIST report determined that "[c]urrently, there is not enough publicly available data to enable an external

[93] The PCAST noted the "PROVEDit Initiative (Project Research Opennes for Validation with Experimental Data) at Boston University has made available a resource of 25,000 profiles from DNA mixtures." The PCAST also claimed "NIST should play a leadership role in this process, by ensuring the creation and dissemination of materials and stimulating studies by independent groups through grants, contracts, and prizes; and by evaluating the results of these studies." [13 pp82-83].
[94] The NIST report defines DNA mixtures as "samples that contain comingled DNA from two or more contributors in which stochastic effects or allele sharing cause uncertainty in determining contributor genotypes … [and include factors such as: n]umber of contributors and degree of overlapping alleles, [l]ow-quantity DNA from one or more minor contributors, [or] … degradation or inhibition of the DNA sample" [130 p12].
[95] The NIST report continued, noting "[f]actors that contribute to complexity include the number of contributors, the quantity of DNA from each contributor, contributor mixture ratios, sample quality, and the degree of allele sharing" [130 p32].
[96] The term "system" refers to the "entire process—starting from sample acquisition … and ending with an interpretation of results and expressing the strength of evidence" [130 p57].
[97] The term "factor space" is used "to describe the totality of scenarios and associated variables (*factors*) that are considered likely to occur in actual casework" [130 p60].
[98] E.g., see [130 pp89, 95] for more details about a "bracketing approach" and "performance testing with case-similar data," respectively.

and independent assessment of the degree of reliability of DNA mixture interpretation practices, including the use of probabilistic genotyping software (PGS) systems" [130 p75]. The NIST report continued, "[t]o allow for external independent assessments of reliability going forward, … forensic laboratories [are encouraged] to make their underlying PGS validation data publicly available and to regularly participate in interlaboratory studies" [130 p75].

In addition to outlining the criteria for assessing reliability of complex DNA mixtures, the NIST report also recognized that "[i]n the end, the reliability of [results] produced by a PGS system means little if relevance of the DNA evidence has not been established first" [130 p96]. The NIST report noted that "the possibility of [DNA] transfer cannot be ignored when interpreting DNA evidence [because i]f it is ignored, DNA findings, when considered in isolation, have the potential to be misleading" [130 p129]. The reasons for this are threefold:

- "[i]t is possible to handle an item without transferring any detectable DNA to it";
- "[g]enetic material may have been deposited before or after the crime and therefore may not be relevant to it";
- "[d]etected DNA might not be present due to indirect (secondary or tertiary) transfer, whether by a person or an object" [130 pp129-130].

Consequently, the NIST report cautioned that "[h]ighly sensitive DNA methods increase the likelihood of detecting irrelevant [and/or contaminating] DNA[; therefore, w]hen assessing evidence that involves very small quantities of DNA, it is especially important to consider relevance" [130 p131]. The NIST report also provided the caveat that "[t]he fact that DNA transfers easily between objects does not negate the value of DNA evidence[; h]owever, the value of DNA evidence depends on the circumstances of the case" [130 p139]. That said, the NIST report also noted that "[t]here is a growing body of knowledge about DNA transfer and persistence, but significant knowledge gaps remain" [130 p140].[99]

In addition to discussing issues concerning reliability and relevance, the NIST report also explored "the potential and limitations of new technologies to assist with DNA mixture interpretation" [130 pp143-154]. Specifically, the NIST report summarized "the challenges that are fundamental to DNA mixtures" as follows:

> *From a measurement and interpretation standpoint, … with any [polymerase chain reaction (PCR)] system, there will be stochastic variation when small amounts of DNA are analyzed … [which] impact the recovery of alleles and genotypes from mixture samples and lead to uncertainty in assigning alleles to genotypes and genotypes to contributor profiles[; w]hen STR markers are examined, stutter products add noise to the system … [which] impact uncertainty when alleles from minor contributor(s) overlap with stutter peaks of alleles from major contributor(s)[; and] sharing of common alleles can*

[99] While the NIST report noted "specific knowledge gaps remain," recommendations for future research are not specified. Instead, reference was made to [295].

> *mask the presence of contributor alleles and affect the ability to estimate the number of contributors.* [130 p145]

With those challenges in mind, the NIST report identified the following two broad areas of possible improvement via new technologies: physical separation of cells and sequencing. Specifically, "[p]hysically separating cells from different contributors prior to DNA extraction and [short tandem repeat (STR)] typing can reduce the need for DNA mixture interpretation[; however, while it is an attractive concept, it] presents new challenges of working directly with cells prior to DNA extraction" [130 p146] and "[n]ext-generation sequencing (NGS) … [which, c]ompared to existing [capillary electrophoresis (CE)]-based methods, NGS provides an additional dimension and more detailed resolution of genetic information, which includes the sequence of targeted PCR amplicons and accompanying stutter products with STR alleles" [130 pp147-149]; however, NGS also has other benefits and limitations that would need to be further explored (e.g., sequencing forensic STR markers and alternate markers) [130 pp149-153].

As of 2023, recommendations from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- DNA collection, extraction, and analysis techniques
- software solutions for interpreting DNA profiles
- exploring new avenues for DNA analysis
- developing procedures for using genetic genealogy in crime labs
- studying the transfer and persistence of DNA in real-world scenarios
- improving methods for identifying and interpreting DNA mixtures
- optimizing DNA sequencing methods
- identifying biological evidence
- assessing limitations and variability of probabilistic genotyping software [150, 158–160]

5.2.3.1.2 Wildlife Forensic Biology

Specific challenges, needs, and recommendations for research and development relating to wildlife forensic biology have been raised within the practitioner community, specifically through the OSAC. As of 2023, those recommendations include studies relating to:

- developing and validating DNA panels for current species of forensic interest
- developing new technologies for identifying species and sex when DNA is not available
- developing methods for distinguishing captive-bred from wild individuals of a given species or taxonomy
- verifying the accuracy of DNA sequences in public databases for taxonomic identification

- developing species-specific mitochondrial primers for identifying contributors in comingled samples
- using probabilistic genotyping methods for STRs found in wildlife samples [150]

5.2.3.2. Chemistry (Controlled Substances and Toxicology)

In 2009, the NRC found that "[t]he chemical foundations for the analysis of controlled substances are sound, and there exists an adequate understanding of the uncertainties and potential errors" [6 p135]. The main issue raised by the NRC was "whether all of the possible combinations [of analytical methods] recommended by SWGDRUG[100] would be acceptable in a scientific sense, if one's goal were to identify and classify a completely unknown substance" [6 p136]. However, the NRC noted that "[t]his ambiguity would be a less significant issue if the reports presented in court contained sufficient detail about the methods of analysis" [6 p136]. The NRC did not, however, evaluate forensic toxicology methods.

In 2019, the NIJ needs assessment highlighted the "opioid crisis and emerging drug threats" as a special topic of concern that has created "soaring growth in U.S. forensic laboratory expenditures associated with the opioid crisis" [42 p85]. The NIJ needs assessment noted "[t]he opioid crisis has been specifically implicated in the dramatic rise in drug overdose deaths in the United States … [and] in 2017, approximately three-times the rate of drug overdose deaths [were] reported [compared to those] reported in 1999" [42 p86]. NIJ continued, noting "[d]ue to the opioid crisis and the emergence of fentanyl and other drug threats from novel psychoactive substances (NPS), forensic laboratories have seen tremendous increases in workloads, and autopsy totals are threatening the accreditation status of [medical examiner/coroner (ME/C)] offices; … [a major issue being that t]he chemical structures of NPS are similar to those of known controlled substances, and NPS are being designed to stay ahead of federal and international laws that restrict the distribution and sale of specific chemicals" [42 p86]. Further, NIJ noted "[a]t the same time, the emergence of other drug threats, such as synthetic cannabinoids, synthetic cathinones ('bath salts'), and stimulants such as methamphetamine and cocaine, continues to affect communities as demonstrated by regional drug trends and mortality rates" [42 pp86-87]. NIJ highlighted a challenge is that "[e]ach newly identified substance requires additional research, development and implementation of laboratory methods, testing protocols, and advanced technologies and equipment to ensure sufficient sensitivity and specificity to detect these emerging drugs in forensic casework" [42 pp87, 89-93]. As a result of recent trends in the drug landscape, specifically relating to the growth of opioids and emerging drug threats, NIJ recommended the following science and technology related issues be addressed:

- development, validation, and implementation of means to "address the technical challenges of analyzing synthetic analogues … [including] new laboratory methods and testing protocols and advanced technologies and equipment";

[100] SWGDRUG refers to the Scientific Working Group for the Analysis of Seized Drugs [296].

- provide "[a]ccess to certified reference materials to accurately identify fentanyl analogues and other novel psychoactive substances in both drug and toxicology evidence";
- facilitation of better "coordination among forensic laboratories and medical examiner/coroner offices and other public safety and public health stakeholders to advance information and data sharing efforts";
- "resources and laboratory instrumentation to implement standardized toxicology analysis and death investigations to advance interpretation and reporting";
- development, validation, and "implement[ation of] field detection equipment for developing actionable information and sharing timely data" relating to opioids and fentanyl related analogues as well as for "increasing trends for other controlled substances including methamphetamine, cocaine, and other stimulants";
- means to facilitate safe "[h]andling of unknown and potentially hazardous substances, … particularly for crime scene and death investigation teams … [as well as] laboratory personnel";
- development, validation, and "implementation of new testing strategies … [for] marijuana and hemp" related cases and related technical challenges associated with their analyses;
- development, validation, and implementation of "[t]esting methods … [for] THC (tetrahydrocannabinol) from a wide variety of plant-based materials, edibles, and extracts and toxicology samples from driving while impaired (DWI) cases" [42 pp85-86].

Some of the recommendations highlighted by NIJ echo those set forth in the 2017 report published by the President's Commission on Combatting Drug Addiction and the Opioid Crisis [170]. For example, two specific yet far-reaching recommendations made by the Commission that were reinforced by NIJ include:

- "a federal effort to strengthen data collection activities enabling real-time surveillance of the opioid crisis at the national, state, local, and tribal levels";
- "develop[ment] and implement[ation of] standardized rigorous drug testing procedures, forensic methods, and use of appropriate toxicology instrumentation in the investigation of drug-related deaths" [170 p59].

As of 2023, recommendations directed toward controlled substances analysis from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- identifying potential isomers
- limitations of field techniques
- optimal derivatization techniques
- differentiation between THC-rich and cannabidiol (CBD)-rich cannabis plants
- alternative methods for identifying fentanyl-related substances
- DNA analysis of marijuana
- forensic laboratory process optimization

- long-term storage conditions of THC/marijuana material
- challenges in identifying new psychoactive substances (NPS)
- limitations of using only gas chromatography/mass spectrometry (GC/MS) analysis
- correlation of analytical findings to the legal status of seized drug analogs [150, 158–160]

Recommendations directed toward toxicology analysis from OSAC, NIJ, and ASCLD include studies relating to:

- improved methods for to the collection and analysis of toxicology samples
- use of data analytics in toxicology analysis
- better methods for the analysis of cannabinoids and other drugs of abuse
- development of road-side devices to test for marijuana use
- better understanding of the stability of drugs in different matrices
- development of reference materials to support further research, development, calibration and validation
- emerging drugs and their effects [150, 158, 159]

5.2.3.3. Trace Materials

Trace materials is a broad class of subdisciplines with varying challenges. Topics covered in this section include analysis of hairs, fibers, paint and coatings, ignitable liquids, explosives, and gunshot residue.

As of 2023, recommendations for research and development identified by OSAC, NIJ, ASCLD, and the CFSO include:

- discrimination and interpretation studies on different trace materials (including glass, soils, tapes, lubricants, hairs, fibers, and paints)
- evaluations of transfer and persistence for various trace materials
- environmental factors and their effects on trace materials
- identification and characterization of nanomaterials
- construction of new databases containing properties of manufactured materials [150, 158–160]

5.2.3.3.1 Analysis of Hair Evidence

In 2009, the NRC noted "[n]o scientifically accepted statistics exist about the frequency with which particular characteristics of hair are distributed in the population … [and t]here appear to be no uniform standards on the number of features on which hairs must agree before an examiner may declare a 'match'[; rather, t]he categorization of hair features depends heavily on examiner proficiency and practical experience" [6 p160]. The NRC continued, claiming "[t]he committee found no scientific support for the use of hair comparisons for individualization in the absence of nuclear DNA … [and while m]icroscopy and [mitochondrial DNA (mtDNA)] analysis can be used in tandem and may add to one another's value for classifying a common source, … no studies have been performed specifically to quantify the reliability of their joint use" [6 p161].

In 2016, the PCAST noted they had "not undertaken a comprehensive review of the discipline"; however, they "undertook a review of the supporting document [provided by the DOJ relating to testimony guidance] … that included supporting documents addressing the validity and reliability of the discipline … in order to shed further light on the standards for conducting a scientific evaluation of a forensic feature-comparison discipline" [13 p118]. Although the PCAST did not provide additional specific recommendations for future research, they express concerns with the research that is relied upon to base the claims of scientific validity and reliability [13 pp118-122]. Specifically, the "PCAST finds that, based on [researchers'] methodology and results, the papers described in the DOJ supporting document [relating to testimony guidance and claims of scientific validity and reliability] do not provide a scientific basis for concluding that microscopic hair examination is a valid and reliable process" [13 p120].[101]

As of 2023, recommendations for research and development identified by OSAC and NIJ include:

- evaluation of hair traits, their reliability and variability
- interdisciplinary studies on human hair morphology in combination with proteomics and DNA [150, 158]

5.2.3.3.2 Analysis of Fiber Evidence

In 2009, the NRC noted that "guidelines, but no set standards, for the number and quality of characteristics that must correspond in order to conclude that two fibers came from the same manufacturing batch … [and t]here have been no studies of fibers (e.g., the variability of their characteristics during and after manufacturing) on which to base such a threshold" [6 pp162-163]. Further, the NRC noted "there have been no studies to inform judgments about whether environmentally related changes discerned in particular fibers are distinctive enough to reliably individualize their source, and there have been no studies that characterize either reliability or error rates in the procedures … [t]hus, a 'match' means only that the fibers could have come from the same type of garment, carpet, or furniture; it can provide only class

[101] The PCAST concluded their discussion of hair analysis with the following statement: "Our brief review is intended to illustrate potential pitfalls in evaluations of the foundational validity and reliability of a method. PCAST is mindful of the constraints that DOJ faces in undertaking scientific evaluations of the validity and reliability of forensic methods, because critical evaluations by DOJ might be taken as admissions that could be used to challenge past convictions or current prosecutions. These issues highlight why it is important for evaluations of scientific validity and reliability to be carried out by a science-based agency that is not itself involved in the application of forensic science within the legal system" [13 pp121-122].

evidence" [6 p163]. Ultimately, the NRC claimed "[b]ecause the analysis of fibers is made largely through well-characterized methods of chemistry, it would be possible in principle to develop an understanding of uncertainties associated with [fiber] analyses[; h]owever, to date [as of 2009], that has not been done" [6 p163].

As of 2023, recommendations for research and development identified by OSAC and NIJ include:

- environmental factors and their effects on fibers, fiber populations, fiber variability and discrimination
- reliability of analytical methods and interpretation [150, 158]

5.2.3.3.3 Analysis of Paint and Coatings Evidence

In 2009, the NRC noted "[a]s is the case with fiber evidence, analysis of paints and coatings is based on a solid foundation of chemistry to enable class identification[; however,] … the community has not defined precise criteria for determining whether two samples come from a common source class" [6 p170].

As of 2023, recommendations for research and development identified by OSAC include:

- environmental factors and their effects on paints and coatings
- discrimination and interpretation of non-automotive (i.e., architectural and maintenance) paints
- characterizations of automotive coatings [150]

5.2.3.3.4 Analysis of Ignitable Liquids / Explosives / Gunshot Residue

In 2009, the NRC noted "[t]he scientific foundations exist to support the analysis of explosions, because such analysis is based primarily on well-established chemistry[; however, a]s part of the laboratory work, an analyst often will try to reconstruct the bomb, which introduces procedural complications, but not scientific ones" [6 pp172-173].

In 2017, as it relates to fire investigation/analysis of ignitable liquids (fire debris analysis), the AAAS report recommended:

- research and development of "[e]nhanced field tools … to optimize sample identification and sample collection at a fire scene … that allow for rapid feedback to investigators, and more sensitive and specific electronic 'noses' that can also detect the broad spectrum of potential [ignitable liquid residues (ILRs)]";

- "[a]dditional research [be conducted] to determine the performance of … [American Society for Testing and Materials (ASTM)] methods under various real-world case scenarios";
- interlaboratory studies be conducted to determine "[e]rror rates in fire debris analysis … [and to enable] a more quantitative assessment of the extent of false positive or false negative determinations of ILRs";
- "studies [be conducted relating to] differentiation of intentionally added ILRs from pyrolysis/combustion products and from products innocently present in materials at the fire scene";
- improvements to "classification of ILRs … to accommodate new products on the market (e.g., more environmentally friendly fuels such as biodiesel and plant-based lamp oils)";
- "[e]xperiements [be conducted] that explore the effects of weathering on different types of ILRs";
- studies be conducted relating to "[t]he impacts of potential microbial degradation on fire debris … [with respect to] substrates and situations [other than soil] in which these effects may be encountered";
- research be conducted relating to "the use of more sensitive methods … when examining ILRs … [and the impact to] the frequency of false positives and false negatives given some determination (absolute concentration) criteria, … [including r]esearch to determine the background levels of ILRs present in substrates encountered in a wide variety of settings" [128 pp10-11].

As of 2023, recommendations from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- the concentration of extracts containing triacetone triperoxide (TATP)
- evaluation and comparison of different extraction methodologies for ignitable liquids
- better means for identifying post-blast residue of liquid explosives
- evaluation of preservation and storage conditions for fire debris samples
- prevalence of characteristic and consistent GSR particles
- feasibility of organic GSR analysis
- persistence and transfer studies for GSR
- development of characterized GSR reference stubs
- new/novel methods for specific identification of shooters
- quantitative metrics for GSR analysis
- statistical framework for method optimization and development of GC-MS and liquid chromatography (LC)-MS systems [150, 158–160]

5.2.3.4. Pattern and Impression Evidence

Topics covered in this section include friction ridge, footwear, firearms and toolmarks, forensic documents, and bloodstain pattern analysis.

5.2.3.4.1 Friction Ridge Examination

In 2009, the NRC noted that "[b]ecause of the amount of detail available in friction ridges, it seems plausible that a careful comparison of two impressions can accurately discern whether or not they had a common source" [6 p142]. However, the NRC also claimed:

> *ACE-V[102] provides a broadly stated framework for conducting friction ridge analyses. However, this framework is not specific enough to qualify as a validated method for this type of analysis. ACE-V does not guard against bias; is too broad to ensure repeatability and transparency; and does not guarantee that two analysts following it will obtain the same results. For these reasons, merely following the steps of ACE-V does not imply that one is proceeding in a scientific manner or producing reliable results.* [6 p142]

The NRC continued, noting that although "[s]ome scientific evidence supports the presumption that friction ridge patterns are unique to each person and persist unchanged throughout a lifetime[,] … those conditions do not imply that anyone can reliably discern whether two friction ridge impressions were made by the same person … [or that] prints from two different people are always sufficiently different that they cannot be confused, or that two impressions made by the same finger will also be sufficiently similar to be discerned as coming from the same source" [6 pp143-144]. Further, the NRC noted that "impression[s] left by a given finger will differ every time, because of inevitable variations in pressure, which change the degree of contact between each part of the ridge structure and the impression medium … [and n]one of these variabilities—of features across a population of fingers or of repeated impressions left by the same finger—has been characterized, quantified, or compared" [6 p144]. Consequently, the NRC concluded "additional research is … needed into ridge flow and crease pattern distribution on the hands and feet[;] … more research is needed regarding the discriminating value of the various ridge formations and clusters of ridge formations … to provide examiners with a solid basis for the intuitive knowledge they have gained through experience and … lead to a good framework for future statistical models[;] … and [more] research on the various factors that affect the quality of latent prints (e.g., condition of the skin, residue, mechanics of touch) … [to] provide examiners with additional tools to support or refute distortion explanations … [and help examiners avoid criticism that they] can too easily explain a 'difference' as an 'acceptable distortion' in order to make an identification" [6 pp144-145].

Following the NRC report, in 2012, the Expert Working Group on Human Factors in Latent Print Analysis released their report entitled *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* [146]. The Expert Working Group recommended research and development to address the following topics related to latent print examination:

---

[102] ACE-V is the acronym that refers to the examination methodology consisting of Analysis, Comparison, Evaluation, and Verification.

- “methods to reduce the variation of feature selection and tools and technologies to help identify the most reliable [and useful] features” [146 p52];
- “understand[ing] the degree to which utility determinations are reliable … [and f]actors affecting utility determinations [such as] … context effects … [and] comparator effects … [as well as] culture norms and other human factors” [146 p54];
- “systematic studies pertaining to the reproducibility and discriminating strength of fingerprint features … [including] minutiae configurations … [and other] features such as creases, lines, and scars, which are useful to support the evaluation process when these features are present” [146 p63];
- “[s]tudies to measure the variability of distortion and the extreme limits of distortion[,] … [whether] an examiner’s working assumptions regarding the effects and degree of distortion have an empirical basis[,] … if, or in what circumstances, a misattribution of distortion may lead to an incorrect conclusion[, and] … the extent to which contextual information affects the interpretation of dissimilarities … [or] erroneous conclusions” [146 p66];
- “more knowledge about error rates, what affects them, and the extent to which they are correlated to the relative difficulty of comparisons[,] … what influences affect not only the ultimate decisions of examiners but also their decision-making thresholds … [and a]t what point does this impact on thresholds change outcomes[,] … [and] whether considering certain features makes examiners more vulnerable to inaccurate decisions” [146 p74];
- understanding “the relationship between risk and cost to examiners when making decisions[, f]or example, does the cultural default position of law enforcement and its relationship with forensic practitioners create an environment in which latent print examiners shun the inconclusiveness of evidence in favor of more definitive conclusions that are more conducive to current law enforcement expectations?” [146 p74];
- understanding “confidence associated with decision making in latent print examination … to see if confidence levels are consistent when making identifications as opposed to exclusions … [and] under what circumstances a qualified conclusion would be warranted” [146 p74];
- “[d]evelop[ment of] measures and metrics relevant to the analysis of latent prints; … [u]se [of] such metrics to assess the reproducibility, reliability, and validity of various interpretive stages of latent print analysis; … [and i]dentif[ication of] key factors related to variations in performance of latent print examiners during the interpretation process” [146 p76];
- “improve[ment of] automated fingerprint identification systems … [by e]xpanding the algorithms used to match prints to account for the fact that the diagnostic value of minutiae depends on the region in which they are located; [m]aking fingerprint and palm print databases interoperable among local, state, and federal automated identification systems; and [i]ncreasing compatibility between automated identification systems and other latent print software tools, including digital enhancement programs, probability calculation programs, and automated quality assessment programs” [146 p79];

- "creat[ion of] large, anonymous databases of exemplars and latent prints … [t]o facilitate the validation of probabilistic models and other statistical research" [146 p85];
- enablement of "automation of the initial quality assessment step in latent print analysis" [146 p88];
- "determin[ation of] the most appropriate tests of visual function for friction ridge examiners" [146 p149] and "what educational and cognitive abilities should be prerequisites for training a latent print examiner" [146 p166].

In 2016, the PCAST recognized the significant progress that had been made since the 2009 NRC report through black-box and white-box studies and highlighted the "nascent efforts to begin to move the field from a purely subjective method toward an objective method—although there is still a considerable way to go to achieve this important goal" [13 p88]. Specifically, the PCAST concluded:

> *[T]he FBI Laboratory black-box study has significantly advanced the field. There is a need for on-going studies of the reliability of latent print analysis, building on its study design. Studies should ideally estimate error rates for latent prints of varying "quality" levels, using well defined measures (ideally, objective measures implemented by automated software) [and these] studies should be designed and conducted in conjunction with third parties with no stake in the outcome [which is an] important feature [that] was not present in the FBI study.* [13 pp97, 102]

Further, the PCAST stated "[t]he most important resource to propel the development of objective methods would be the creation of huge databases containing known prints, each with many corresponding 'simulated' latent prints of varying qualities and completeness, which would be made available to scientifically-trained researchers in academia and industry [and t]he simulated latent prints could be created by 'morphing' the known prints, based on transformations derived from collections of actual latent print-record print pairs" [13 pp102-103].

In 2017, an AAAS report expanded on the prior evaluations completed by the Expert Working Group and the PCAST and recommended:

- "research [be conducted] on possible quantitative methods for estimating the probative value or weight of fingerprint evidence" [129 p5] and for "[d]eveloping better quantitative measures of the quality of latent prints … [for] assessing and improving AFIS as well as for evaluating the performance of human examiners" [129 p7];
- "research [be conducted] on the performance of latent fingerprint examiners under typical laboratory conditions … by introducing known-source prints into the flow of casework in a manner that makes test samples indistinguishable from casework samples" [129 pp5, 9];

- "research [be conducted] … on how accurately latent print examiners can assess intra-finger variability—that is, the degree to which prints may be changed due to distortion[—and] … ways to reduce the probability of false exclusions" [129 p6];
- evaluations of "the performance of commercial AFIS systems, particularly their performance in identifying latent prints" [129 p7];
- engagements be facilitated between law enforcement agencies and AFIS vendors "to better assure interoperability of AFIS systems and avoid compatibility problems" between different systems [129 p7];
- "creation of research test sets—i.e., latent print specimens of known source that can be used for testing examiner performance" [129 p9];
- "research … on how lay people, such as police officers, lawyers, judges, and jurors evaluate and respond to fingerprint evidence … [to] evaluate how best to present fingerprint evidence in reports and expert testimony … [and to] help ensure that the statements made in reports and testimony will be interpreted in the intended manner" [129 p12].

As of 2023, recommendations from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- accuracy and reliability
- examiner consistency
- sources and impacts of bias
- impacts to performance as a function of evidence complexity
- development and validation of standardized methods
- practical statistical approaches for interpretation
- evaluation of review and verification processes
- improved evidence collection and visualization tools
- understanding cognitive processes involved in pattern recognition
- development of tools to quantitatively assess aptitude of candidates [150, 158–160]

5.2.3.4.2 Footwear Examination

In 2009, the NRC noted that "[t]he scientific basis for the evaluation of impression evidence is that mass-produced items (e.g., shoes, tires) pick up features of wear that, over time, individualize them[; h]owever, because these features continue to change as they are worn, elapsed time after a crime can undercut the forensic scientist's certainty" [6 p149]. While the NRC recognized "[a]t the least, class characteristics can be identified, and with sufficiently distinctive patterns of wear, one might hope for specific individualization," the NRC also noted "there is no consensus regarding the number of individual characteristics needed to make a positive identification, and the committee is not aware of any data about the variability of class or individual characteristics or the validity or reliability of the method … [and w]ithout such population studies, it is impossible to assess the number of characteristics that must match in order to have any particular degree of confidence about the source of the impression" [6 p149]. The NRC continued, "it is difficult to avoid biases in experience-based judgments, especially the absence of a feedback mechanism to correct an erroneous judgment" [6 p149]. Ultimately, the NRC concluded that "critical questions that should be addressed include the persistence of individual characteristics, the rarity of certain characteristic types, and the appropriate statistical standards to apply to the significance of individual characteristics" [6 p150].

In 2016, as it relates to "identifying characteristics" [13 pp114-115],[103] the PCAST noted "there is little research on which to build with respect to conclusions that seek to associate a shoeprint with a particular shoe (identification conclusions) … [and n]ew approaches will be needed to develop paradigms" [13 p117]. Specifically, the PCAST suggested there is a need for the creation of "impressions generated from … footwear [to] provide an initial dataset for (1) a pilot black-box study and (2) a pilot database of feature frequencies" [13 p117]. Further, although the PCAST noted that determinations of class characteristics are not "inherently a challenging measurement problem" [13 pp114-115], and therefore is not a topic addressed in their report, nevertheless, they recommend "evaluations should be undertaken concerning the accuracy and reliability of determinations about class characteristics" [13 p117].

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- assessment of examiner reliability
- algorithms for automated searching of patterns within a database
- understanding the relationship between manufacturing techniques and features used for comparisons
- validation of interpretation scales
- determining the size of the smallest detail required for tire and shoe comparisons
- development of improved casting materials
- understanding the variability of dimensional characteristics during replication of impressions
- determining relevant populations for interpretation of class associations
- evaluating the probative value of general wear during examinations [150, 158]

5.2.3.4.3 Firearms and Toolmarks Examination

In 2009, the NRC noted "[t]oolmark and firearm analysis suffers from the same limitations discussed … for [footwear and tire] impression evidence[; b]ecause not enough is

[103] The PCAST noted their report "do[es] not address the question of whether examiners can reliably determine class characteristics—for example, whether a particular shoeprint was made by a size 12 shoe of a particular make. While it is important that that [sic] studies be undertaken to estimate the reliability of footwear analysis aimed at determining class characteristics, PCAST chose not to focus on this aspect of footwear examination because it is not inherently a challenging measurement problem to determine class characteristics, to estimate the frequency of shoes having a particular class characteristics, or (for jurors) to understand the nature of the features in question." [13 pp114-115].

known about the variabilities among individual tools and guns, [the committee] is not able to specify how many points of similarity are necessary for a given level of confidence in the result" [6 p154]. The NRC continued, noting "[s]ufficient studies have not been done to understand the reliability and repeatability of the methods" [6 p154]. The NRC also indicated that "[a] fundamental problem with toolmark and firearms analysis is the lack of a precisely defined process … [and t]he meaning of 'exceeds the best agreement' and 'consistent with' are not specified, and the examiner is expected to draw on his or her own experience" [6 p155]. The NRC claimed that "questions regarding variability, reliability, repeatability, or the number of correlations needed to achieve a given degree of confidence … is not addressed [in the best guidance available]." Ultimately, the NRC found that "[a]lthough some studies have been performed on the degree of similarity that can be found between marks made by different tools and the variability in marks made by an individual tool, the scientific knowledge base is fairly limited" [6 p155].

In 2016, the PCAST noted that "there is a need for additional black-box studies based on the study design of the Ames Laboratory black-box study … [and] the studies should be designed and conducted in conjunction with third parties with no stake in the outcome" [13 p113]. The PCAST also stated there is a need for "developing and testing image-analysis algorithms for comparing the similarity of tool marks on bullets" while recognizing some progress in this direction [13 p113]. However, the PCAST also noted "efforts are currently hampered by lack of access to realistically large and complex databases that can be used to continue development of these methods and validate initial proposals … [and called for] creating and dissemination [of] appropriately large datasets" [13 pp113-114].[104]

As of 2023, recommendations from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- assessing and improving the accuracy and reliability of firearm and toolmark examinations conducted by examiners
- studying the extent to which cognitive bias affects firearm and toolmark comparison outcomes
- consistency of examiners' distance determinations
- developing optimal methods and materials for the preservation, visualization, recovery, and comparison of tool marks in cartilage and bone [150, 158–160]

5.2.3.4.4 Forensic Document Examination

In 2009, the NRC noted "[r]ecent studies have increased our understanding of the individuality and consistency of handwriting and computer studies and suggest that there may be a scientific basis for handwriting comparison, at least in the absence of intentional

[104] The PCAST also claimed "NIST, in coordination with the FBI Laboratory, should play a leadership role in propelling this transformation by creating and disseminating appropriate large datasets … [and t]hese agencies should also provide grants and contracts to support work—and systematic processes to evaluate methods" [13 p114].

obfuscation or forgery" [6 pp166-167]. However, the NRC did not offer specific recommendations other than claiming "[t]he scientific basis for handwriting comparisons need[s] to be strengthened" [6 p166].

In 2021, the Expert Working Group for Human Factors in Handwriting Examination released their report entitled *Forensic Handwriting Examination and Human Factors: Improving the Practice through a Systems Approach* [154]. They recommended research and development to address the following topics related to forensic handwriting examination:

- understanding "[t]he impact of various sources of contextual information on forensic handwriting examinations, and [h]ow to balance the risks of bias and information loss with respect to all levels of contextual information[, including] … [w]hether some sources of contextual information are more biasing than others[;] … [t]he optimal order for [forensic document examiners] to perform their tasks and receive task-relevant information[;] … [t]he efficacy of [contextual information management] protocols[, such as] … whether or not redacting potentially biasing information during examinations is an effective way of increasing [forensic document examiner] objectivity and reducing bias[;] … and [a] cost/benefit analysis of the threshold at which information loss has a greater detrimental impact than risk of bias" [154 p52];
- "design[ing] and participat[ing] in 'black box' and 'white box' studies" [154 p63];
- "understand[ing] how to best convey [probabilistic] concepts to [forensic document examiners] and to consumers of handwriting examinations" [154 p70] as well as "how the [forensic document examiner's] presentation of evidence in court impacts the judge and jury's comprehension of the forensic evidence to avoid potential misunderstandings or miscommunication" [154 p140];
- "conduct[ing] empirical studies … to characterize the extent of scientific support for those claims" made by the forensic document examiner community about the opinions they claim they can render, such as claims that forensic document examiners can accurately and reliably render opinions under various conditions of writing and sample sources, e.g., "when both materials are uppercase print; … lowercase cursive; … non-originals; and … simulation or disguise behavior" [154 pp71-72];
- understanding "baseline occurrences of particular features in a population … addressing[: o]ccurence of features by geographic area[, o]ccurence of combinations of features[, i]dentification of rarely occurring features[, and i]dentification of characteristics common among and specific to population subgroups" [154 pp72-73];
- "design[ing] and construct[ing] publicly available, large databases of representative handwriting features to facilitate research in and improve the accuracy of handwriting examination" [154 p73];
- "defin[ing] how complexity is measured and the level to which complexity is sufficient for meaningful comparisons for all types of writing, like hand printing, numerals, signatures, or foreign writing systems" [154 p73];
- "[d]eveloping methods of quantifying and measuring inter-writer and intra-writer variability … [that] include cross-cultural writing and longitudinal studies of changes in

writing across time and … writing characteristics that arise in the absence of formal instruction in cursive writing and penmanship" [154 p73];

- "[a]mount of writing required to reach a conclusion about the writership of the questioned writings … [that] include the degree of writing complexity required to establish the presence or absence of diagnostic features, the minimum quantity of writing needed to form reliable opinions, cross-cultural studies, and studies specifically addressing writing forms like numerals, signatures, initials, and hand-printed materials" [154 p73];
- "[c]omparability of types of writing … [that] include forms of writing like initials, signatures, hand printing, and foreign writing" [154 p73];
- "[r]elevant information (features) identified in writing samples and the extent of the consistencies in how such information is interpreted … [and that] address the extent to which information in the written materials has the potential to reliably indicate whether the writing is genuine or non-genuine (i.e., disguised, traced, or produced by some other method of simulation) and how consistently such information is used to establish the writership of a questioned writing" [154 p73];
- "develop[ing] and validat[ing] applicable, user-friendly automated systems" for handwriting examination [154 p80];
- understanding "about the assumptions and principles underlying the elements of forensic handwriting examinations" [154 p115];
- "conduct[ing] … studies to determine the optimal content and frequency of proficiency tests to properly evaluate forensic document examiners' ability to perform the range of tasks encountered in casework" [154 p147];
- "develop[ing] collaborative testing programs [such as interlaboratory studies] aimed at monitoring and providing performance improvement opportunities related to specific claims and sub-claims" made by forensic document examiners [154 p151].

As of 2023, recommendations from OSAC, NIJ, ASCLD, and the CFSO include studies relating to:

- development of a national database of handwriting
- validation of conclusion scales
- assessment of handwriting and hand printing complexity and comparability
- evaluation of automated handwriting identification systems
- quantitative assessment of intra- and inter-person handwriting variation
- the assessment of comparability of different forms of writing from individuals
- understanding of the kinematics of handwriting and digitally captured signatures
- creating reference collection databases of handwriting samples and typewriter and computer font styles for forensic document examination [150, 158–160]

5.2.3.4.5 Bloodstain Pattern Analysis

In 2009, the NRC noted "[s]cientific studies support some aspects of bloodstain pattern analysis[, such as] for example, … if the blood spattered quickly or slowly, but some experts extrapolate far beyond what can be supported" [6 p178]. The NRC continued, noting "[a]lthough the trajectories of bullets are linear, the damage that they cause in soft tissue and the complex patterns that fluids make when exiting wounds are highly variable[; thus, f]or such situations, many experiments must be conducted to determine what characteristics of a bloodstain pattern are caused by particular actions during a crime and to inform the interpretation of those causal links and their variabilities" [6 pp178-179]. Ultimately, the NRC claimed that "extra care must be given to the way in which analyses are presented in court [because t]he uncertainties associated with bloodstain pattern analysis are enormous" [6 pp178-179].

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- developing objective and validated methods to classify spatter patterns
- differentiating spatter from transfer
- studying droplet formation and trajectory
- understanding the interaction of blood with fabrics and textiles [150, 158]

5.2.3.5. Scene Examination

Topics covered in this section include crime scene investigation and reconstruction, dogs and sensors, and fire investigation.

5.2.3.5.1 Crime Scene Investigation and Reconstruction

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- culture and contextual bias in scene evidence collection, handling and processing
- required personal protective equipment (PPE) at crime scenes
- minimum staffing for crime scene response
- decontamination of crime scene equipment
- effects of on-scene fatigue on investigators
- laboratory techniques and technologies at the crime scene
- quality assurance framework for crime scene investigation
- use of technology for crime scene documentation
- development of novel, improved, or enhanced presumptive tests for evidence analysis and interpretation at the scene and in the lab [150, 158]

5.2.3.5.2 Dogs and Sensors

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- canine performance and training
- development of methods for monitoring levels of contamination of training aids
- development of reliable surrogate aids
- evaluation of dissipation of odorants
- identification of odorant chemicals present in and above targets
- integration of canine and instrumental detectors [150, 158]

5.2.3.5.3 Fire and Explosion Investigation

In 2009, the NRC noted “much more research is needed on the natural variability of burn patterns and damage characteristics and how they are affected by the presence of various accelerants” [6 p173]. The NRC continued, noting “[d]espite the paucity of research, some arson investigators continue to make determinations about whether or not a particular fire was set[; h]owever, … many of the rules of thumb that are typically assumed to indicate that an accelerant was used (e.g., ‘alligatoring’ of wood, specific char patterns) have been shown not to be true[; thus, e]xperiments should be designed to put arson investigation on a more solid scientific footing” [6 p173].

In 2017, as it relates to fire scene investigation, an AAAS report recommended:

- physical fire tests be run: “in both reduced and full scale, using multiple compartments and multiple openings, fully documenting the aftermath; with the burning of different materials under a range of realistic fire conditions; and by lighting fires in identically constructed compartments … [and] information, such as temperature at various layers of the room and radiant heat fluxes … [are] measured” to better understand a fire’s origin and cause;
- concurrent with physical fire tests, “the fire scenario being tested should also be simulated with a deterministic fire model to evaluate the accuracy of the model and to better understand uncertainties associated with the model … [so that] computer-based deterministic fire models can be continually refined to produce more accurate results, and over time may find an expanded role as a useful tool in actual investigations”;
- “[n]ew technologies, as well as additional training aids and research on new methods … be developed for measuring canine performance, … [and conduct c]omparative research assessing the effectiveness of technologically more innovative field tools against the effectiveness of canine use”;
- study “[t]he reliability of conclusions when fire investigators are presented with similar data of fire origin and cause … [to enable] the calculation of both error rates and the reliability of investigators’ conclusions [and understand] decision points that cause divergent findings among investigators”;

- study “[t]he effects of education, training, and certification on fire investigators’ ability to determine fire origin and cause” [128 pp6-7, 9].

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- understanding of the creation and obscuration of fire patterns due to ventilation effects
- evaluation of methods for origin and cause determination
- standardized procedures for collecting, preserving, and analyzing building system electronic data
- tools for fire investigators to evaluate the effects of fuel characteristics on the growth and spread of fires
- the repeatability and reproducibility of test measurements of large-scale structure fires
- characterization of electrical system response as a means to study fire progression
- evaluation of incident heat flux profiles
- dissemination of fire research laboratory data [150, 158]

5.2.3.6. Forensic Medicine

Topics covered in this section include forensic anthropology, nursing, odontology, and medico-legal death investigation.

5.2.3.6.1 Forensic Anthropology

As of 2023, recommendations from OSAC and NIJ include studies relating to:

- controlled experimentation of bone trauma and bone healing rates
- stable isotope analysis for geospatial identification
- development of statistical models for personal identification
- enhancement of unidentified decedent systems
- difficulty in identifying geographical origin of remains [150, 158]

5.2.3.6.2 Forensic Nursing

As of 2023, recommendations from OSAC include studies relating to:

- optimal use of alternate light sources (ALS) to maximize findings during a forensic examination
- indicators for computer tomography angiography following non-fatal strangulation
- optimal oral areas to swab for DNA evidence collection

- optimization of DNA evidence collection following oral sexual assault and activities that inhibit DNA detection [150]

5.2.3.6.3 Forensic Odontology

Over the years, much of the attention relating to forensic odontology have been directed toward bitemark analysis specifically. As of 2023, recommendations from OSAC relating to forensic odontology generally include studies relating to:

- developing 3D databases of human dentitions
- improving dental age assessment methods
- phenotyping of tooth shape and color
- assessing the reliability of bitemark analysis methodology [150]

Assessments of bitemark analysis have been undertaken by several scientific entities over the years with largely consistent results; these are described in the section that follows.

5.2.3.6.3.1 Bitemark Analysis

In 2009, the NRC noted “[a]lthough the methods of collection of bite mark evidence are relatively noncontroversial, there is considerable dispute about the value and reliability of the collected data for interpretation … [such as] the accuracy of human skin as a reliable registration material for bite marks, the uniqueness of human dentition, the techniques used for analysis, and the role of examiner bias” [6 p176]. The NRC continued, noting that “[a]lthough the majority of forensic odontologists are satisfied that bite marks can demonstrate sufficient detail for identification, no scientific studies support this assessment, and no large population studies have been conducted[; furthermore, i]n numerous instances, experts diverge widely in their evaluations of the same bite mark evidence, which has led to questioning of the value and scientific objectivity of such evidence” [6 p176]. Overall, the NRC found “no evidence of an existing scientific basis for identifying an individual to the exclusion of all others [and] … research is warranted in order to identify the circumstances within which the methods of forensic odontology can provide probative value” [6 p176].

In 2016, the PCAST noted “[f]ew empirical studies have been undertaken to study the ability of examiners to accurately identify the source of a bitemark … [and a]mong those studies that have been undertaken, the observed false positive rates were so high that the method is clearly scientifically unreliable at present” [13 p87]. The PCAST continued, concluding “bitemark analysis does not meet the scientific standards for foundational validity, and is far from meeting such standards[; rather,] … [t]o the contrary, available scientific evidence strongly suggests that examiners cannot consistently agree on whether an injury is a human bitemark and cannot identify the source of bitemark with reasonable accuracy” [13 p87]. Ultimately, the PCAST considered “the prospects of developing bitemark analysis into a scientifically valid method to be low" and therefore advised "against devoting significant resources to such efforts" [13 p87].

Finally, in the 2023 foundation review, the NIST report focused on the overarching question: "Can bitemarks be accurately associated with teeth that left them?" [132 p4]. Ultimately, the NIST report reinforced the findings from the NRC and the PCAST and concluded: "[t]he data available does not support the accurate use of bitemark analysis to exclude or not exclude individuals as the source of a bitemark" [132 p24]. Although the NIST report identified several challenges facing the discipline and noted that "[c]alls have been made for empirical studies to assess the limitations of bitemark analysis for decades … and the need to address reliability concerns in bitemark methods" [132 p24], the NIST report also found that "these calls have largely gone unheeded" [132 p24]. Specifically, the NIST report claimed that bitemark "patterns are not accurately transferred to human skin consistently" [132 p4]; "[c]omparisons between bitemark patterns made on skin … have shown that there exists intra-individual variation in bitemark morphology on the human body such that bitemarks from the same biter may not appear consistent" [132 p18]; "bitemark examiners may not agree on the interpretation of a specific bitemark, including whether the injury is a bitemark, the features present, and the exclusion or non-exclusion of potential biters" [132 p23]; and "[r]epeated calls for additional data by critics and practitioners (since at least 1960) suggest insufficient support for the accurate use of bitemark analysis and a lack of consensus from the community on a way forward" [132 p24].

5.2.3.6.4 Medico-Legal Death Investigation

In 2009, the NRC noted "[c]urrently, little research is being conducted in the areas of death investigation and forensic pathology in the United States … [and i]ndividual [medical examiner / coroner (ME/C)] offices mainly utilize their databases for epidemiological retrospective reviews" [6 p261]. Further, the NRC noted "[o]ccasionally, a specific case may inspire 'litigation research' directed to the elucidation of a specific problem related to a case that is being litigated actively, but this does not replace broad and systematic research of a forensic issue … [and f]ew university pathology departments promote basic pathology research in forensic problems such as time of death, injury response and timing, or tissue response to poisoning" [6 p261]. The NRC also found that "[o]f the many impediments to academic research in forensic pathology in the United States, [two of] the most significant are the lack of understanding of forensic research challenges [and] the lack of perceived need and national goals" [6 p262]. The NRC continued "[g]iven the large numbers of autopsies performed in the United States in medical examiner offices, there is a great need for new knowledge that will filter down to the autopsy pathologist and for opportunities for practicing forensic pathologists to identify problems that need basic research" [6 pp262-263]. Finally, as it relates to science and technology issues, the NRC concluded that "[m]any ME/C systems do not utilize up-to-date technologies that would help in making accurate medical diagnoses [and b]asic and translational forensic pathology research are nearly nonexistent" [6 p265].

As of 2023, recommendations from OSAC and CFSO include studies relating to:

- the usefulness of autopsies in contentious medicolegal categories of death
- improving the accuracy of cause and manner of death certification
- understanding cognitive bias in disaster victim identification (DVI) operations
- improving the collection of reliable and well-documented toxicology samples
- developing new biometric capture techniques for decedent data
- determining the precise time since death
- innovative methods for trauma analysis
- detection of subtle injuries
- the use of advanced imaging technologies
- understanding the consequences of differing levels of post-mortem investigation
- assessment of current and appropriate caseloads of forensic pathologists and medicolegal death investigators [150, 160]

5.2.3.7. Digital / Multimedia

Topics covered in this section include digital evidence, video/imaging technology analysis, facial identification, and speaker recognition.

5.2.3.7.1 Digital Evidence

The need to address digital evidence (i.e., digital forensics) was recognized by the NCFS. Although their initial charter in 2013 specifically excluded digital evidence from their scope of review, "[t]he charter was later amended in 2015 to allow the Commission to consider digital forensics" [149 p7]. As the NCFS noted in their final report in 2017, "[w]hat became obvious right from the beginning is that the challenges facing digital forensics are in some ways unique[; for example, t]his area of practice is fast paced, often done in law enforcement settings by technicians rather than scientists, and has security issues that may not be of concern in other areas of forensics" [149 p7]. Further, the NCFS noted "[d]igital forensics, as a fairly new yet pervasive area of forensic science, can benefit from guidance of the Commission or similar group regarding quality assurance, foundational reliability, evidence preservation, and more[; simply put,] … [t]his entire area of forensic science needs more study and significant input from subject-matter experts" [149 p7].

In addition to the NCFS, the need to address digital forensics was also noted by NIJ in their 2019 needs assessment [42 pp98-110]. Specifically, NIJ noted "[t]he ubiquity of digital devices, including devices with digital camera and video capabilities, closed-circuit television, and officer body-worn cameras; the advent of device portability and instantaneous sharing via social media; internet-connected devices; and integrated devices such as drones and autonomous vehicles all exemplify the types and quantities of digital data that may provide investigative or forensic value" [42 p99]. NIJ continued, noting that digital evidence examinations "may [also] encounter technological challenges, such as encrypted files or devices, or require additional expertise to examine 'dark web' content and cryptocurrency transactions" [42 p99]. Although digital evidence examinations "continue to play a prominent role in everyday activities … [and i]t is now common for every case to have some form of [digital evidence,]" in 2009, the NRC treated digital evidence "as an emerging forensic science discipline" [42 pp99-100]. In the last 10 years, however, the field of digital forensics "has grown tremendously with the surge along with the growth in data storage capacities, data speeds, types of data, and the methods through which data can be shared and accessed" [42 p100]. As a result of the growth of demands for digital evidence, NIJ recommended the following science and technology related issues be addressed:[105]

- means to "continuously respond to new and emerging technologies and devices" [42 p99];
- "methods, tools, and training to triage [digital evidence examinations], so that front-line personnel who respond to crime scenes or work in the field can utilize appropriate tools that provide actionable information early in an investigation and facilitate real-time data analysis, particularly for major investigations" [42 p103];
- "[t]ool testing and validation … in a timely and consistent manner in order to verify their suitability for use in casework" due to frequent updates by commercial vendors and in-house developed tools within the digital evidence community" [42 p103].

In 2018, NIST began their technical merit evaluations and included a review of the foundation for digital forensics examinations. In June 2022, NIST published its final report, which focused primarily on "techniques for examining digital data stored in mobile device memory, computer memory, or secondary storage in an active computer[, such as] … hard drives, flash drives, removable drives, or "external" storage media such as CD, DVD, or memory cards" [131 p5]. The NIST report did not, however, focus on "how well the techniques are used in practice, the best practices for implementing techniques, or limitations placed on usage by the courts[, as well as o]ther digital forensics topics such as network analysis and multimedia (video, audio) … [and] other issues such as improved methods for tool validation and verification, privacy, or legal issues, and managing forms of bias within forensic practice" [131 p5]. Ultimately, the NIST report concluded: "[t]he overall finding … is that digital evidence examination rests on a firm foundation based in computer science … [and t]he application of these computer science techniques to digital investigations is sound and only limited by the difficulties of keeping up with the complexity and rapid pace of change in [information technology]" [131 pp52-53].

Although the NIST report concluded "digital evidence examination rests on a firm foundation based in computer science" [131 p52], it identified several challenges and needs

---

[105] NIJ recommendations also built on challenges and needs identified and prioritized by community members during a joint workshop in 2014 hosted by RAND and the Police Executive Research Forum, which was funded by NIJ through the Priority Criminal Justice Needs Initiatives [42 pp105-106].

facing the discipline which might have implications for future research and development, including:

- "[d]igital forensics is dependent on an understanding of computers and how they work" [131 p20];
- "[c]omputer technology evolves rapidly; however, some attributes of computers last for decades and some only for a few weeks" [131 p20];
- "forensic examiner[s] need[] to be aware of changes in computing technology relevant to the examination being performed [because c]hanges in digital technology introduce the possibility for incomplete analysis or for misunderstanding of the meaning of artifacts" [131 p20];
- "[e]very digital forensic technique should undergo peer review, formal testing, or error rate analysis … [and although helpful informal mechanisms exist,] … this process is not comprehensive … [and therefore, e]fforts to promote additional rapid peer assessment should be promoted" [131 p28];
- "[w]hen using techniques to recover deleted or hidden artifacts the examiner must determine the relevance of the recovered information as it may be incomplete or merged with irrelevant information" [131 p36];
- "[s]earching tools have limitations based on multiple ways that computers store information … [and, although] digital search tools are very effective at finding information … there is a possibility that the data will be missed because a tool does not have the capability to find it" [131 p37];
- "[i]f someone has taken steps to change information in digital evidence to mislead an examiner, it may be difficult to detect the changes [and therefore i]dentification of deliberate obfuscating changes relies on the skill of the examiner" [131 p40];
- "[d]igital processes tend to have systematic errors rather than random errors[; t]herefore, an error mitigation analysis provides more information and is the correct way to manage uncertainty [because a]n error rate is only useful where there are random errors" [131 p44];
- "[w]hen error rates are provided, it is important for the user to understand the context of the numbers[, because f]or some forensic techniques, the error rates may vary significantly based on attributes of the technology and usage patterns" [131 p44];
- "[i]t is not feasible to test all combinations of tools, run time environments, and digital evidence sources" [131 p45];
- the need for "[b]etter sharing of forensic knowledge including new and changed artifacts, new techniques, tool limitations and workarounds, and other forensic insight … [with] a more structured approach" [131 p52];
- the need for "[m]ore efficient and consistent approaches to testing forensic tools … [using a] more structured approach [to] increase efficiency [and] … [less variability in] test coverage and test results" [131 p52];
- the need for "[b]etter sharing of forensic reference data … [to enable better] tool testing, training and education, and research and development of new tools and techniques" [131 p52];

- the need for "[b]etter analysis of how digital evidence is used and whether there have been incorrect or misleading conclusions … [and h]aving this information centrally collected" [131 p52];
- the need for "[b]etter understanding of bias and effective bias minimization measures … [because] forensic examiners are exposed to knowledge about people involved in a case, such as seeing their photos and reading their text messages … [and] the forensic examiner may need to interact with an investigator" [131 p52];
- the need for "[b]etter understanding of the types and characterizations of mistakes examiners make in interpreting tool results" [131 p52].

As of 2023, recommendations from OSAC include studies relating to:

- tools and techniques for analyzing virtual machines and virtual file systems
- analysis of hash authentications
- triaging mobile applications [150]

5.2.3.7.2 Video / Imaging Technology and Analysis

Methods relating to forensic video and imaging technology and analysis have had some relation to the assessments made relating to digital evidence. As of 2023, recommendations from OSAC include studies relating to:

- detecting deepfake multimedia content
- factors affecting image quality when extracting a still image from video
- authentication of PDF documents
- development of a software validation repository
- black box and white box studies relating to vehicle comparisons, including analyses of vehicle year, make, and model
- determination of the size of the smallest detail required for tire and shoe comparisons [150]

5.2.3.7.3 Facial and Iris Identification

As of 2023, recommendations from OSAC include studies relating to:

- human factors in facial image comparison
- proper monitor selection and setup for forensic image examinations
- validation of physical stability of facial features of adults [150]

In 2023, the OSAC expanded the scope of the Facial Identification subcommittee to include iris identification. While iris patterns have been widely used as a biometric for nonforensic applications, the use of it in a forensic context is relatively new [171]. In 2022, NIST published a review of scientific literature relating to forensic iris and noted that despite the technology limitations from a decade ago, "[t]here are now demonstrations of iris recognition that can be plausibly applied to matters of forensic interest"; however, "[i]t is an open question whether the science of forensic iris is sufficiently established to satisfy the criteria that the National Research Council (NRC)/National Academy of Science (NAS) and the Presidential Council of Advisors on Science and Technology (PCAST) put forth for forensic science in general" [171 piii]. Specific recommendations for research and development include the need for:

- "Research to better characterize the human ability to adjudicate iris image pairs";
- "Further studies of the statistics of visible iris features … to provide the underpinnings for the science of human comparison of iris images";
- "Datasets appropriate for scenarios of interest … to provide material suitable for training and testing iris image examiners" [171 pv].

5.2.3.7.4 Speaker Recognition

As of 2023, recommendations from OSAC include studies relating to:

- effects of data variations on speaker recognition results
- development of standard test data for speaker recognition
- detection of synthetic speech [150]

5.2.4. Considerations for Forensic Science Research and Standards Programs

Over the years since the 2009 NRC report, there has been increased attention on the need for research and development in forensic science. While the NRC expressed the need for stronger research foundations underpinning many forensic science disciplines and methods [6], others followed with similar findings and recommendations [13]. Some of the challenges and needs identified by the NRC in 2009 have been addressed to different extents while others remain [5]. To be more effective and impactful in addressing science and technology issues moving forward, there are three major considerations for forensic science research and standards programs (Box 5.2.4).

Box 5.2.4. Forensic science research and standards programs should:

(a) Create discipline-specific research roadmaps that clearly prioritize the research needs and objectives that most directly align with addressing knowledge gaps related to the validity and reliability of commonly practiced forensic science methods.

(b) Assess their internal capabilities and competencies to identify what science and technology challenges and needs are within scope for intramural or extramural research and consider optimal ways of structuring and administering programmatic activities to ensure they have the capacity to support unanticipated or time-sensitive projects.

(c) Establish strategic relationships through interagency agreements (IAAs) and cooperative research and development agreements (CRADAs), as appropriate, to provide the administrative underpinnings and enable efficient coordination, collaboration, and sharing of research and development activities, data, technologies, and related materials between other major or strategically significant government, academic, or industry partners.

Research priorities should not supplant development and validation of new and improved methods that yield advanced capabilities for the forensic science community; however, there should be higher prioritization placed on ensuring existing methods commonly used within the forensic science community are based on sound scientific foundations (or ensuring that the extent to which those existing methods are based on sound scientific foundations is known).[106] Further, the prioritizations should include discipline-specific gap assessments to determine the extent to which the research and development challenges and needs previously identified are still needed or relevant.

Research priorities should also consider the magnitude and significance of the intended outcome and impact, among other relevant factors. The research roadmaps should specify how current near-term (e.g., 1-2 years out), mid-term (e.g., 3-5 years out), and long-term (6-10 years out) research efforts align with those priorities and other research and development challenges and needs facing the community.[107] Additionally, the research roadmaps should be developed in a way that they can fold into a larger effort to establish a shared roadmap and strategy across the government for addressing the broader research and development needs of the forensic science community.

[106] Including those methods reflected in published or proposed documentary standards and best practices recognized, or under consideration for recognition, by the OSAC.

[107] The research roadmaps should align to a hierarchy of needs that clearly define the purpose and overarching outcome and impact the specific research project(s) are intended to enable. Further, the roadmaps should clarify how the specific research project(s) fit within a broader architecture of related research projects that, when combined, yield the intended outcome and impact. For more details relating to the development of a R&D strategic plan, see also [56 pp62-65].

5.3. Translation and Implementation

Translation and implementation of innovations in science and technology have been a challenging issue across many domains, and one that has been the subject of study for decades (e.g., [172]).[108] Driven by the challenges with implementing evidence-based practices in healthcare, the field of "implementation science" emerged in 2006 [173 p3], defined as "the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services and care" [174].[109] Since 2006, attention has been directed toward principles of implementation science across several domains and identifying and understanding key challenges, barriers, and facilitators affecting the uptake and adoption of information and technologies from research and development efforts.

5.3.1. Outside the Context of Forensic Science

The healthcare industry has translation and implementation challenges similar to those seen in forensic sciences [175]. Those issues have been the subject of extensive research, and can serve as a model for forensics, which is less well-studied. In 2010, a survey of dentists found "[t]he most common barriers to implementation [of evidence-based practices] were *difficulty in changing current practice model*, *resistance and criticism from colleagues*, and *lack of trust in evidence or research*" [176 p195], and in 2011, a leading textbook on evidence-based practices (EBP) in nursing and healthcare cites the most common barriers observed include:

- *Lack of EBP knowledge and skills*
- *Misperceptions or negative attitudes about research and evidence-based care*
- *Lack of belief that EBP will result in more positive outcomes than traditional care*
- *Voluminous amounts of information in professional journals*
- *Lack of time and resources to search for and appraise evidence*
- *Overwhelming patient [work] loads*
- *Organizational constraints, such as lack of administrative support or incentives*
- *Lack of EBP mentors*
- *Peer pressure to continue with practices that are steeped in tradition*
- *Resistance to change*
- *Lack of consequences for not implementing EBP*
- *Lack of autonomy over practice and incentives* [177 p17]

By 2012, the term "knowledge translation," defined as "ensuring that stakeholders are aware of and use research evidence to inform their health and healthcare decision-making …

[108] Rogers' classic social science theory describes the spread, or diffusion of innovation, as a social process with multiple determinants that extend beyond the innovation itself [172].
[109] Eccles and Mittman note: "Research continually produces new findings that can contribute to effective and efficient healthcare. However, such research cannot change outcomes unless health services and healthcare professionals adopt the findings into practice. Uneven update of research findings—and thus inappropriate care—occurs across settings, specialties and countries" [174 p1].

[which] recognizes that there are a wide range of stakeholders or target audiences for knowledge translation, including policy makers, professionals (practitioners), consumers, … researchers, and industry," had gained traction [178 p2]. The act of "knowledge translation" began to be framed as a strategic versus tactical effort to promote more effective outcomes, taking into account five key questions that are context-specific to the information, target audience, domain, and purpose for which the knowledge is being transferred:

1. *What should be transferred?*
2. *To whom should research knowledge be transferred?*
3. *By whom should research knowledge be transferred?*
4. *How should research knowledge be transferred?*
5. *With what effect should research knowledge be transferred?* [178 p2]

These questions provide a framework for approaching knowledge translation strategies, and when developing a strategy "[i]ndividuals involved in knowledge translation need to: identify modifiable and non-modifiable barriers relating to behavior; identify potential adopters and practice environments; and prioritise which barriers to target based upon consideration of 'mission critical' barriers" [178 p5]. By 2015, over 70 different strategies had been proposed with different levels of perceived feasibility and importance [179].[110]

In 2022, a "systematic review of reviews" was published that aimed to "synthesize existing literature [predominantly within the last decade] to elucidate the barriers and facilitators to the translation of health research into clinical practice" [180 pe3265], considering the issues across three different levels: "micro (individuals)," "meso (systems or organizations)," and "macro (economic and political)" [180 pe3272].

At the micro-level (individuals), key barriers included: "limited professional engagement in the research process, lack of time, insufficient critical appraisal skills and an inability of healthcare professionals to use the research findings and recommendations in clinical practice" [180 pe3272]. Other barriers included: "professionals' unfamiliarity with evidence-based practice concepts, lack of interest in updating knowledge on emerging best practices, and underestimation of the value of research" [180 pe3272]. Facilitators involved challenging "professionals' motivation and interests in addressing and studying research findings, given suitable packaging and targeted communication of results" [180 pe3273].

At the meso-level (systems or organizations), barriers included "[t]ime constraints, insufficient organizational resources, poor knowledge dissemination and lack of access to evidence and research." Other barriers included: "[i]nsufficient resources (materials and equipment) required for the implementation of research and inadequate facilities to conduct research[, as well as] … [w]orkforce shortage[s], … [i]nappropriate management, organisation of staff and workload density" [180 pe3273]. Facilitators required institutional support "through a

[110] Interactive webtools and various other implementation theories, models, and frameworks are also available (e.g., [297]).

policy or plan to implement" and included "[d]issemination of primary research findings across organisations and budgeting for research activities" to support implementation [180 pe3273].

At the macro-level (economic and political), the findings suggested that "[p]olicy makers [are] not sufficiently trained and skilled in research methods and [might] not perceive or observe alignment or integration between research and policy … [and] they remain doubtful about the utility of research findings" [180 pe3273]. Facilitators involved "identifying the stakeholders and developing robust collaboration and connections between policy makers and research staff, … [which] provides policy makers with updated information and knowledge in research and engages them in all research priorities … [as well as] helps them make evidence-based decisions" [180 pe3274]. Other facilitators included "developing trust across policy makers and researchers and developing guidelines that promote clinical best practices[, as well as] … involv[ing] stakeholders early in the research design and initiation process, as they are most likely to be affected by research output" [180 pe3274].

The challenges of translating research into practice are further complicated when the outputs of research and development activities include sophisticated tools and technologies. The full potential of technologies (e.g., those containing algorithms for statistical prediction or forecasting tasks, such as probabilistic models or decision tools) cannot be harnessed unless people are willing to rely on them [181]. Attempts to understand human-algorithm interactions from psychological and behavioral science perspectives gained increased interest following a study published in 2015 in which the term "algorithm aversion" was first coined, describing a phenomenon in which people rely on their own judgment for a variety of prediction and forecasting tasks, even while knowing that their own judgment is inferior to that of an algorithm [181]. Although facilitators have been proposed to promote overcoming barriers to "algorithm aversion," and include "giving people some control—even a slight amount—over an imperfect algorithm's forecast" [182], the specific reasons and contexts for which the phenomenon manifests have been debated [183].

Separate from the challenges relating to the adoption of algorithmic systems, when algorithms are based on or otherwise incorporate artificial intelligence and machine learning (AI/ML) methods, the issues are further complicated from a socio-technical perspective [184, 185]. Attention to these issues has accelerated in the last several years as AI/ML methods have become more prominent [186]. For example, in 2018, the AI Now Institute published its Algorithmic Impact Assessment Report focused on the growing use of AI and algorithmic decision-making in public agencies and promoting a means of helping communities assess when and where AI/ML based systems are appropriate [184]. In December 2018, the first G7 Multi-Stakeholder Conference on Artificial Intelligence was held in Montreal, Canada, bringing together AI researchers, advocates, and policy makers representing diverse perspectives across industry, academia, civil society, and government focused on "Enabling the Responsible Adoption of AI" [187]. In 2023, as directed by the National Artificial Intelligence Initiative Act of 2020 [188], NIST published the AI Risk Management Framework (AI RMF) "to offer a resource to the organizations designing, developing, deploying, or using AI systems to help manage the many risks of AI and promote trustworthy and responsible development and use of AI systems"

[185 p2]. The NIST AI RMF analyzes AI risks and trustworthiness, "outlining the characteristics of trustworthy AI systems, which include valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy enhanced, and fair with their harmful biases managed" [185 p12].[111] Further, the NIST AI RMF describes four specific functions (govern, map, measure, and manage), referred to as the AI RMF Core, that "provides outcomes and actions that enable dialogue, understanding, and activities to manage AI risks and responsibly develop trustworthy AI systems" [185 p20].[112] Finally, in October 2023, the President published the "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," which advances a "coordinated, Federal Government-wide approach" to "governing the development and use of AI safely and responsibly" [186].

5.3.2. Within the Context of Forensic Science

The literature on knowledge translation in forensic science is much more limited; however, over the last decade there has been increasing recognition of the need to bridge the gap between research and practice [52]. In June 2019, NIST hosted the Forensic Science Research Innovation to Implementation Symposium (RI2I), which brought together researchers, forensic science practitioners, and business and legal professionals, many of whom had experience in technology transfer, with the goal of "explor[ing] the challenges of transferring forensic science research into operations" [52]. Symposium participants discussed barriers and facilitators from four different perspectives: research, laboratory management, business, and legal.[113]

From a *research* perspective, key barriers to the translation of research innovations included:

- lack of incentives among researchers to implement their innovations in an operational environment beyond publication in peer-review journals;
- risk aversion among laboratories to implement new technologies due to small changes having potentially big consequences;
- little alignment in terms of research being conducted and the specific needs of the practitioner community;
- little collaboration between researchers and practitioners;
- ambiguity relating to how to conduct validation studies and how robust they need to be to implement new methods (e.g., number and variety of samples);
- lack of centralized resources and support for validation activities;
- limited budgets, time and resources to stay abreast of relevant literature, support new validation studies, and train on new methods or practices;

[111] A description of each of the AI RMF characteristics of trustworthy AI systems is also provided [185 pp12-18].
[112] A description of each of the AI RMF Core functions is also provided [185 pp20-23].
[113] Summaries of individual plenary speaker perspectives are provided in [52 pp4-16]. Summaries of "key points" from group breakout sessions are provided in [52 pp20-38].

- lack of standardization and information and data sharing across laboratories [52 pp20-24].

"The overall outcome of the discussions highlighted the need for a centralized resource(s) that would offer implementation services to forensic laboratories [relating to]: [c]reation of a centralized area or agency for protocol sharing, [c]reation of a federally funded validation services team, [and] [c]reation of a forensic clearinghouse or analogous mechanism for sharing information from field practitioners" [52 p24].

From a *laboratory management* perspective, key barriers to the translation of research innovations included:

- little alignment in terms of technology available and the specific laboratory needs (e.g., technology is often developed without close coordination from laboratory practitioners, it is often "too fancy" for what the lab needs and all the extra "bells and whistles" drive up the cost, and, when multiple options are available, laboratory managers often don't have the time to research which specific option is best;
- training and education required as part of adoption of new methods or technologies often create steep learning curves for staff, cause added stress among staff and management to identify and secure training resources and sources, and contribute to increased case backlogs;
- increased caseload due to higher demands to use the new methods, tools, or technologies that were not previously available or not available by laboratories in other jurisdictions;
- procurement processes, costs, and other secondary requirements related to the procurement (e.g., information technology (IT) infrastructure, maintenance, security) of new technologies can cause them to be difficult to justify compared to traditional methods and status quo;
- validation and standardization guidelines are lacking, which leads to time consuming and costly efforts that can be resource-prohibitive and discouraging for many laboratories;
- communication between laboratories and researchers is limited and oftentimes researchers develop methods that cannot be implemented due to IT, security, or limitations in infrastructure [52 pp25-28].

The discussions highlighted the value of (i) resource sharing, in which "a central organization could identify trends and evaluate research, then disseminate this information to laboratories with recommendations for new technology," "help create validation guidelines and standards for use that laboratories could follow," and conduct "the main validation on a new technology and then disseminate[] that information to other laboratories who can do site validations when they procure that new technology" [52 p28] and (ii) leveraging "external pressures from media, prosecutors, and legislators to hasten procurement" so that laboratories can "adhere to the *Daubert* standard or the *Frye* test" and maintain enduring admissibility of evidence, "which may require the acquisition of new technology so that the laboratory maintains the best technology" [52 p28].

From a *business* perspective, key barriers to the translation of research innovations included:

- risk in adopting methods or technologies that are not widely used by others, that have not yet gained legal acceptance, that have not yet been standardized, or which may be outdated (e.g., replaced by a newer innovation or version) by the time validation, implementation, and training is completed;
- trust that the new method or technology is sound, reliable, and can be defended during testimony;
- budget limitations often make the combination of upfront and secondary costs such as long-term costs, time considerations, IT impact, quality system changes, and training challenging to justify or afford;
- mis-aligned priorities and different perspectives on what is needed or why (e.g., laboratory quality/technical managers focused on failure rates and limits of detection versus laboratory operational managers focused on costs or throughputs; small commercial markets and low market demand with explicit requirements make investments more risky) [52 pp28-30].

The discussions emphasized the importance of (i) greater transparency into vendor life cycles (e.g., "more insights into expected life cycle times or help [from vendors to] create backwards compatibility with new innovations"); (ii) availability of independent, third-party validation, especially from a clear authoritative source … [to] help lower the perception of risk and court acceptance"; (iii) subject matter expert(s) to provide "an outside, unbiased perspective on what new technology might be beneficial to operations"; (iv) vendor training and/or proficiency testing so that laboratories can reduce the burden and "help them be more confident that their training is correct"; and (v) development of "better tools and systems for publicly funded entities to regularly measure the market demand of its service area" [52 pp30-31].

From a *legal* perspective, key barriers to the translation of research innovations included:

- the adversarial nature of the criminal justice system can create cultural challenges between scientists and litigators, with scientists, litigators, and judges often differing in their scientific literacy and educational backgrounds;
- the need for validations to be done properly, thoroughly, and be "bullet proof," which is often challenging given lack of guidelines, standards, and resources;
- limited resources to support proper validation, training (including litigators, judges), and admissibility hearings that would likely be required as part of the implementation of the new methods or technologies;

- risk aversion among courts to embrace new methods or technologies due to potential for errors or unintended consequences that might contribute to injustice, and the need for finality in legal decisions that can be challenged when new methods or technologies might have implications for previously resolved cases;
- miscommunications and misunderstandings caused by differences in terms used, expertise, and backgrounds between scientists and litigators can create confusion and increase the potential for biased presentation of information or interpretation of testimony, particularly if the scientist is unable to adequately answer questions or provide relevant information or when litigators do not fully understand the strengths or limitations relating to the method or technology;
- variable admissibility criteria both within and between jurisdictions due to differences in legal requirements, interpretations of law, legal rulings, and definitions/interpretations of "general acceptance" as it relates to the *Frye* or *Daubert* standard;
- the potential for retrospective testing or retesting of old cases,[114] which might require resources beyond what is available to the laboratory based on existing budgets [52 pp31-38].

The discussions underscored the need for (i) better education and interdisciplinary familiarity with concepts and issues between scientists and court personnel (e.g., scientists needing better education concerning legal requirements and court personnel needing better education on scientific issues) and (ii) sufficient resources to support all aspects of implementation, including the need for robust and defensible validations, training of practitioners and others within the criminal justice community, and support for admissibility hearings, if necessary [52 pp31-38 and 50-51].

While the NIST RI2I symposium dealt with the implementation of innovations within the context of forensic science across a broad spectrum of methods, technologies, and disciplines, in recent years, there has also been a growing debate over the use of technologies in the criminal justice system that include algorithms [189–192]. Many of the concerns tend to be framed from a socio-technical perspective and focused on those algorithms that rely on AI/ML methods [189–191]. Within the last few years, law enforcement leaders have strategized on how the benefits of these algorithms can be leveraged in an operational context while maintaining public trust and upholding societal values by ensuring the algorithms are characterized by fairness, accountability, transparency, and explainability [189–191]. The risks associated with algorithmic biases have also been raised as a point of concern [193], as well as ethical considerations related to data privacy and the use of law enforcement data and databases and the risks they pose for perpetuating systemic biases in surveillance and investigative practices [194]. As calls for statistical models and computational methods become more prominent in forensic science (e.g., [13]), efforts to understand the benefits and risks of

[114] Although this is listed as a challenge with implementation of new technologies, there was consensus that "this factor would have minimal impact on the initial decision to implement or admit the new technology" in court; and, "[w]hile the ability to do testing in old cases may raise the concerns of 'opening the floodgates' for additional testing, the group agreed that taking a retrospective look at older cases needed to proceed and that any resulting issues would need to be dealt with in an appropriate manner. Not doing retrospective testing was viewed as an unacceptable decision" [52 pp37-38].

algorithmic systems have been brought into focus [195] and studies have sought to elicit different perspectives [28] and identify distinct barriers to their translation and implementation and propose approaches that promote a practical and responsible implementation [175]. Although the barriers to the translation and implementation of algorithms are similar to those of other methods and technologies intended for forensic science applications, they have additional nuances to consider, and the manner in which algorithms are developed and deployed might have implications on whether, and to what extent, practitioners are willing to use the algorithms and courts are willing to rely on evidence produced by the algorithms [175].

While the translation and implementation of science and technology innovations within the context of forensic science requires consideration of multiple complex issues through several different lenses, many of the barriers can be summarized as relating to:

- individual (micro) issues: lack of incentives, risk aversion, resistance to change;
- organizational (meso) issues: limited resources to support evaluations and validations, training, and related costs associated with quality management systems, information technology systems, legal hearings, and increased casework demands, lack of administrative support/priorities;
- community-wide (macro) issues: lack of guidelines and standards concerning appropriate and acceptable use, lack of actual requirements/demands or consequences, socio-technical and ethical considerations, legal implications [28, 52, 175].

Many of these issues are similar to the barriers cited in domains outside of forensic science. Although there are distinct differences between the domains and applications, the significance of the overlap suggests the forensic science community can learn from and adapt strategies from the principles of "implementation science" and apply them toward specific issues relating to the translation and implementation of science and technology innovations within the context of forensic science.[115]

5.3.3. Considerations for Forensic Science Research and Standards Programs

While there has been increased attention to the need for research and development to underpin many forensic science disciplines in the years following the 2009 NRC report, so too has there been recognition that the outcomes and impact of research and development activities will not be realized unless they are translated and implemented into practice [52]. Although many of the challenges facing the forensic science community relating to the

[115] E.g., The importance of collaborations and partnerships has been identified as a common and domain agnostic facilitator for translation and implementation of innovations. Calls for collaborative approaches to promote cost sharing across entities and for instrument and equipment validations have been proposed as potential strategies over the last few years [42 p5, 45, 50–52] but often lacked a central entity to lead the collaboration. In recent years, however, the ASCLD Forensic Research Committee (FRC) has established several initiatives to promote information sharing and collaborative partnerships (e.g., [298]). Further, in 2022, several laboratories announced a partnership to establish the National Technology Validation and Implementation Collaborative (NTVIC) with the mission "to share resources and strategies to rapidly implement technology and new methods into publicly funded forensic science service provider (FSSP) and forensic science medical provider (FSMP) facilities in a scientifically sound and defensible manner" [299]. Related, in 2023, additional attention has been directed toward understanding major barriers and facilitators to promote stronger academic-practitioner partnerships forensic science [300].

translation and implementation of science and technology innovations are outside the direct control of forensic science research and standards programs, there are several ways that those programs might contribute to lowering the barriers and reducing the burden on forensic science laboratories, end users, and others within the criminal justice system. Accordingly, to be more effective in this respect moving forward, there are four major considerations for forensic science research and standards programs (Box 5.3.3).

Box 5.3.3. Forensic science research and standards programs should:

(a) Create a clear translation pipeline from research to practice that defines the categories, phases, and stages of research as they progress from ideation to implementation to impact.

(b) Prioritize investments that support the translation and implementation of outputs from their programmatic activities into practice, such as:

i. establishing strategic collaborations and partnerships with forensic science laboratories, end-users, and other government and academic researchers to ensure research and development activities are coordinated, tailored to the specific needs of the forensic science community, and have a clear pathway for the translation and implementation of the outputs by early-adopters that can further assess the methods or technologies in operational contexts and champion the adoption across the community;

ii. providing a centralized means of sharing resources, such as data, datasets, validation guidelines and reports, protocols, and training materials relating to the performance characterization and validation of methods and technologies having broad applicability across the forensic science community;

iii. supporting a centralized means of conducting research, development, testing and evaluation activities relating to the performance characterization and validation of methods and technologies having broad applicability across the forensic science community, including the creation of model protocols and training materials that can be leveraged by the forensic science community.

(c) Establish strategic outreach plans that prioritize dissemination of their research and development activities, outputs, and key findings using various approaches to reach broad and diverse audiences.

(d) Identify an appropriate set of procedures and metrics to regularly measure outcomes and impacts of their programmatic activities resulting from the translation and implementation of their outputs.

As part of the translation pipeline, there should be clearly defined "translational readiness levels (TRL)" and criteria that describe the maturity of specific research and development efforts that can be tracked as they progress along the pipeline.[116] The TRLs should be created with coordination among other programs to promote consistency, be specific to forensic science applications, and provide a means of not only communicating the "readiness" of research and development innovations for operational use in a manner that is commonly understood across all members of the forensic science community (e.g., researchers, practitioners, academia, policy makers, litigators), but also prioritizing activities and investments relating to those efforts as it relates to a broader research and development strategy.

Activities that support translation and implementation should be included in the development of strategic roadmaps, be grounded in evidence-based approaches from concepts of "implementation science," and align to the needs and recommendations of community members to enable the adoption of new methods and technologies with less burden.[117] Similarly, outreach activities should be geared toward informing, instructing, and inspiring the translation and implementation of programmatic outputs among researchers, practitioners, and policy makers and include evidence-based approaches from concepts of "implementation science" to maximize receptivity and consumption of the information.

Finally, procedures and metrics to measure outcomes and impacts should also be included in the development of strategic roadmaps as a means of assessing progress toward achieving the strategic priorities for the programs and maintaining accountability of the investments.[118]

6. Legal and Regulatory Landscape

Issues and trends relating to the complex framework of laws, regulations, standards, and guidelines that govern the practice of forensic science and its application in legal proceedings, and how that might impact the forensic science community, are evaluated in the legal and regulatory landscape. Although specific laws and regulations vary across jurisdictions, there are common elements that are broadly applicable across the forensic science community. Looking through a legal and regulatory lens, there are two key issues that are most relevant and potentially impactful to forensic science research and standards programs: (i) admissibility of evidence and (ii) oversight and accountability.

[116] E.g., similar efforts have been undertaken by the National Institute of Health and further adaptations to the traditional "Technology Readiness Levels" used for assessing the "readiness" of technologies for use in acquisition programs and projects [301, 302].
[117] E.g., Such as those key-pints and recommendations identified during the NIST-hosted 2019 R2I2 symposium [52].
[118] These metrics should be developed with coordination across other programs to provide a common foundation for assessing the outcomes and impacts of a broader shared strategy, and to provide the potential for these metrics to be tied to other initiatives (such as extramural research grants and capacity building grants). However, in the absence of a shared set of appropriate metrics across the government, forensic science research and standards programs need not wait for such coordination to occur. A set of metrics tailored to the outputs of individual programs can be established and serve as a basis for coordinating across other entities when appropriate.

6.1. Admissibility of Evidence

In 2009, the NRC noted "[t]he law's greatest dilemma in its heavy reliance on forensic evidence … concerns the question of whether—and to what extent—there is *science* in any given forensic science discipline" [6 p87]. The NRC continued "[t]here are two very important questions that should underlie the law's admission of and reliance upon forensic evidence in criminal trials: (1) the extent to which a particular forensic discipline is founded on a reliable scientific methodology that gives it the capacity to accurately analyze evidence and report findings and (2) the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, or the absence of sound operational procedures and robust performance standards" [6 p87].

In the United States, the admission of evidence in the federal court system is governed by the Federal Rules of Evidence, which affect both criminal and civil litigation. Although each state has its own evidence admissibility rules, they are often similar to the Federal Rules [196]. For purposes of this assessment, therefore, attention will be directed towards the admissibility of evidence in a federal context. Admissibility is also governed by Constitutional provisions. While evidence might be admissible according to evidentiary rules, it might be excluded if it violates Constitutional rights of the defendant. Thus, decisions relating to the admissibility of forensic science evidence (for purposes of criminal prosecution) require evaluation of several issues.

6.1.1. Evidentiary Rules

When considering the admissibility of forensic science evidence under evidentiary rules, the court must weigh several issues. One of the first issues is whether the evidence proffered is "relevant" in accordance with Rule 401 [197]. Rule 401 provides that "evidence is relevant if: (a) it has any tendency to make a fact more probable or less probable than it would be without the evidence; and (b) the fact is of consequence in determining the action" [197]. A second issue to consider is whether the evidence is admissible under Rule 403 [198]. Rule 403 provides that "[t]he court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence" [198]. A third issue to consider, specifically relating to testimony by expert witnesses, is whether the evidence is admissible under Rule 702 [199]. Prior to December 2023,[119] Rule 702 provided that:

> *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:*
>
> *(a) 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;*

[119] In December 2023, Rule 702 was updated [214]. The update to Rule 702 is discussed later in this section.

(b) *the testimony is based on sufficient facts or data;*

(c) *the testimony is the product of reliable principles and methods; and*

(d) *the expert has reliably applied the principles and methods to the facts of the case.* [199]

When the Federal Rules were first promulgated in 1975, the prevailing means for determining the admissibility of scientific evidence was governed by the 1923 landmark decision in *Frye v. United States* [30]. According to *Frye*, the admissibility of scientific evidence is based on whether or not it had gained "general acceptance" in the relevant scientific community [30]. However, "[a]fter the promulgation of Rule 702, litigants, judges, and legal scholars remained at odds over whether the rule embraced the *Frye* standard or established a new standard" [6 p89].[120] It was not until 1993 when the U.S. Supreme Court finally clarified that Rule 702, not *Frye*, controlled the admission of expert testimony of scientific evidence in federal courts, as part of their decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* [31 p589].[121] According to *Daubert*, the Court ruled that a "trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable" [31 p589]. The Court continued, noting "*evidentiary reliability* will be based upon *scientific validity*" [31 pp590-591].[122] In detailing the *Daubert* standard, the Court pointed to several factors that might be considered by a trial judge, stating:

> *Faced with a proffer of expert scientific testimony under Rule 702, the trial judge, pursuant to Rule 104(a), must make a preliminary assessment of whether the testimony's underlying reasoning or methodology is scientifically valid and properly can be applied to the facts at issue. Many considerations will bear on the inquiry, including 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 and the existence and maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community. The inquiry is a flexible one, and its focus must be solely on principles and methodology, not on the conclusions they generate.* [31 p580]

However, the Court also noted "[c]ross-examination, presentation of contrary evidence, and careful instruction on the burden of proof, rather than wholesale exclusion under an

[120] NRC noted that the first version of Rule 702 provided that "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training or education, may testify thereto in the form of an opinion or otherwise" [6 p89].

[121] The Court noted: "*Frye* made 'general acceptance' the exclusive test for admitting expert scientific testimony. That austere standard, absent from, and incompatible with, the Federal Rules of Evidence, should not be applied in federal trials" [31 p589].

[122] See footnote 9 in [31 pp590-591].

uncompromising 'general acceptance' standard, is the appropriate means by which evidence based on valid principles may be challenged" [31 p580].[123]

Subsequently, in 2000, Rule 702 was amended "in response to *Daubert*" and other cases that had applied the *Daubert* standard [200]. In addition to codifying the factors provided by the *Daubert* standard, the Committee noted "several other factors relevant in determining whether expert testimony is sufficiently reliable to be considered by the trier of fact" [200], which include:

(1) Whether experts are "proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying." Daubert v. Merrell Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1317 (9th Cir. 1995) [31].
(2) Whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion. See General Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997) (noting that in some cases a trial court 'may conclude that there is simply too great an analytical gap between the data and the opinion proffered' [201].
(3) Whether the expert has adequately accounted for obvious alternative explanations. See Claar v. Burlington N.R.R., 29 F.3d 499 (9th Cir. 1994) (testimony excluded where the expert failed to consider other obvious causes for the plaintiff's condition) [202]. *Compare Ambrosini v. Labarraque, 101 F.3d 129 (D.C.Cir. 1996) (the possibility of some uneliminated causes presents a question of weight, so long as the most obvious causes have been considered and reasonably ruled out by the expert)* [203].
(4) Whether the expert "is being as careful as [they] would be in [their] regular professional work outside [their] paid litigation consulting." Sheehan v. Daily Racing Form, Inc., 104 F.3d 940, 942 (7th Cir. 1997) [204]. *See Kumho Tire Co. v. Carmichael, 119 S.Ct. 1167, 1176 (1999) (Daubert requires the trial court to assure itself that the expert 'employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field')* [205].
(5) Whether the field of expertise claimed by the expert is known to reach reliable results for the type of opinion the expert would give. See Kumho Tire Co. v. Carmichael, 119 S.Ct. 1167, 1175 (1999) (Daubert's general acceptance factor does not 'help show that an expert's testimony is reliable where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called generally accepted principles of astrology or necromancy') [205]*; Moore v. Ashland Chemical, Inc., 151 F.3d 269 (5th Cir. 1998) (en banc) (clinical doctor was properly precluded from testifying to the toxicological cause of the plaintiff's respiratory problem, where the opinion was not sufficiently grounded in scientific methodology)* [206]*; Sterling v. Velsicol Chem. Corp., 855 F.2d 1188 (6th Cir. 1988) (rejecting testimony based on 'clinical ecology' as unfounded and unreliable)* [207]. [200]

[123] The Court also framed this as "[v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence" [31 p596].

Although the 2000 amendment to Rule 702 provided more explicit criteria for considering the admissibility of expert testimony and connected decisions of admissibility with principles of scientific validity, when the NRC released their report in 2009, they stated: "[n]o judgment is made about past convictions and no view is expressed as to whether courts should reassess cases that already have been tried" [6 p85]. However, they also note "[t]he report finds that the existing legal regime—including the rules governing the admissibility of forensic evidence, the applicable standards governing appellate review of trial court decisions, the limitations of the adversary process, and judges and lawyers who often lack the scientific expertise necessary to comprehend and evaluate forensic evidence—is inadequate to the task of curing the documented ills of the forensic science disciplines" [6 p85]. The NRC continued, noting "[t]his matters a great deal … [and] every effort must be made to limit the risk of having the reliability of certain forensic science methodologies judicially certified before the techniques have been properly studied and their accuracy verified" [6 pp85-86]. The NRC concluded:

> *Law enforcement officials and the members of society they serve need to be assured that forensic techniques are reliable. Therefore, we must limit the risk of having the reliability of certain forensic science methodologies condoned by the courts before the techniques have been properly studied and their accuracy verified. … [However,] the adversarial process relating to the admission and exclusion of scientific evidence is not suited to the task of finding "scientific truth." The judicial system is encumbered by, among other things, judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner, trial judges (sitting alone) who must decide evidentiary issues without the benefit of judicial colleagues and often with little time for extensive research and reflection, and the highly deferential nature of the appellate review afforded trial courts' Daubert rulings. Furthermore, the judicial system embodies a case-by-case adjudicatory approach that is not well suited to address systematic problems in many of the various forensic science disciplines. Given these realities, there is a tremendous need for the forensic science community to improve. Judicial review, by itself, will not cure the infirmities of the forensic science community. … With more and better educational programs, accredited laboratories, certified forensic practitioners, sound operational principles and procedures, and serious research to establish the limits and measures of performance in each discipline, forensic science experts will be better able to analyze evidence and coherently report their findings in the courts. The present situation, however, is seriously wanting, both because of the limitations of the judicial system and because of the many problems faced by the forensic science community.* [6 pp109-110]

In the years following the release of the NRC report, it was raised in legal arguments on both sides [208]. Defense litigators often used the report as a basis for challenging admissibility, attempting to illustrate the lack of studies demonstrating the scientific validity underlying the methods proffered by prosecutors; however, prosecutors argued the report was never intended to affect decisions concerning admissibility [208 pp591-592]. Although different viewpoints exist around the extent to which the report ought to affect admissibility decisions, following a review of cases that cited or referenced the NRC report in their legal filings between 2008 and 2014, courts appear to have been hesitant to fully defer to the findings of the report and gave "relatively little weight to 'science' even when available as an official report from an authoritative institution" [208 p585].

In 2016, however, the PCAST report was much more explicit about the extent to which legal decisions concerning reliability are (or should be) based on scientific principles of validity [13]. Specifically, the PCAST noted:

> *The admissibility of expert testimony depends on a threshold test of, among other things, whether it meets certain legal standards embodied in Rule 702. These decisions about admissibility are exclusively the province of the courts. Yet, … the overarching subject of the judge's inquiry under Rule 702 is "scientific validity." It is the proper province of the scientific community to provide guidance concerning scientific standards for scientific validity. PCAST does not opine here on legal standards, but seeks only to clarify the scientific standards that underlie them. For complete clarity about our intent, we have adopted specific terms to refer to the scientific standards for two key types of scientific validity, which we mean to correspond, as scientific standards, to the legal standards in Rule 702 (c,d):*
>
> *(1) by "foundational validity," we mean the scientific standard corresponding to the legal standard of evidence being based on "reliable principles and methods," and*
>
> *(2) by "validity as applied," we mean the scientific standard corresponding to the legal standard of an expert having "reliably applied the principles and methods."* [13 p43]

The PCAST went further, making several recommendations to the federal judiciary on these issues, including:

> *When deciding the admissibility of expert testimony, Federal judges should take into account the appropriate scientific criteria for assessing scientific validity including:*
>
> *(i) foundational validity, with respect to the requirement under Rule 702(c) that testimony is the product of reliable principles and methods; and*
>
> *(ii) validity as applied, with respect to [the] requirement under Rule 702(d) that an expert has reliably applied the principles and methods to the facts of the case.* [13 p145]

In the years following the release of the PCAST report, it too was raised in legal arguments, and has had some effect on admissibility decisions [209].[124] Motivated in part by

[124] A list of post-PCAST cases between 2016 and 2022 that cited or referenced the PCAST report can be found online [209].

the fact that “a number of recent federal and state court opinions have cited the [PCAST] Report as support for limiting the admissibility of firearms/toolmarks evidence in criminal cases,” in 2021, the DOJ issued a statement arguing the claims made by the PCAST were “incorrect” and “erroneous” [20].

The PCAST report was also influential in the decision to consider amendments to Rule 702 affecting admissibility decisions systemically [210]. Shortly after its release, efforts were taken to amend Rule 702 again to provide clearer direction to judges relating to decisions of admissibility of expert testimony [210].[125] Between 2017 and 2022, the issues were debated by the Judicial Conference Committee on Rules of Practice and Procedure [210]. In summarizing the issue, the Committee noted:

> *The proposed amendments to Rule 702’s first paragraph and to Rule 702(d) are the product of Advisory Committee work dating back to 2016. As amended, Rule 702(d) would require the proponent to demonstrate to the court that “the expert’s opinion reflects a reliable application of the principles and methods to the facts of the case.” This language would more clearly empower the court to pass judgment on the conclusion that the expert has drawn from the methodology. In addition, the proposed amendments as published would have required that “the proponent has demonstrated by a preponderance of the evidence” that the requirements in Rule 702(a) – (d) have been met. This language was designed to reject the view of some courts that the reliability requirements set forth in Rule 702(b) and (d) – that the expert has relied on sufficient facts or data and has reliably applied a reliable methodology to the facts – are questions of weight and not admissibility, and more broadly that expert testimony is presumed to be admissible.* [211 pp23-24]

On April 24, 2023, the U.S. Supreme Court notified Congress that it had approved the proposed amendment to Rule 702 [212, 213], which took effect on December 1, 2023 [199]. The 2023 amendment to Rule 702 is reflected as follows, with the new language underlined and bolded and the omitted language struck through:

> *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* ***<u>the proponent demonstrates to the court that it is more likely than not that</u>****:*
>
> *(a) 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;*
>
> *(b) the testimony is based on sufficient facts or data;*

[125] “The project began with a symposium on forensic experts and *Daubert* held at Boston College School of Law in October, 2017. That Symposium addressed, among other things, the challenges to forensic evidence raised in a report by the President’s Council of Advisors on Science and Technology. A Subcommittee on Rule 702 was appointed to consider possible treatment of forensic experts … the Subcommittee did express interest in considering an amendment to Rule 702 that would focus on one important aspect of forensic expert testimony --- the problem of overstating results (for example, an expert claiming that her opinion has a ‘zero error rate’, where that conclusion is not supportable by the expert’s methodology)” [210].

(c) *the testimony is the product of reliable principles and methods; and*

(d) *the ~~expert has reliably applied~~ **expert's opinion reflects a reliable application of** the principles and methods to the facts of the case.* [199, 211][126]

The amendment accomplishes two things. First, it makes explicit that the "preponderance of evidence" standard applies to all four prongs of the Rule [214]. Second, it emphasizes the need for courts to focus not only on the "principles and methods" but also the opinion stemming from them[214], which is a shift from the original decision passed down from *Daubert*.[127] Collectively, these changes are "especially pertinent to the testimony of forensic experts in both criminal and civil cases" [214].

6.1.2. Constitutional Provisions

In the United States the admissibility of evidence in criminal litigation also has a Constitutional dimension:

Amendment V provides that "[n]o person shall be … deprived of life, liberty, or property, without due process of law" [215]. Amendment XIV provides that "[n]o state shall … deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws" [216]. Effectively, amendment XIV extends the legal obligation of "due process" to all states and not just limited to the federal government [216]. The concept of "due process" provides "assurance that all levels of American government must operate within the law ('legality') and provide fair procedures" [217].

Amendment VI provides that "[i]n all criminal prosecutions, the accused shall enjoy the right … to be confronted with the witnesses against him" [218]. The "Confrontation Clause" of amendment VI applies specifically to testimonial statements [218]. In recent years, the extent to which this applies to forensic evidence has been addressed[219, 220]. In *Melendez-Diaz v. Massachusetts* (2009) the Court held that the admission of forensic lab analysts' affidavits—reporting that material seized from the defendant was cocaine—violated the Confrontation Clause because affidavits were testimonial and the "analysts were 'witnesses' for purposes of the Sixth Amendment" [219]. Then, in *Bullcoming v. New Mexico* (2011), the Court clarified that when the government proffers forensic laboratory reports containing testimonial certifications "made for the purpose of proving a particular fact," the Confrontation Clause applies, and the accused has the right to be confronted by the analyst who made the certification [220]. In *Bullcoming v. New Mexico* (2011), the Court also clarified that surrogate testimony by another analyst who is familiar with the procedures but did not "sign the

[126] See: Proposed Amendments to the Federal Rules of Evidence, Rule 702. Testimony by Expert Witnesses, Rules Appendix E-10 in [211].

[127] Noting that "its focus must be solely on principles and methodology, not on the conclusions they generate" [31 p580]. See Committee Notes on Rule (2023 Amendment) [214].

certification or perform or observe the test reported in the certification," is insufficient to satisfy this right [220]. In 2023, *Smith v. Arizona* raised another dimension to the applicability of the Confrontation Clause and the use of substitute witnesses which has been brought before the U.S. Supreme Court [221]. Oral arguments occurred in January 2024 and the key question presented before the Court was: "Whether the Confrontation Clause of the Sixth Amendment permits the prosecution in a criminal trial to present testimony by a substitute expert conveying the testimonial statements of a nontestifying forensic analyst, on the grounds that (a) the testifying expert offers some independent opinion and the analyst's statements are offered not for their truth but to explain the expert's opinion, and (b) the defendant did not independently seek to subpoena the analyst" [221].

6.1.3. Admissibility Challenges

In recent years, there has been several admissibility challenges involving forensic evidence, often pointing to findings from the NRC and PCAST reports as a basis for the claims [208, 209]. After the PCAST report, firearms and toolmark analysis, latent fingerprint analysis, and DNA mixture interpretation in particular have faced increased scrutiny [209].[128] Challenges have not only focused on the validity and reliability of the methods overall, but also whether they were appropriately applied in the case at hand and the extent to which the conclusions reported are supported by the evidence [209]. Judicial dispositions of these challenges, however, have been mixed (both in trial courts and appellate courts) [209]. In some cases, judges have taken judicial notice of past precedent and dismissed the findings of these reports, while in other cases, judges have raised concerns and echoed the findings from these reports yet still ruled in favor of admissibility, arguing the issues go to the weight of the evidence rather than admissibility [208, 209]. Others, however, have either excluded the evidence all together or limited the testimony in some way to mediate the certainty that can be expressed by the expert (e.g., "more probable than not" versus "100% certainty") [208, 209].

In addition to challenges that have been raised over the years relating to the validity and reliability of traditional forensic science methods, the use of sophisticated software programs and computational algorithms have also been the subject of challenges [192, 195, 222–225]. These challenges focused not only on the validity and reliability of the underlying methods, but also whether the methods were reliably implemented into the software applications and free of errors and algorithmic biases [192, 195, 222–225]. Further, these challenges have raised potential issues concerning admissibility based on Rule 702 and *Daubert*, as well as issues relating to potential violations of Constitutional rights to due process and confrontation [192, 222–225]. Many legal scholars have discussed the lack of transparency around the innerworkings of these algorithms, which can stifle meaningful scrutiny and accountability and therefore infringe on Constitutional rights of criminal defendants [192, 222–225]. Although initial concerns centered around the accessibility to source-code for breath alcohol machines in the early 2000s, attention within the last decade has been directed toward probabilistic genotyping software [192, 222–224]. Although some legal scholars have argued accessibility to

[128] Reference [209] provides a list of post-PCAST cases between 2016 and 2022 that cited or referenced the PCAST report.

source-code is not an absolute requirement for the admissibility of these algorithms [224], the issues are more complicated when those algorithms are based on AI/ML methods for which source-code is often uninterpretable and the innerworkings are abstract [225].[129]

Finally, although much of the discussion relating to legal challenges has been aimed at traditional forensic science disciplines (e.g., DNA, fingerprints), which were the focus of both the NRC and the PCAST reports, similar issues are applicable to the growing field of digital forensics, which is further complicated by the technical complexities of the systems and tools [131, 226]. Although the scientific principles underlying digital forensics are considered sound since they rely on well-established principles of computer science, the rate at which digital technologies are evolving requires digital forensics tools to evolve at a similar rate [131]. "This poses serious problems for meeting requirements of *Daubert*—i.e., being able to demonstrate that digital evidence presented in court is reliable" [226 p13]. Consequently, not only is there a need for these tools to constantly be tested to evaluate and demonstrate their validity and reliability, but also the rate at which the technologies and digital forensics tools are evolving has resulted in challenges relating to analysts' expertise and the extent to which the practices are governed by accepted standards [227, 228].

6.1.4. Considerations for Forensic Science Research and Standards Programs

Although decisions relating to the admissibility of forensic science evidence rest solely within the purview of the courts, many of the issues surrounding those decisions relate to the extent to which the analytical methods and technologies underlying that evidence have been shown to be scientifically valid, reliable, and fair, causing courts to evaluate these legal and technical complexities in their decision-making. Thus, consideration of the various issues that might impact admissibility as they relate to matters of science and technology are relevant to forensic science research and standards programs. Accordingly, there are four major considerations for forensic science research and standards programs (Box 6.1.4).

[129] Although attention has been directed toward probabilistic genotyping software, similar issues are expected to be faced by other disciplines, such as pattern evidence examination (e.g., [28, 175]).

Box 6.1.4. Forensic science research and standards programs should:

(a) Be cognizant of the various legal rules and standards governing the admissibility of scientific evidence and their implications when developing new methods, technologies, and tools intended to be used by the forensic science community.

(b) Prioritize activities that promote greater objectivity, validity, reliability, and consistency to forensic science practices as well as evaluate the extent to which those outcomes are achieved.

(c) Prioritize activities that promote more efficient and effective testing of technologies and computational algorithms proposed for court use as it relates to assessing their validity, reliability, and fairness.

(d) Prioritize activities that provide guidance to officers of the court on issues relating to scientific principles of validity, reliability, and fairness.

While admissibility rules and standards governing should not drive or restrict the conduct of certain scientific activities or development of certain methods, they might inform how such research and development activities are approached and designed such that the outputs of those activities have greater likelihood of being adopted within the criminal justice system. For example, efforts to develop new technologies to promote greater objectivity and statistical underpinnings to the evaluation of evidence often require the use of computational methods. Consideration of the legal rules and standards governing the admissibility of evidence produced by those technologies will need to be considered during the design stage of the research activities, such that it can be demonstrated that the methods are not only scientifically valid and reliable, but also fair and trustworthy. This requires consideration of factors affecting transparency, accountability, explainability, and interpretability, among others, within the design and operation of the method.

Efforts to promote greater objectivity, validity, reliability, and consistency include research and development activities relating to improving existing methods or developing new methods. They also include evaluating and measuring the extent to which existing methods and practices conform to these principles (e.g., such as through supporting or conducting black-box studies, interlaboratory studies, and blind proficiency testing schemes).[130]

Effective testing of technologies and computational algorithms is important for ensuring their validity, reliability, and fairness. When done by a neutral and unbiased entity, results from such testing can provide a resource for forensic science practitioners and litigators to consider issues surrounding the use of specific technologies and algorithms, and for courts to make

[130] This is not to suggest that a single entity be solely responsible for administering such activities; rather, guidance for how such activities ought to be conducted is also relevant and important to promote and enable these activities to be conducted at scale across various entities.

informed decisions concerning the admissibility of evidence produced by those technologies and algorithms. With the rapid evolution of technologies, testing schemes must also be efficient and enable the activities to be conducted at scale consistently and systematically across the community.[131]

Additionally, efforts to promote more rigorous methods and practices also require prioritizing activities that provide guidance to the members of the forensic science and criminal justice communities so that attorneys and judges can be more effective at litigating forensic science evidence and making informed decisions related to admissibility. Such guidance could take many forms but might include how to assess basic principles of scientific validity, reliability, and fairness in general contexts as well as in specific applications and practices (i.e., methods or technologies employed in a specific case or circumstance).

6.2. Oversight and Accountability

In 2009, the NRC highlighted that "the quality of forensic practice in most disciplines varies greatly because of the absence of adequate training and continuing education, rigorous mandatory certification and accreditation programs, adherence to robust performance standards, and effective oversight," and "[t]hese shortcomings obviously pose a continuing and serious threat to the quality and credibility of forensic science practice" [6 p6]. The NRC continued, noting "the fact is that there are no requirements, except in a few states (New York, Oklahoma, and Texas), for forensics laboratories to meet specific standards for quality assurance or for practitioners to be certified according to an agreed set of standards" [6 p193].

There is no single entity that has the responsibility and authority to oversee and enforce specific practices relating to the examination, interpretation, and reporting of forensic evidence in a consistent and systematic manner across the United States [6, 229]. Rather, this responsibility is distributed across hundreds to thousands of different jurisdictions and has often fallen on the shoulders of litigators and judges to be accomplished through judicial scrutiny in case specific contexts [6, 229]. Shortly after the release of the NRC Report, the Hon. Harry Edwards, co-chair of the NRC Committee on Identifying the Needs of the Forensic Sciences Community, indicated that this was not an effective solution, stating:

> *Unfortunately, the adversarial approach to the submission of evidence in court is not well suited to the task of finding "scientific truth." The judicial system is encumbered by, among other things, judges, lawyers, and jurors who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner; defense attorneys who often do not have the resources to challenge*

[131] E.g., such activities might include conducting the testing of those technologies and algorithms or providing guidance for how such testing should be performed to promote and enable these activities to be conducted at scale across various entities. Consideration could be given to developing guidance and tools to enable tailoring the AI Risk Management Framework to algorithms developed for forensic science applications, irrespective whether they are based on AI/ML methods or not as the principles outlined are relevant for all types of algorithms used in forensic science for court purposes (e.g., [185 p2]). Finally, such activities might also include consideration of providing a central source and means for disseminating the results of such testing throughout the forensic science community to enable greater access to the information.

> *prosecutors' forensic experts; trial judges (sitting alone) who must decide evidentiary issues without the benefit of judicial colleagues and often with little time for extensive research and reflection; and very limited appellate review of trial court rulings admitting disputed forensic evidence. … Furthermore, the judicial system embodies a case-by-case adjudicatory approach that is not well suited to address systematic problems in many of the various forensic science disciplines. … [Consequently,] I am of the view that judicial review, by itself, will not cure the infirmities of the forensic science community … [and b]ecause of the many problems presently faced by the forensic science community and the inherent limitations of the judicial system, the forensic science community as it is now constituted cannot consistently serve the judicial system as well as it might. … [L]awyers and judges should not be counted on to fix the science problem. What we need is for the forensic science community to improve so that it better serves the needs of justice.* [7 pp11-12]

Over the years, the forensic science community has worked to self-regulate through the development of standards and guidelines, accreditation programs, certification programs, and proficiency testing programs, all of which are foundational to a quality management system [6, 229]. However, in the United States, there are no requirements for compliance or systematic means of enforcing conformance, which has raised questions as to their impact and effectiveness [6, 7, 229].[132]

6.2.1. Standards and Guidelines

In 2009, the NRC highlighted the role that standards and guidelines play in facilitating and promoting oversight and accountability of forensic science practices, noting:

> *Standards provide the foundation against which performance, reliability, and validity can be assessed. Adherence to standards reduces bias, improves consistency, and enhances the validity and reliability of results. Standards reduce variability resulting from the idiosyncratic tendencies of the individual examiner—for example, setting conditions under which one can declare a "match" in forensic identifications. They make it possible to replicate and empirically test procedures and help disentangle method errors from practitioner errors. Importantly, standards not only guide practice but also can serve as guideposts in accreditation and certification programs. Many forensic science disciplines have developed standards, but others have not, which contributes to questions about the validity of conclusions.* [6 p201]

In the years leading up to the 2009 NRC report, standards and guidelines (i.e., guidance documents) relating specifically to forensic science practices[133] were produced primarily by: the

[132] This is in contrast to the UK Forensic Science Regulator, for example, which, as of 2021 has enforcement powers [303].
[133] The International Organization for Standardization (ISO) has also published standards (jointly with the International Electrotechnical Commission, IEC) (ISO/IEC 17025 and ISO/IEC 17020) that are commonly used by forensic service providers for conformity assessment; however, they are not forensic or discipline-specific [239, 240].

FBI,[134] NIST,[135] ASTM International,[136] and various Scientific Working Groups (SWGs) [6].[137] Although the FBI, NIST, and ASTM International have produced guidance documents, the SWGs have been the most prolific in the number of disciplines covered and quantity of guidance documents produced [6, 143, 229]. While the NRC recognized "the SWGs have been a source of improved standards for the forensic science disciplines and represent the results of a profession that is working to strengthen its professional services with only limited resources," the NRC also noted "some standards and guidelines lack the level of specificity needed to ensure consistency … [and] their voluntary nature and inconsistent application make it difficult to assess their impact" [6 pp202-203]. The NRC noted:

> *Ideally, standards should be consistently applicable and measurable. In addition, mechanisms should be in place for their enforcement, with sanctions imposed against those who fail to comply. As such, standards should be developed with a consideration of the relevant measures that will be used to provide a meaningful evaluation of an organization's or individual's level of compliance. Appropriate standards must be coupled with effective systems of accreditation and/or certification that include strong enforcement mechanisms and sanctions.* [6 pp203-206]

Shortly after the publication of the NRC report, Hon. Harry Edwards commented further about the concerns identified by the Committee, noting:

> *[S]ome SWGs undoubtedly incorporate good technical protocols that should enhance forensic science analyses; however, … as a general matter, SWGs are of questionable value. Why? Because:*
>
> - *SWG committees meet irregularly and have no clear or regular sources of funding.*
> - *There are no clear standards in place to determine who gains membership on SWG committees.*
> - *Neither SWGs nor their recommendations are mandated by any federal or state law or regulation.*
> - *SWG recommendations are not enforceable.*
> - *A number of SWG guidelines are too general and vague to be of any great practical use.*

---

134 The FBI Quality Assurance Standards (QAS) apply only to DNA laboratories performing DNA testing or utilizing the Combined DNA Index System (CODIS). The FBI is authorized by the DNA Identification Act of 1994 [304] to issue quality assurance standards governing forensic DNA testing laboratories and to establish and administer the National DNA Index System (NDIS). Conformity assessment to the FBI QAS is provided by third-party conformity assessment bodies (e.g., [305]).

135 The NRC noted that NIST "conducts research to establish [physical] standards in a limited number of forensic areas, for example, organic gunshot residue analysis, trace explosives detectors, and improvised explosive devices … [and t]hey also develop guides to help forensic organizations formulate appropriate policies and procedures, such as those concerning mobile phone forensic examinations" [6 p201].

136 ASTM International is an international standards development organization that produces and publishes voluntary technical standards for a wide range of materials, products, systems, and services [306]. In the area of forensic science, it offers limited standards relating to a variety of forensic disciplines, including but not limited to, forensic documents examination, trace evidence examination, drug chemistry, explosives analysis, and fire debris analysis [306].

137 The SWGs were first initiated by the FBI in the late 1980s and early 1990s "to facilitate consensus around forensic science operations among federal, state, and local agencies" [6 p202].

- *SWG committees have no way of knowing whether state or local agencies even endorse the standards.*
- *Complaints are not filed when a practitioner violates a SWG standard.*
- *SWG committees do not attempt to measure the impact of their standards by formal study or survey.*

> *In other words, even if it is true that some SWG standards make sense and might result in good practice, there is nothing to indicate that the standards are routinely followed and enforced in a way to ensure best practices in the forensic science community.* [7 pp10-11]

In 2013, many SWGs ceased operations[138] and, in 2014, OSAC was established through a cooperative agreement between NIST and DOJ [22, 230]. OSAC provided funding, structure, and coordination around the process of developing discipline-specific standards and guidelines [23]. Although the mission of OSAC has received widespread support across the forensic science community [231], questions have been raised regarding its overall effectiveness [5, 13, 229]. For example, in 2016, the PCAST noted "[t]he creation by NIST of OSAC was an important step in strengthening forensic science practice[; however,] … initial lessons from its first years of operation have revealed some important shortcomings" [13 p126]. The PCAST continued "OSAC's membership includes relatively few independent scientists: it is dominated by forensic professionals, who make up more than two-thirds of its members" [13 p126]. Consequently, the "PCAST concludes that OSAC lacks sufficient independent scientific expertise and oversight to overcome the serious flaws in forensic science … [and s]ome restructuring is necessary to ensure that independent scientists and statisticians have a greater voice in the standards development process, a requirement for meaningful scientific validity" [13 p126].

Looking at OSAC from a sociological perspective of standards development [232], similar concerns were raised in 2018, noting that the purpose of standards, the composition of committees developing the standards, and the implementation of standards can all impact their effectiveness [229]. "The goals of standardization can range from aspirational to reflecting the status quo: 'standards can imply a lowest common denominator of available options, the power of the strongest party in standardization, a negotiated order among some or all stakeholders, or a confirmation of how things are done by most parties' … [and a] key question that has emerged already is whether the OSAC standards should be aspirational—should articulate where we want forensic science eventually to be—or should reflect the status quo" [229 p577]. Additionally, consensus-based committees, such as OSAC, can "lead[] to compromises, bitterly contested power plays, and negotiations, … [which] has already emerged as an issue for OSAC … [by] heavily favor[ing] practitioners who may be oriented toward the aforementioned status

[138] SWGDAM and SWGDE remained. A summary of the history of the SWGs and actions that occurred immediately following the 2009 NRC report are provided in [143, 230].

quo" [229 p577].[139] Finally, "[r]egardless of the composition of committees and the content of standards, their enforcement and adoption is far from assured[; rather,] … the voluntary nature of many standards makes it difficult to develop momentum unless built-in incentives promote compliance, [such as an] … auxiliary system that provides internal or external incentives, audits, and certification" [229 p580].

In 2020, concerns were raised regarding OSAC standards that were similar to those raised in 2009 by the NRC in relation to the SWGs [233]. Specifically, commentators claimed "that there are instances in which its [OSAC] standards-development process is being subverted … [and] some standards initially developed by OSAC are detrimental to the goal of improving the scientific validity of forensic practice" [233 p206]. They claimed that OSAC standards were "vacuous" in that "[t]hey usually state few requirements"; "[t]heir stated requirements are often vague"; "[c]ompliance with their stated requirements can be achieved with little effort—the bar is set very low"; and "[c]ompliance with their stated requirements would not be sufficient to lead to scientifically valid results" [233 p206]. They claimed the OSAC standards "appear to be designed to allow laboratories and practitioners to continue with existing poor practice, and if challenged to be able to respond that they are following established standards" [233 p207]. They continued, suggesting that "[t]here is danger, however, that a court may not look further than the fact that a standard exists, and be misled into believing that conformity to a vacuous standard is indicative of scientific validity, even though it is not[; therefore, f]or this reason, it would be better to have no standard than to have a vacuous standard" [233 p207].

Responses to these concerns highlighted the nature of the voluntary consensus process, pointing to structure and process of standards development, noting that "[c]onsensus, achieved through due process, is the very heart of the voluntary standards process … [and c]onsensus allows for critical thinking and implementation to be evaluated in a structured domain that includes public review[; however, r]arely does the content of a standard meet the expectations of all the members of a community" [234 pp1-2].

In addition to the creation of OSAC, a global effort to develop ISO standards for forensic science has been initiated through ISO Technical Committee (ISO/TC) 272 [235]. It was initially created in 2012 with work commencing in 2013 as a Project Committee responsible for developing the standard ISO/IEC 18385:2016 Minimizing the Risk of Human DNA Contamination in Products Used to Collect, Store and Analyze Biological Material for Forensic Purposes—Requirements [236]. In December 2015, the committee was converted to a Technical Committee responsible for developing a broader range of forensic standards, which include "laboratory and field-based forensic science techniques and methodology in broad general areas such as the detection and collection of physical evidence, the subsequent analysis and interpretation of the evidence, and the reporting of results and findings" [235]. As of 2023, the

[139] Although Cole highlighted the question of whether OSAC standards ought to be aspirational or reinforce status-quo [229], others, looking specifically at standards developed through government recognized Standards Development Organizations (SDOs), have argued that SDO standards "are practical and do not set aspirational targets; they are not meant to articulate a gold standard, or a minimum standard, but rather a benchmark that balances community expectations against what can realistically be delivered (taking into account technology, market forces, safety etc.)" [307 p2].

ISO/TC 272 is in the process of developing and publishing several standards under the ISO 21043 series [237].[140] The five parts of the ISO 21043 series of standards include: terms and definitions (Part I); recognition, recording, collecting, transport and storage of items (Part 2); analysis (Part 3); interpretation (Part 4); and reporting (Part 5) [237].

6.2.2. Accreditation, Certification, and Proficiency Testing

In 2009, the NRC noted "standards are used to measure the quality of institutions or organizations, either in terms of their policies and procedures or in terms of the proficiency and skills of an individual practicing the discipline … [and] standards are enforced through systems of accreditation and certification, wherein independent examiners and auditors test and audit the performance, policies, and procedures of both laboratories and service providers" [6 p194]. The NRC continued "[t]his cannot be a self-assessing program"; rather, "[o]versight must come from outside the participating laboratory to ensure that standards are not self-serving and superficial and to remove the option of taking shortcuts when other demands compete with quality assurance" [6 p195]. Although accreditation and certification are important mechanisms for demonstrating conformance to established standards of practice, the NRC found that "with the exception of three states mandating accreditation (New York, Oklahoma, and Texas), the accreditation of laboratories and certification of forensic examiners remains voluntary" [6 p194]. Finally, the NRC concluded "[s]tandards should reflect best practices and serve as accreditation tools for laboratories and as guides for the education, training, and certification of professionals" and "[l]aboratory accreditation and individual certification of forensic science professionals should be mandatory" [6 pp214-215].

6.2.2.1. Accreditation

Accreditation "primarily addresses the management system, technical methods, and quality of the work of a laboratory" [6 p208]. Although laboratory accreditation is voluntary, in 2009, approximately 83% of publicly funded crime laboratories were accredited by a professional forensic science organization (i.e., accreditation body) [36 p1] and by 2014, approximately 88% were accredited [238 p1].[141] Among those accredited in 2014, the majority (approximately 83%) were accredited by either the American Society of Crime Laboratory Directors / Laboratory Accreditation Board (ASCLD/LAB) (73%) or Forensic Quality Services-International (FQS-International) (10%) [238 p1]. In 2009, laboratories were primarily accredited for demonstrating conformance to ISO/IEC 17025 [239]. However, in 2011, ISO/IEC 17020 [240] was offered as an optional alternative to ISO/IEC 17025 for accreditation among many of the pattern evidence disciplines that rely on visual inspection as the basis for their examination and interpretation (e.g., friction ridge, firearms and toolmarks, scene investigation)

[140] A full list of standards published and under development by ISO/TC 272 is provided in [237]. A detailed review of international development of forensic science standards, including activity relating to ISO/TC 272, is provided in [307].

[141] These data, however, exclude forensic service providers that do not work in traditional crime laboratory settings. In 2009, the NRC noted a survey of International Association for Identification (IAI) members, who tend to work in settings other than traditional crime laboratories, revealed that only 15% of respondents were accredited [6 p199]. This is contrasted with 83% reflected by BJS in the same year (2009) [36].

[241]. In 2016, ASCLD/LAB and FQS-International merged under ANSI National Accreditation Board [242, 243]. Collectively, ANSI National Accreditation Board and the American Association for Laboratory Accreditation (A2LA) represent the two largest accrediting bodies offering forensic accreditation to either ISO/IEC 17025 or ISO/IEC 17020 in the United States [238 pp2-3, 244].[142] In 2020, approximately 90% of publicly funded crime laboratories were accredited in at least one discipline [38 p16]. Almost all state labs (98%) had some form of accreditation, as did 94% of federal labs, 86% of county labs, and 82% of municipal labs [38 p16].

In April 2015, the NCFS recommended universal accreditation [245], and in November 2015, the U.S. Attorney General accepted those recommendations, issuing a directive across the DOJ mandating that by December 2020, all DOJ forensic service providers must obtain or maintain accreditation and that DOJ attorneys must use accredited forensic service providers when requesting testing or evidence [246].[143] Further, the Attorney General also called for the DOJ to redraft its grant solicitations to provide incentives for state, local, and tribal forensic testing entities to apply for and use discretionary funding to seek and maintain accreditation [246]. In 2017, the NCFS extended the recommendation for accreditation to digital and multimedia evidence forensic service providers [247]; however, the Attorney General did not explicitly act on those recommendations.

Achieving universal accreditation on a voluntary basis (for non-DOJ laboratories) has been challenging [42]. In 2019, NIJ engaged with the forensic science community and noted "some agencies may not feel that accreditation is the highest priority or a reasonable investment if not mandatory or incentivized" and "securing funding sources for accreditation can be a challenge, and some agencies may not want to divert time and workforce efforts from casework" [42 p66]. Also in 2019, in response to concerns over barriers to accreditation, ASCLD launched the ASCLD Accreditation Initiative (AAI) in collaboration with the NIJ FTCoE at RTI International, which was "designed to leverage expertise and knowledge from laboratory directors, quality assurance managers, and technical subject matter experts who serve as mentors to support [forensic service providers] committed to achieving international accreditation within an 18-month timeline" [244 p1]. In 2022, RTI International published a report of the impact and lessons learned from the AAI to date, noting that the AAI was successful, and in the three-year timeframe since its inception had assisted eight forensic service providers to become accredited to either ISO/IEC 17025 or ISO/IEC 17020 [244 pp1, 6].

While achieving accreditation is an important component in a laboratory's quality assurance program, accreditation status alone does not necessarily ensure that discipline-specific practices are scientifically valid and reliable [13, 229]. Accreditation has been argued to be "insufficiently rigorous" [229 p567] and the PCAST claimed that accreditation alone "cannot substitute for empirical evidence of scientific validity and reliability" [13 p66]. Neither ISO/IEC 17025 nor ISO/IEC 17020 provide discipline-specific standards relating to forensic science examination, interpretation, and reporting [239, 240].[144] Additionally, the 2009 NRC report

[142] For a brief history of forensic laboratory accreditation in the years prior to 2009, see [6 pp197-200].
[143] Excluding digital analysis [246].
[144] The purpose of establishing OSAC was "to address a lack of discipline-specific forensic science standards" [23].

highlighted deficiencies in the validity and reliability of several discipline-specific practices employed by many forensic service providers [6] despite over 80% of publicly funded crime laboratories claiming to have been accredited [36]. In 2016, similar concerns relating to the validity and reliability of discipline-specific practices were asserted by the PCAST [13] even though nearly 90% of publicly funded crime laboratories claimed to have been accredited [238].[145] Although accrediting bodies have supplemental requirements specific to their forensic accreditation programs [248, 249], they have not fully adopted requirements or enforced the use of discipline-specific standards, such as those developed or recognized by OSAC [169].

6.2.2.2. Certification

In 2009, the NRC noted that "certification of individuals complements the accreditation of laboratories for a total quality assurance program … [and] is specifically designed to ensure the competency of the individual examiner" [6 p208]. In their report, the NRC raised similar concerns regarding certification as they did with accreditation, stating "[i]n other realms of science and technology, professionals, including nurses, physicians, professional engineers, and some laboratorians, typically must be certified before they can practice"; and "[t]he same should be true for forensic scientists who practice and testify" [6 p208]. However, the NRC highlighted that, despite entities offering certification for some disciplines, "[c]ertification, while broadly accepted by the forensic science community, is not uniformly offered or required" [6 p214]. The NRC concluded, stating "[n]o person (public or private) should be allowed to practice in a forensic science discipline or testify as a forensic science professional without certification" [6 p215]. Despite its voluntary nature, by 2014, approximately 72% of publicly funded crime laboratories employed at least one externally certified analyst, up from 60% in 2009 [238 p5].[146] In 2016, the NCFS also called for widespread certification, stating:

> *Professional certification is the recognition by an independent certification body that an individual has acquired and demonstrated specialized knowledge, skills, and abilities in the standard practices necessary to execute the duties of his or her profession. Certification also provides the general public and the judicial system with a means of identifying those practitioners who have successfully demonstrated compliance with established requirements. … Requiring [forensic service providers] to mandate the certification of their forensic science practitioners would improve the quality of services provided and enhance confidence in the judicial system. … Finally, certification provides another means of external oversight for practitioners.* [250][147]

Alongside calls for mandatory certification were discussions surrounding the challenges of implementing it [42, 251]. For example, a survey of forensic practitioners on their views of

[145] While ISO/IEC 17025 and ISO/IEC 17020 require methods to be validated prior to use, in their reports, both the NRC and the PCAST claimed that adequate studies demonstrating the foundational validity of some methods has yet to be conducted [6, 13].
[146] These data, however, exclude forensic service providers that do not work in traditional crime laboratory settings. Data relating to certification were not available in BJS data for 2020.
[147] For a list of forensic certification bodies and the various categories of testing and disciplines that certification is offered, see Appendix A in [250].

certification reported in 2019 revealed mixed results [251]. Generally, respondents "supported mandatory certification and agreed that requiring universal mandatory certification was a positive thing" [251 p166]. However, "[t]hose not in support of mandatory certification generally reported it was not necessary or had concerns with the certification examination itself … [such that] certification did not actually qualify someone as competent in their field and that mandatory certification could result in overlooking quality candidates" [251 p166]. "A majority of [respondents] (52%) also reported that a major reason individuals do not want to take the exam is the fear of failing the certification examination" and "the fear of appearing less competent in court if attorneys bring up prior failed certification exams" [251 p166]. The survey results also found that "cost for both certification and recertification factors into the decision not to become certified … [and] most laboratories or organizations do not have the means to pay for analysts to become certified, leaving the practitioner to pay not only for the initial examination seating but also travel expenses to the examination site … [which] can particularly affect practitioners in the early part of their career where pay may be low and large monthly student loan payments are often an issue" [251 pp166-167].

Also in 2019, similar findings relating to logistical burdens of obtaining certification were reported by NIJ following their engagement with the practitioner community [42]. While NIJ recognized the value of certification in advancing national standards, they also noted that "[a]s positions requiring these certifications become more common, there will be funding needs associated with gaining and maintaining certification, including costs of examinations, costs and time for continuing education and professional development, and consideration for higher salaries for certified personnel" [42 p33]. NIJ continued "[o]verall, [participants] identified the primary barriers to certification to be a lack of funding and the inability to remove staff from casework during the time required to become certified" [42 p33]. Further, NIJ noted "[s]ome agencies cannot afford to send their examiners to … training [required for certification], and in smaller agencies training is done in-house" [42 p33]. NIJ concluded "[a]s the field continues to develop and focus on standardization and professionalization, certification may be required by more agencies for certain positions, meaning that the challenges [relating to] training and education … will likely be amplified" [42 p34].[148]

6.2.2.3. Proficiency Testing

Proficiency testing provides a means of monitoring laboratory performance on a continual basis [252]. Proficiency testing is often included in performance monitoring requirements as part of laboratory accreditation schemes [239, 240, 248, 249]. As noted by the NCFS:

> *Proficiency testing is intended as an evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons for the determination of*

[148] Similar considerations were highlighted by the NCFS, noting that "[e]ducational programs or preparatory courses should be developed to help practitioners prepare for certification examinations" and "[b]udgetary constraints may impact the ability to obtain and maintain certification" (Appendix D in [250]).

> *service provider performance. Proficiency testing is commonly used by [forensic service providers] management to evaluate staff, training, and method validation; appropriateness of test methods; traceability of measurements and calibrations to national standards; calibration and maintenance of test equipment; documentation, sampling, and handling of test items; and quality assurance of data, including reporting of results. In forensic science, proficiency testing is used not only as a measure of the [Forensic Science Service Provider's] overall performance and quality system (e.g., facility, equipment, procedures, and training programs) but also as a tool for monitoring an individual [Forensic Science Practitioner's] continued ability to perform work in a specific discipline or tasks. The use of proficiency testing to evaluate individual examiners' continuing ability to perform specific tasks should not be confused with competency testing … [which] is the demonstration that a[] [Forensic Science Practitioner] has acquired and demonstrated specialized knowledge, skills, and abilities in the standard practices necessary to conduct examinations in a discipline and/or category of testing prior to performing independent casework.* [252]

In 2009, the NRC recognized the importance of proficiency testing; however, they also commented that "[a]lthough many forensic science disciplines have engaged in proficiency testing for the past several decades, several courts have noted that proficiency testing in some disciplines is not sufficiently rigorous" [6 p206]. The NRC continued, noting that although blind proficiency testing is not required, "proficiency testing should include blind testing" [6 p207]. This recommendation from the NRC for blind proficiency testing is not new—in 1992, the NRC Committee on DNA Technology in Forensic Science made similar comments, claiming:

> *Most importantly, there is no substitute for rigorous external proficiency testing via blind trials. Such proficiency testing constitutes scientific confirmation that a laboratory's implementation of a method is valid not only in theory, but also in practice. No laboratory should let its results with a new DNA typing method be used in court, unless it has undergone such proficiency testing via blind trials.* [10 p55]

According to the BJS, the percentage of laboratories that report participating in proficiency testing has remained consistent over the years at over 95% since 2009 (97% in 2009 and 98% in both 2014 and 2020) [38 p16, 238 p4].[149] The number of laboratories participating in blind proficiency testing, however, is lower, with the BJS noting the percentage of laboratories conducting blind proficiency tests being approximately 10% for both 2009 and 2014 [238 p4] and 11% in 2020 [38 p16]. Additionally, the BJS found that in 2014, "federal crime labs (39%) were more likely than county (8%), state (7%), and municipal (5%) labs to test the proficiency of employees through blind examinations" [238 p4].[150] Although the BJS reported approximately 98% of publicly funded crime laboratories participating in proficiency testing in 2020, they also noted that fewer (approximately 87%) participated in competency

[149] These data exclude forensic service providers that do not work in traditional crime laboratory settings.
[150] Similar data for 2020 are available; however, they must be "interpret[ed] with caution" because the "[e]stimate is based on 10 or fewer sample cases, or coefficient of variation is greater than 50%" [38 p17].

testing "to evaluate the knowledge and abilities of their analysts or examiners before they perform independent forensic casework" [38 pp16-17].

In 2016, the NCFS made similar comments as the NRC, stating "[i]t is important for proficiency tests to be sufficiently rigorous and representative of the challenges of forensic casework" and recommended universal participation in proficiency testing programs [253]. Also in 2016, the PCAST reinforced the importance of proficiency testing and pointed to it as a key component for demonstrating the validity "as applied" [13]. Specifically, the PCAST claimed: "the *only* way to establish scientifically that an examiner is capable of applying a foundationally valid method is through appropriate empirical testing to measure how often the examiner gets the correct answer" and "[s]uch empirical testing is often referred to as 'proficiency testing'" [13 p57]. The PCAST continued, noting "[p]roficiency testing should be performed under conditions that are representative of casework and on samples, for which the true answer is known, that are representative of the full range of sample types and quality likely to be encountered in casework in the intended application" [13 p57]. Reinforcing the concerns raised by the NRC and NCFS, a common finding by the PCAST across multiple disciplines was the need for proficiency tests to improve in their overall rigor [13].[151]

The PCAST also suggested that "[t]o ensure integrity, proficiency testing should be overseen by a disinterested third party that has no institutional or financial incentive to skew performance" [13 p57]. As part of this assertion, the PCAST highlighted comments made to the NCFS in 2015 in that commercial proficiency test providers often weigh test difficulty with commercial demands, noting that "'easy tests are favored by the community,' with the result that tests that are too challenging could jeopardize repeat business for a commercial vendor" [13 p68].[152] Critiques of the relative ease of commercial proficiency tests have increased in recent years, with particular emphasis on the friction ridge discipline [254–256]. For example, in 2018, one study reported that "the quality levels of latent fingerprints from proficiency tests are generally higher quality, less complex, and do not represent the quality levels observed in routine casework" [254 p379]. Then, in 2019, another article reported that three public defenders with no formal training in friction ridge examination took the test and passed, concluding "participating in [current proficiency] tests should be considered little more than window dressing" [255 p294]. Finally, in 2020, another study found "[t]he low observed error rate, examiner perceptions of relative ease, and high objective print quality metrics together suggest that latent print proficiency testing is not especially challenging" [256 p120].

Ultimately, the PCAST offered a similar recommendation as the NRC and pointed to blind testing as a path forward, noting "proficiency testing should *ideally* be conducted in a 'test-blind' manner—that is, with samples inserted into the flow of casework such that examiners do not know that they are being tested" [13 p58]. Although the PCAST recognized

[151] See specifically pages 75, 102, 113, 149 in reference [13].

[152] The PCAST cited Collaborative Testing Services, Inc. (CTS) President Christopher Czyryca. Czyryca noted that the forensic community disfavors more challenging tests—and that testing companies are concerned that they could lose business if their tests are viewed as too challenging (see pages 57 and 68 in [13]). In a presentation to the NCFS, Czyryca also stated that "[e]asy tests are favored by the community" [308].

that "there is disagreement in the forensic community about its feasibility in all settings … [and] blinded, inter-laboratory proficiency tests may be difficult to design and orchestrate on a large scale[,] … PCAST believes that test-blind proficiency testing of forensic examiners should be vigorously pursued, with the expectation that it should be in wide use, at least in large laboratories, within the next five years" [13 pp58-59]. Although the challenges associated with implementing blind proficiency testing schemes have been recognized, the benefits of doing so have continued to be argued in recent years [257, 258], and some laboratories have demonstrated its feasibility with successful implementation [259–261]. Efforts to promote broad-based adoption of blind proficiency testing have also been undertaken, with laboratory directors and quality managers convening to discuss obstacles to the adoption of blind testing and to assess successful and potential strategies to overcome them [262].

6.2.3. Boards and Commissions

In 2009, the NRC highlighted that in the United States, there is no centralized body that has the authority or capacity to govern or provide coherent advisement or oversight of forensic service providers [6]. Consequently, a principal recommendation by the NRC was a call for Congress to develop NIFS as an independent federal entity that could focus on: "establishing and enforcing best practices for forensic science professionals and laboratories" and "establishing standards for the mandatory accreditation of forensic science laboratories and the mandatory certification of forensic scientists" [6 p81]. In making such a recommendation, however, the NRC recognized that "[m]ost forensic science methods, programs, and evidence are within the regulatory province of state and local law enforcement entities or are covered by statutes and rules governing state judicial proceedings"; thus, "Congress cannot directly fix all the deficiencies in the forensic science community" [6 p13]. The NRC continued, noting "Congress does not have free reign to amend state criminal codes, rules of evidence, and statutes governing civil actions; nor may it easily and directly regulate local law enforcement practices, state and local medical examiner units, or state policies covering the accreditation of crime laboratories and the certification of forensic practitioners" [6 p13]. At the same time, the NRC also noted, that "oversight and enforcement of operating standards, certification, accreditation, and ethics are lacking in most local and state jurisdictions" [6 p23]. Thus, the NRC concluded that "[i]n the end, … state and local authorities must be willing to enforce change if it is going to happen" [6 p14].

At the time the NRC report was released in 2009, only three states required accreditation (New York, Oklahoma, Texas) [6 p193]. Although the principal recommendation to establish NIFS never came to fruition, greater attention was directed toward the establishment of federal and state forensic science boards and commissions to provide advisory and oversight functions to forensic science laboratories [148, 263, 264]. At the federal level, the NCFS was established in 2013 as a federal advisory committee to the United States Attorney General on forensic science policies and practices and co-chaired by NIST and the DOJ [148]. During its tenure, the NCFS adopted a total of 43 work products: 20 Recommendation documents and 23 Views documents on a range of issues relating to forensic science research, policy, and practice, including endorsements of universal accreditation, certification, and proficiency testing across DOJ forensic science laboratories [149]. In 2017, however, the NCFS charter ended and was not renewed, causing the NCFS to terminate [149]—a decision that was met with criticism and viewed as premature given the challenges continuing to face the forensic science community [265]. At the state and local levels, the establishment of boards and commissions has increased in recent years [263, 264].

In 2016, the NIJ FTCoE first published a landscape review of forensic science state boards and commissions, with the intent of serving as a resource "for states wishing to create and maintain a state forensic science commission" [263 p1]. The report noted that, as of 2016, "10 states and the District of Columbia have a legislatively created commission to provide support, guidance or oversight to state and local crime laboratories"; these states are Arkansas, Delaware, Maryland, Missouri, New York, North Carolina, Texas, Virginia, Rhode Island, and Washington [263 p7]. However, the report also noted that those commissions "vary considerably" [263 p1] with respect to the functions they fulfill and "[t]he responsibilities and duties … include a wide range of activities" [263 p7]. As of 2016, at least six states required accreditation (Maryland, Missouri, New York, North Carolina, Oklahoma,[153] and Texas)[154] and two states required individual certification or licensure (North Carolina, Texas) [263 pp29-40].

In 2022, the NIJ FTCoE published an update to the 2016 landscape review of forensic science state boards and commissions [264], which "include[ed] state forensic science commissions, task forces, oversight and advisory boards, or investigative councils that may improve the field of forensic sciences through oversight and coordination of forensic science resources" [264 p6]. The report noted that, as of 2022, "[t]wenty-one states and Washington, D.C. currently have statutorily created or created by another means forensic science state commissions or oversight bodies" [264 p7]. Specifically, those with bodies created by statute included: Alabama, Arkansas, Delaware, Illinois, Indiana, Maryland, Massachusetts, Michigan, Missouri, Montana, New York, North Carolina, Oklahoma, Rhode Island, Texas, Virginia, Washington, and Washington, D.C. [264 p10]. States with bodies created by other means included: Arizona, Kansas, Nebraska, and Wisconsin [264 p10]. Although "many states have established forensic science oversight bodies," they "take many forms and have myriad roles and responsibilities" [264 p8]. As of 2022, at least seven states required accreditation (Maryland, Massachusetts, Missouri, New York, North Carolina, Oklahoma,[155] and Texas)[156] and two states required individual certification or licensure (North Carolina, Texas) [264 pp31-54]. Although the number of state commissions and oversight bodies doubled between 2016 and

[153] Oklahoma does not have a commission per se; rather, Oklahoma has a Forensic Sciences Improvement Task Force hosted by the Oklahoma District Attorneys Council. Oklahoma requires the accreditation of public crime laboratories through statute 74 OK Stat. § 74-150.37 (2014) [309]. The report notes: "This statute exempts alcohol/breath, CSI, digital, crime scene reconstruction, marijuana and latent print analysis. For latent print analysis to be admitted into evidence, it must be conducted by an International Association for Identification (IAI)-certified examiner" [263 p36].

[154] Although not listed in reference [263], Washington, D.C. has required accreditation since the establishment of the Department of Forensic Sciences in 2012 through D.C. Code § 5–1501.06 [310].

[155] "As of 2005, Oklahoma requires the accreditation of public crime laboratories: 74 OK Stat § 74-150.36 (2020). This statute exempts alcohol/breath, CSI, digital, crime scene reconstruction, marijuana, and latent print analysis. For latent print analysis to be admitted into evidence, it must be conducted by an International Association for Identification (IAI)-certified examiner" [264 p49].

[156] Although not listed in reference [264], Washington, D.C. has required accreditation since the establishment of the Department of Forensic Sciences in 2012 through D.C. Code § 5–1501.06 [310].

2022, only one additional state mandated accreditation (Massachusetts) [263, 264]. Most bodies focused on functions relating to advising on key personnel qualifications, providing guidance on programs and protocols, or reviewing complaints [264 pp31-54].

In 2023, the National Association of Forensic Science Boards (NAFSB) was established as "a grass-roots initiative to ensure that State-level forensic science boards are best positioned to benefit forensic science" by providing "a forum for these groups to communicate and share their different experiences and identify best practices that are most applicable to each state's unique circumstances" [266]. The inaugural conference of the NAFSB was hosted by the Texas Forensic Science Commission in Austin, TX in November 2023 [267].

6.2.4. Considerations for Forensic Science Research and Standards Programs

Over the years, efforts to strengthen the oversight and accountability of forensic science practices has been made. This is often accomplished by promoting more effective and comprehensive quality management systems focused on the use of standards and guidelines and participation in accreditation programs, certification programs, and proficiency testing programs. The extent to which a forensic service provider employs rigorous principles of quality assurance and conforms to a robust and comprehensive quality management system is a decision that rests solely on the shoulders of the forensic service provider. That decision is outside the purview of forensic science research and standards programs, which instead must ensure that the components of a comprehensive quality management system (e.g., standards, accreditation, certification, proficiency testing) are available and effective. Accordingly, there are three major considerations for forensic science research and standards programs (Box 6.2.4).

Box 6.2.4. Forensic science research and standards programs should:

(a) Strengthen the rigor of discipline-specific standards and guidelines recommended for use by the forensic science community.

(b) Provide a means for forensic service providers to demonstrate conformance to discipline-specific standards and guidelines (or for external bodies with advisory or oversight responsibilities and authorities to enforce compliance) through external audits conducted by conformity assessment bodies as part of an accreditation program.[157]

(c) Promote and support rigorous proficiency testing schemes, including efforts by forensic service providers to institute blind testing.

Discipline-specific standards and guidelines recommended for the forensic science community should be developed with the values of inclusivity, transparency, and objectivity (among others) central to the process and written with sufficient detail, clarity, and specificity to ensure they can be applied in a reliable and reproducible manner so that, in turn, they can more effectively promote better quality, consistency and standardization across the forensic science community. While consensus is central to the standards development process, attention should be directed toward "ensur[ing] independent scientists and statisticians have a greater voice in the standards development process" [13 p126] so that the final consensus is more defensible and robust.

Discipline-specific standards and guidelines are only effective if they are used. Thus, supporting the development of auditing schemes to assess conformance will not only lay a foundation for forensic service providers to provide objective evidence of their adherence to discipline-specific standards and guidelines but also provide a more effective means of promoting and enabling greater consistency and standardization of discipline-specific practices across the forensic science community.

Support for blind testing could take the form of expanding interlaboratory testing schemes to cover a broad range of disciplines relevant to a majority of forensic service providers throughout the United States; providing recommendations and guidelines on designing, executing, and interpreting the results from proficiency testing schemes that could be applied by other entities (including scientifically sound methods for determining the assigned values or expected results from examinations of test specimens); or providing assistance with the analysis of test specimen for use in proficiency testing exercises. Such

[157] Conformity assessment is traditionally carried out by accrediting bodies independent of the forensic service provider. Non-government accrediting bodies provide fee-based conformity assessment services to forensic service providers that voluntarily choose to use those services to demonstrate conformance to national or international standards, such as ISO/IEC 17025:2017 and ISO/IEC 17020:2012; however, those assessments do not necessarily require conformance to discipline-specific forensic science standards such as those recognized by the OSAC Registry [169]. Instead, forensic service providers provide self-declarations of conformance to discipline-specific standards and guidelines [273].

support would not only be responsive to the recommendations from the NRC,[158] but also help address concerns raised by the PCAST, claiming "[t]o ensure integrity, proficiency testing should be overseen by a disinterested third party that has no institutional or financial incentive to skew performance" [13 p57].

[158] E.g., to "develop tools for advancing measurement, validation, reliability, information sharing, and proficiency testing in forensic science and to establish protocols for forensic examinations, methods, and practices" [6 p214].

7. List of Considerations for Forensic Science Research and Standards Programs

This list of considerations for forensic science research and standards programs correspond to those identified across the various landscapes in this report. The parenthetical alphanumeric identifier next to each implication provides the reference to the corresponding section in this report. Although some redundancies exist between the considerations, the frame of reference differs as a function of the landscape to which they correspond. This scan was developed through the lens of the NIST Forensic Science Program; however, the considerations are not intended to be specific to any particular forensic science research or standards program.

Forensic Science Research and Standards Programs should:

1. Ensure transparency and balance across diverse perspectives when administering programmatic activities so that the outputs align to the needs of all members of the forensic science community and are not unduly influenced (real or perceived) by any individual groups (researchers, practitioners, investigators, prosecutors, defense attorneys, and judges) (2.1.1.a).

2. Take a leading role and affirmative stance on matters relating to good scientific principles and practices—even in the face of adversarial viewpoints (2.1.1.b).

3. Ensure program outputs are both applicable and accessible to all forensic service providers despite differences in resources and account for those resource limitations when carrying out their missions (2.2.1.a).

4. Provide coordinated mechanisms to promote sound scientific practices that define how forensic science activities should be practiced by forensic service providers throughout the nation and consider ways to help enable and promote mechanisms to ensure forensic service providers conform to those practices (2.2.1.b).

5. Take a neutral but affirmative stance on science-based issues affecting the use of forensic science in the criminal justice system, such as standards, conformity assessment, quality management systems, and accreditation (2.3.1.a).

6. Ensure more guidance is made available and accessible to members of the forensic science and criminal justice communities on issues relating to scientific validity and reliability and how they can be properly assessed using science-based standards, guidelines, and scientific information concerning forensic science practice (2.3.1.b).

7. Prioritize efforts that are most relevant, practical, and impactful to addressing the day-to-day challenges faced by the broader forensic science community (3.1.1.a).

8. Strengthen partnerships and collaborations among government, academia, professional organizations, and private industry to support forensic service providers in efforts to translate and implement new or improved analytical methods and technologies into practice (3.1.1.b).

9. Assess the strengths and limitations of their existing research and development infrastructure, including their competencies, capabilities, and available resources to identify the most meaningful ways they can contribute to the broader forensic science research and development ecosystem (e.g., balancing their efforts across a spectrum of basic and applied research, short-term and long-term priorities, intramural and extramural funding, and support for method development and translation, including commercialization) (3.2.1.a).

10. Coordinate across different research and standards programs to establish a shared strategy for addressing the broader research and development needs of the forensic science community and, where relevant, align their respective resources toward those shared priorities and objectives (3.2.1.b).

11. Ensure that investments in forensic science research and development include support for translation and implementation of their outputs so that they can be impactful and yield the greatest practical returns for the forensic science community and criminal justice system (3.2.1.c).

12. Ensure transparency, objectivity, neutrality, and impartiality when administering programmatic activities, such that the public can build and maintain trust that their scientists are credible, and that their outputs are not unduly influenced (real or perceived) by individual interests or agendas (4.1.4.a).

13. Take an active and assertive stance on matters relating to good scientific principles or practices—even when issues are controversial, or viewpoints are adversarial (4.1.4.b).

14. Speak with a consistent voice on matters related to science and standards and be able to communicate and engage with all members of the criminal justice system and the public on how to recognize sound scientific and technical practices (4.1.4.c).

15. Establish ways for forensic service providers to demonstrate their adherence to sound scientific and technical practices and conformance to recognized standards, guidelines, and recommendations through impartial and independent third-party entities rather than relying solely on self-declarations (4.1.4.d).

16. Ensure scientific evaluations of forensic methods and practices are conducted in a consistent and systematic manner (4.2.1.a).

17. Approach scientific evaluations of forensic methods and practices in a manner that has a clearly defined scope that is broad enough to produce meaningful information but narrow enough to ensure it can be completed within a timeframe that is acceptable to the forensic science community given the resources available (4.2.1.b).

18. evaluations of forensic methods and practices not only focus on the capabilities and limitations to accurately and reliably answer specific and clearly defined forensic questions, but also include a discussion of whether there are other forensic questions for which the forensic methods or practices could produce accurate and reliable answers (4.2.1.c).

19. Have well-defined research plans that include contributions to the development of actual measures of accuracy and reliability of different forensic methods and practices to promote and enable better calibration of public perception (4.2.1.d).

20. Prioritize outreach, using various approaches to reach broad and diverse audiences, to inform the public about the results of scientific evaluations and results of research that will contribute to calibrating what the public perception of the accuracy and reliability of forensic science methods should be, which is grounded by actual measures of accuracy and reliability, to ensure the public is not misled and unduly over- or under-values the evidence presented in the context of a particular case (4.2.1.e).

21. Provide explicit criteria on how scientific validity and reliability ought to be assessed and the basis for such determinations (5.1.1.a).

22. Develop guidelines and tools for the forensic science community that describe how research that is intended to measure validity and reliability of forensic science methods should be designed and executed (5.1.1.b).

23. Develop guidelines and tools that enable the forensic science community to assess the validity and reliability of a particular method as applied in the context of a particular case (5.1.1.c).

24. Create discipline-specific research roadmaps that clearly prioritize the research needs and objectives that most directly align to addressing knowledge gaps related to the validity and reliability of commonly practiced forensic science methods (5.2.4.a).

25. Assess their internal capabilities and competencies to identify what science and technology challenges and needs are within scope for intramural or extramural research and consider optimal ways of structuring and administering programmatic activities to ensure they have the capacity to support unanticipated or time-sensitive projects (5.2.4.b).

26. Establish strategic relationships through interagency agreements (IAAs) and cooperative research and development agreements (CRADAs), as appropriate, to provide the administrative underpinnings and enable efficient coordination, collaboration, and sharing of research and development activities, data, technologies, and related materials between other major or strategically significant government, academic, or industry partners (5.2.4.c).

27. Create a clear translation pipeline from research to practice that defines the categories, phases, and stages of research as they progress from ideation to implementation to impact (5.3.3.a).

28. Prioritize investments that support the translation and implementation of outputs from their programmatic activities into practice, such as:

    a. establishing strategic collaborations and partnerships with forensic science laboratories, end-users, and other government and academic researchers to ensure research and development activities are coordinated, tailored to the specific needs of the forensic science community, and have a clear pathway for the translation and implementation of the outputs by early-adopters that can further assess the methods or technologies in operational contexts and champion the adoption across the community;
    b. providing a centralized means of sharing resources, such as data, datasets, validation guidelines and reports, protocols, and training materials relating to the performance characterization and validation of methods and technologies having broad applicability across the forensic science community;
    c. supporting a centralized means of conducting research, development, testing and evaluation activities relating to the performance characterization and validation of methods and technologies having broad applicability across the forensic science community, including the creation of model protocols and training materials that can be leveraged by the forensic science community (5.3.3.b).

29. Establish strategic outreach plans that prioritize dissemination of their research and development activities, outputs, and key findings using various approaches to reach broad and diverse audiences (5.3.3.c).

30. Identify an appropriate set of procedures and metrics to regularly measure outcomes and impacts of their programmatic activities resulting from the translation and implementation of their outputs (5.3.3.d).

31. Be cognizant of the various legal rules and standards governing the admissibility of scientific evidence and their implications when developing new methods, technologies, and tools intended to be used by the forensic science community (6.1.4.a).

32. Prioritize activities that promote greater objectivity, validity, reliability, and consistency to forensic science practices as well as evaluate the extent to which those outcomes are achieved (6.1.4.b).

33. Prioritize activities that promote more efficient and effective testing of technologies and computational algorithms proposed for court use as it relates to assessing their validity, reliability, and fairness (6.1.4.c).

34. Prioritize activities that provide guidance to officers of the court on issues relating to scientific principles of validity, reliability, and fairness (6.1.4.d).

35. Strengthen the rigor of discipline-specific standards and guidelines recommended for use by the forensic science community (6.2.4.a).

36. Provide a means for forensic service providers to demonstrate conformance to discipline-specific standards and guidelines (or for external bodies with advisory or oversight responsibilities and authorities to enforce compliance) through external audits conducted by conformity assessment bodies as part of an accreditation program (6.2.4.b).

37. Promote and support rigorous proficiency testing schemes, including efforts by forensic service providers to institute blind testing (6.2.4.c).

8. References

[1] Eurostat Quality of Life Indicators - Measuring Quality of Life. Available at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Quality_of_life_indicators_-_measuring_quality_of_life
[2] President's Management Agenda U.S. Social Indicators. Available at https://www.performance.gov/explore/social-indicators/
[3] National Institute of Standards and Technology Forensic Science Program. Available at https://www.nist.gov/spo/forensic-science-program
[4] Kirk PL (1953) *Crime Investigation: Physical Evidence and the Police Laboratory* (Interscience New York).
[5] Koehler JJ, Mnookin JL, Saks MJ (2023) The Scientific Reinvention of Forensic Science. *Proceedings of the National Academy of Sciences of the United States of America* 120(41):e2301840120. Available at https://doi.org/10.1073/pnas.2301840120
[6] National Research Council Committee on Identifying the Needs of the Forensic Sciences Community (2009) *Strengthening Forensic Science in the United States: A Path Forward* (The National Academies Press, Washington, D.C. USA). https://doi.org/10.17226/12589
[7] Edwards HT (2009) Solving The Problems that Plague the Forensic Science Community. *Jurimetrics* 50:5-19.
[8] Ribeiro G, Tangen JM, McKimmie BM (2019) Beliefs About Error Rates and Human Judgment in Forensic Science. *Forensic Science International* 297:138-147.
[9] National Research Council Committee on DNA Forensic Science (1996) *The Evaluation of Forensic DNA Evidence* (National Academies Press).
[10] National Research Council Committee on DNA Technology in Forensic Science (1992) *DNA Technology in Forensic Science* (National Academies Press).
[11] President's DNA Initiative: Advancing Justice Through DNA Technology Act of 2003
[12] Edwards HT (2010) The National Academy of Sciences Report on Forensic Sciences: What it Means for the Bench and Bar. *Jurimetrics* 51(1):1-15.
[13] President's Council of Advisors on Science and Technology (2016) *Report to the President, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-comparison Methods* (U.S. Executive Office of the President, Washington, D.C., USA).
[14] National Association of Criminal Defense Lawyers (2016) President's Council of Advisors on Science and Technology (PCAST) Issues Major Forensic Science Report; Calls for Stronger Scientific Standards. Available at https://www.nacdl.org/newsrelease/NewsRelease-09-20-2016
[15] Innocence Project (2016) Innocence Project Applauds President Obama's Science Advisors' Landmark Report Calling for Essential Improvements to Forensic Disciplines. Available at https://innocenceproject.org/president-obamas-report-forensic-science/
[16] Ramos M (2016) Reference: Report Entitled "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods." *National District Attorneys Association* (National District Attorneys Association). Available at https://ndaa.org/wp-content/uploads/NDAA-PCAST-Response-FINAL.pdf

[17] Statement on September 20, 2016 PCAST Report on Forensic Science (2016) (American Society of Crime Laboratory Directors). Available at https://pceinc.org/wp-content/uploads/2016/10/20160930-Statement-on-PCAST-Report-ASCLD.pdf

[18] Comments on: President's Council of Advisors on Science and Technology REPORT TO THE PRESIDENT Forensic Science in Federal Criminal Courts: Ensuring Scientific Validity of Pattern Comparison Methods (2016) (Federal Bureau of Investigation). Available at https://www.fbi.gov/file-repository/fbi-pcast-response.pdf/view

[19] President's Council of Advisors on Science and Technology (PCAST) Report (2016) Available at https://web.archive.org/web/20221203061458/https://news.aafs.org/policy-statements/presidents-council-of-advisors-on-science-and-technology-pcast-report/

[20] United States Department of Justice Statement on the PCAST Report: 'Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods,' (2021) (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/olp/page/file/1352496/download

[21] National Academies Press A Timeline of Notable Events for Strengthening Forensic Science in the United States. Available at https://nap.nationalacademies.org/resource/12589/interactive/

[22] Memorandum of Understanding Between the Department of Justice and the National Institute of Standards and Technology in Support of the National Commission on Forensic Science and the Organization of Scientific Area Committees Available at https://www.justice.gov/archives/ncfs/file/761051/download

[23] National Institute of Standards and Technology Organization of Scientific Area Committees for Forensic Science (OSAC). Available at https://www.nist.gov/osac

[24] Center for Statistics and Applications in Forensic Evidence (CSAFE) Available at https://forensicstats.org/

[25] Butler J, Iyer H, Press R, Taylor M, Vallone P, Willis S (2018) NIST Scientific Foundation Reviews (NISTIR 8225). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8225

[26] Kulsuski M (2022) An Easy Win for Criminal Justice Reform: Independent Crime Labs. *The Hill*. Available at https://thehill.com/opinion/criminal-justice/3511295-an-easy-win-for-criminal-justice-reform-independent-crime-labs/

[27] American Society of Crime Laboratory Directors (2022) Parent Organizations of Forensic Laboratories. Available at https://www.ascld.org/wp-content/uploads/2022/11/2022_ASCLD-Position-Statement-on-Parent-Organizations_11_08_2022.pdf

[28] Swofford H, Champod C (2022) Probabilistic Reporting and Algorithms in Forensic Science: Stakeholder Perspectives within the American Criminal Justice System. *Forensic Sci Int: Synergy* 4(100220):1-25.

[29] Rakoff JS, Liu G (2023) Forensic Science: A Judicial Perspective. *Proceedings of the National Academy of Sciences of the United States of America* 120(41):e2301838120. Available at https://doi.org/10.1073/pnas.2301838120

[30] Frye v. United States (293 F. 1013 (D.C. Cir. 1923)).

[31] Daubert v. Merrell Dow Pharmaceuticals, Inc. (509 U.S. 579 (1993)).

[32] Berry K (2015) *How Judicial Elections Impact Criminal Cases* (Brennan Center for Justice New York, NY). Available at https://www.brennancenter.org/sites/default/files/publications/How_Judicial_Elections_Impact_Criminal_Cases.pdf
[33] Bannon A (2016) Rethinking Judicial Selection. *Prof Law* 24:1. Available at https://www.americanbar.org/groups/professional_responsibility/publications/professional_lawyer/2016/volume-24-number-1/rethinking_judicial_selection/
[34] Giannelli PC (2017) Forensic Science: Daubert's Failure. *Case W Res L Rev* 68:869-933. Available at https://ssrn.com/abstract=3031227
[35] National Commission on Forensic Science (2016) Recommendation to the Attorney General: Technical Merit Evaluation of Forensic Science Methods and Practices. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/page/file/905541/download
[36] Bureau of Justice Statistics (2012) Census of Publicly Funded Forensic Crime Laboratories, 2009. (U.S. Department of Justice, Washington, D.C.). Available at https://bjs.ojp.gov/content/pub/pdf/cpffcl09.pdf
[37] Bureau of Justice Statistics (2016) Publicly Funded Forensic Crime Laboratories: Resources and Services, 2014. (U.S. Department of Justice, Washington, D.C.). Available at https://bjs.ojp.gov/library/publications/publicly-funded-forensic-crime-laboratories-resources-and-services-2014
[38] Bureau of Justice Statistics (2023) Census of Publicly Funded Forensic Crime Laboratories, 2020. (U.S. Department of Justice, Washington, D.C.).
[39] Bureau of Labor Statistics Consumer Price Index (CPI), CPI Inflation Calculator. Available at https://www.bls.gov/data/inflation_calculator.htm
[40] Idaho State Police Planning Grants and Research (2008) Forensic Funding Alternatives: Criminal Justice Opinion Survey. (State of Idaho). Available at https://web.archive.org/web/20210425144715/https://www.jrsa.org/pubs/sac-digest/documents/idaho-forsensicsurvey.pdf
[41] Speaker P (2022) Project FORESIGHT Annual Report, 2020-2021. *Faculty & Staff Scholarship* 3093. Available at https://researchrepository.wvu.edu/faculty_publications/3093
[42] National Institute of Justice (2019) Report to Congress: Needs Assessment of Forensic Laboratories and Medical Examiner/Coroner Offices. (U.S. Department of Justice, Washington, D.C.). Available at https://www.ncjrs.gov/pdffiles1/nij/253626.pdf
[43] Garrett BL (2020) The Costs and Benefits of Forensics. *Hous L Rev* 57(3):593-616. Available at https://houstonlawreview.org/article/12197-the-costs-and-benefits-of-forensics
[44] Transcript, Presidential Commission on the Administration of Justice Conference (2020) Available at https://www.justice.gov/ag/page/file/1270041/download
[45] Gamette M (2020) Improving Forensic Science Integration: A Director's Perspective. *Forensic Sci Int: Synergy* 2:183-186. Available at https://doi.org/10.1016/j.fsisyn.2020.05.005
[46] Houston Forensic Science Center Impacts of Covid-19 on the Houston Forensic Science Center, Report to the Texas House Jurisprudence Committee. Available at

https://capitol.texas.gov/tlodocs/86R/handouts/C3302020110600001/d01f5c4b-1093-4642-81e1-eedd9b3d3f20.PDF

[47] Esposito M, Salerno M, Scoto E, Di Nunno N, Sessa F (2022) The Impact of the COVID-19 Pandemic on the Practice of Forensic Medicine: An Overview. *Healthcare* 10(2):319-335. Available at https://doi.org/10.3390/healthcare10020319

[48] Boring S (2023) Backlog, Staffing Shortages, Space All Causing Issues at Sedgwick County Forensic Science Center. *KSN News*. Available at https://www.ksn.com/news/local/backlog-staffing-shortages-space-all-causing-issues-at-sedgwick-county-forensic-science-center/

[49] Bureau of Labor Statistics Occupational Outlook Handbook, Forensic Science Technicians, Job Outlook 2021 - 2031. (U.S. Department of Labor , Washington, D.C.). Available at https://www.bls.gov/ooh/life-physical-and-social-science/forensic-science-technicians.htm#tab-6

[50] Wickenheiser RA, Farrell L (2020) Collaborative Versus Traditional Method Validation Approach: Discussion and Business Case. *Forensic Sci Int: Synergy* 2:230-237.

[51] Wickenheiser RA (2021) Reimagining Forensic Science–The Mission of the Forensic Laboratory. *Forensic Sci Int: Synergy* 3(100153):1-6.

[52] Cavanagh R (2019) Notes from the NIST Research Innovation to Implementation in Forensic Science Symposium (RI2I). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.SP.2100-02

[53] National Institute of Justice (2021) The Impact of NIJ Forensic Science Research and Development. (U.S. Department of Justice, Washington, D.C.). Available at https://www.ojp.gov/pdffiles1/nij/300422.pdf

[54] National Institute of Justice Research and Development in Forensic Science for Criminal Justice Purposes, Fiscal Year 2021. Available at https://nij.ojp.gov/funding/opportunities/o-nij-2021-95001

[55] National Institute of Justice Research and Development in Forensic Science for Criminal Justice Purposes, Fiscal Year 2022. Available at https://nij.ojp.gov/funding/opportunities/o-nij-2022-171186

[56] National Research Council Committee on Strengthening Forensic Science at the National Institute of Justice (2015) *Support for Forensic Science Research: Improving the Scientific Role of the National Institute of Justice* (National Academies Press, Washington, D.C.).

[57] Forensic Technology Center of Excellence Available at https://forensiccoe.org/

[58] Dear Colleague Letter: Forensic Science—Opportunity for Breakthroughs in Fundamental and Basic Research and Education (2013) (National Science Foundation, Washington, D.C.). Available at http://www.nsf.gov/pubs/2013/nsf13120/nsf13120.jsp

[59] Dear Colleague Letter—I/UCRCs in Areas Relevant to the Forensic Sciences (2014) (National Science Foundation, Washington, D.C.). Available at https://www.nsf.gov/pubs/2014/nsf14066/nsf14066.pdf

[60] Forensic Technology Center of Excellence (2016) Federal Investment in Forensic Science Research and Development: Landscape Study. (RTI International, Research Triangle Park, NC). Available at https://forensiccoe.org/private/6548099c28d60

[61] Forensic Technology Center of Excellence (2021) Federal Investment in Forensic Science Research and Development: Landscape Study--A 2021 Update. (RTI International, Research Triangle Park, NC). Available at https://forensiccoe.org/private/65481f55c36ab

[62] Sherman LW (2002) Trust and Confidence in Criminal Justice. *National Institute of Justice Journal* 248:22-31. Available at https://www.ojp.gov/sites/g/files/xyckuh241/files/archives/ncjrs/189106-1.pdf?q=

[63] The Gallup Organization (2022) Confidence in Institutions. (Washington, D.C.). Available at https://news.gallup.com/poll/1597/confidence-institutions.aspx

[64] Pew Research Center (2022) Americans' Views of Government: Decades of Distrust, Enduring Support for Its Role. Available at https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2022/06/PP_2022.06.06_views-of-government_REPORT.pdf

[65] The Gallup Organization (2022) Trust in Government. (Washington, D.C.). Available at https://news.gallup.com/poll/5392/trust-government.aspx

[66] Edelman Trust Barometer 2022 Global Report (2022) Available at https://www.edelman.com/sites/g/files/aatuss191/files/2022-01/2022 Edelman Trust Barometer FINAL_Jan25.pdf

[67] Forensic Science Standards Board of the Organization of Scientific Area Committees for Forensic Science OSAC Preferred Terms. Available at https://www.nist.gov/organization-scientific-area-committees-forensic-science/osac-lexicon

[68] American Academy of Forensic Sciences What is Forensic Science. Available at https://www.aafs.org/careers-forensic-science/what-forensic-science

[69] National Opinion Research Center (2021) Confidence in Scientific Community: 1972 – 2021. *General Social Survey*. Available at https://gssdataexplorer.norc.org/trends?category=Politics&measure=consci

[70] Kennedy B, Tyson A, Funk C (2022) Americans' Trust in Scientists, Other Groups Declines. *Pew Research Center*. Available at https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines

[71] Wellcome Global Monitor: How does the World Feel About Science and Health? (2018) (Gallup). Available at https://wellcome.org/sites/default/files/wellcome-global-monitor-2018.pdf

[72] National Opinion Research Center (2013) General Social Survey Final Report: Trends in Public Attitudes about Confidence in Institutions. Available at https://www.norc.org/PDFs/GSS Reports/Trends in Confidence Institutions_Final.pdf

[73] National Academies of Sciences Engineering and Medicine (2015) *Trust and Confidence at the Interfaces of the Life Sciences and Society: Does the Public Trust Science? A Workshop Summary* (National Academies Press, Washington, D.C.). Available at https://nap.nationalacademies.org/catalog/21798/trust-and-confidence-at-the-interfaces-of-the-life-sciences-and-society

[74] National Academies of Sciences Engineering and Medicine (2017) Examining the Mistrust of Science: Proceedings of a Workshop–in Brief. Available at https://nap.nationalacademies.org/catalog/24819/examining-the-mistrust-of-science-proceedings-of-a-workshop-in

[75] Cross C (2021) Will Public Trust in Science Survive the Pandemic? *Chemical and Engineering News* 99(3). Available at https://cen.acs.org/policy/global-health/Will-public-trust-in-science-survive-the-pandemic/99/i3
[76] National Institute of Justice (2013) Race, Trust and Police Legitimacy. Available at https://nij.ojp.gov/topics/articles/race-trust-and-police-legitimacy
[77] Rosenbaum D (2014) Building Trust Inside and Out: The Challenge of Legitimacy Facing Police Leaders. *Research for the Real World Seminar* (National Institute of Justice). Available at https://nij.ojp.gov/library/publications/building-trust-inside-and-out-challenges-facing-police-leaders-research-real
[78] National Opinion Research Center (2021) National Spending on Law Enforcement: 1985 – 2021. *General Social Survey*. Available at https://gssdataexplorer.norc.org/trends?category=Current%20Affairs&measure=natcrimy
[79] Parker K, Hurst K (2021) Growing Share of Americans Say They Want More Spending on Police in Their Area. *Pew Research Center*.
[80] Jones J (2020) Black, White Adults' Confidence Diverges Most on Police. *The Gallup Organization*. Available at https://news.gallup.com/poll/317114/black-white-adults-confidence-diverges-police.aspx
[81] Jones J (2021) In U.S., Black Confidence in Police Recovers From 2020 Low. *The Gallup Organization*. Available at https://news.gallup.com/poll/352304/black-confidence-police-recovers-2020-low.aspx
[82] Bureau of Justice Statistics (2016) Body-Worn Cameras in Law Enforcement Agencies. (U.S. Department of Justice, Washington, D.C.). Available at https://bjs.ojp.gov/content/pub/pdf/bwclea16.pdf
[83] Executive Office of the President (2022) Executive Order on Advancing Effective, Accountable Policing and Criminal Justice Practices to Enhance Public Trust and Public Safety. (Washington, D.C.). Available at https://www.whitehouse.gov/briefing-room/presidential-actions/2022/05/25/executive-order-on-advancing-effective-accountable-policing-and-criminal-justice-practices-to-enhance-public-trust-and-public-safety
[84] Evans H, Greene K, Healy W, Hoffman E, Rimmer K, Sberegaeva A, Zimmerman N (2020) National Institute of Standards and Technology Environmental Scan 2020 (NISTIR 8348). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8348
[85] Ling S, Kaplan J, Berryessa CM (2021) The Importance of Forensic Evidence for Decisions on Criminal Guilt. *Science & Justice* 61(2):142-149.
[86] Peterson JL, Hickman MJ, Strom KJ, Johnson DJ (2013) Effect of Forensic Evidence on Criminal Justice Case Processing. *Journal of Forensic Sciences* 58:S78-S90. Available at https://doi.org/10.1111/1556-4029.12020
[87] Kaplan J, Ling S, Cuellar M (2020) Public Beliefs About the Accuracy and Importance of Forensic Evidence in the United States. *Science & Justice* 60(3):263-272.
[88] Chin JM, Ibaviosa CM (2022) Beyond CSI: Calibrating Public Beliefs About the Reliability of Forensic Science Through Openness and Transparency. *Science & Justice* 62(3):272-283.

[89] Garrett BL, Neufeld PJ (2009) Invalid Forensic Science Testimony and Wrongful Convictions. *Virginia Law Review*:1-97.

[90] Mnookin JL, Cole SA, Dror IE, Fisher BAJ (2011) The Need for a Research Culture in the Forensic Sciences. *UCLA L Rev* 58(3):725-779.

[91] Jones J (2012) Forensic Tools: What's Reliable and What's Not-So-Scientific. *Frontline, Public Broadcasting Station (PBS)*. Available at https://www.pbs.org/wgbh/frontline/article/forensic-tools-whats-reliable-and-whats-not-so-scientific/

[92] Edmond G (2015) Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation. *Melbourne University Law Review* 39(1):77-127.

[93] Mnookin JL (2018) The Uncertain Future of Forensic Science. *Daedalus* 147(4):99-118.

[94] Koehler JJ (2018) How Trial Judges Should Think About Forensic Science Evidence. *Judicature* 102(1):28-38.

[95] Sikorski M (2022) Is Forensic Science in Crisis? *Synthese* 200(3):188-222.

[96] Martire KA, Ballantyne KN, Bali A, Edmond G, Kemp RI, Found B (2019) Forensic Science Evidence: Naive Estimates of False Positive Error Rates and Reliability. *Forensic Science International* 302(109877):1-9.

[97] Redmayne M, Roberts P, Aitken C, Jackson G (2011) Forensic Science Evidence in Question. *Criminal Law Review* 5:347-356.

[98] Lieberman JD, Carrell CA, Miethe TD, Krauss DA (2008) Gold Versus Platinum: Do Jurors Recognize the Superiority and Limitations of DNA Evidence Compared to Other Types of Forensic Evidence? *Psychology, Public Policy, and Law* 14(1):27-62.

[99] Garrett BL, Mitchell G (2016) Forensics and Fallibility: Comparing the Views of Lawyers and Jurors. *W Va L Rev* 119(2):621-637.

[100] Murrie DC, Gardner BO, Kelley S, Dror IE (2019) Perceptions and Estimates of Error Rates in Forensic Science: A Survey of Forensic Analysts. *Forensic Science International* 302(109887):1-9.

[101] Edwards HT (2014) Reflections on the Findings of the National Academy of Sciences Committee on Identifying the Needs of the Forensic Science Community. *First Public Meeting of the National Commission on Forensic Science, Office of Justice Programs* Available at https://www.justice.gov/sites/default/files/ncfs/legacy/2014/05/13/harry-edwards.pdf

[102] Ulery BT, Hicklin RA, Buscaglia J, Roberts MA (2011) Accuracy and Reliability of Forensic Latent Fingerprint Decisions. *Proceedings of the National Academy of Sciences of the United States of America* 108(19):7733-7738.

[103] Pacheco I, Cerchiai B, Stoiloff S (2014) Miami-Dade Research Study for the Reliability of the ACE-V Process: Accuracy & Precision in Latent Fingerprint Examinations (Final Report for NIJ Award 2010-DN-BX-K268). (National Institute of Justice). Available at http://www.ncjrs.gov/pdffiles1/nij/grants/248534.pdf

[104] Baldwin DP, Bajic SJ, Morris M, Zamzow D (2014) A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons. *Ames Laboratory, U.S. Department of Energy* Available at https://www.ojp.gov/pdffiles1/nij/249874.pdf

[105] Ulery BT, Hicklin RA, Buscaglia J, Roberts MA (2012) Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners. *PloS ONE* 7(3):e32800.

[106] Ulery BT, Hicklin RA, Kiebuzinski GI, Roberts MA, Buscaglia J (2013) Understanding the Sufficiency of Information for Latent Fingerprint Value Determinations. *Forensic Science International* 230(1):99-106. Available at https://doi.org/10.1016/j.forsciint.2013.01.012
[107] Ulery BT, Hicklin RA, Roberts MA, Buscaglia J (2014) Measuring What Latent Fingerprint Examiners Consider Sufficient Information for Individualization Determinations. *PloS one* 9(11):e110179. Available at https://doi.org/10.1371/journal.pone.0110179
[108] Ulery BT, Hicklin RA, Roberts MA, Buscaglia J (2015) Changes in Latent Fingerprint Examiners' Markup Between Analysis and Comparison. *Forensic Science International* 247:54-61. Available at https://doi.org/10.1016/j.forsciint.2014.11.021
[109] White House Office of Science and Technology Policy (2016) Published Statements in Response to the PCAST Report on Forensic Science in Criminal Courts. Available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_2016_public_comments.pdf
[110] Speir JA, Richetelli N, Hammer L (2020) Forensic Footwear Reliability: Part I—Participant Demographics and Examiner Agreement. *Journal of Forensic Sciences* 65(6):1852-1870.
[111] Richetelli N, Hammer L, Speir JA (2020) Forensic Footwear Reliability: Part II—Range of Conclusions, Accuracy, and Consensus. *Journal of Forensic Sciences* 65(6):1871-1882.
[112] Richetelli N, Hammer L, Speir JA (2020) Forensic Footwear Reliability: Part III—Positive Predictive Value, Error Rates, and Inter-Rater Reliability. *Journal of Forensic Sciences* 65(6):1883-1893.
[113] Hicklin RA, McVicker BC, Parks C, LeMay J, Richetelli N, Smith M, Buscaglia J, Perlman RS, Peters EM, Eckenrode BA (2022) Accuracy, Reproducibility, and Repeatability of Forensic Footwear Examiner Decisions. *Forensic Science International* 339(111418):1-16. Available at https://doi.org/10.1016/j.forsciint.2022.111418
[114] Eldridge H, De Donno M, Champod C (2021) Testing the Accuracy and Reliability of Palmar Friction Ridge Comparisons–A Black Box Study. *Forensic Science International* 318(110457):1-20.
[115] Hicklin RA, Winer KR, Kish PE, Parks CL, Chapman W, Dunagan K, Richetelli N, Epstein EG, Ausdemore MA, Busey TA (2021) Accuracy and Reproducibility of Conclusions by Forensic Bloodstain Pattern Analysts. *Forensic Science International* 325(110856):1-8.
[116] Chapnick C, Weller TJ, Duez P, Meschke E, Marshall J, Lilien R (2021) Results of the 3D Virtual Comparison Microscopy Error Rate (VCMER) Study for Firearm Forensics. *Journal of Forensic Sciences* 66(2):557-570.
[117] Monson KL, Smith ED, Bajic SJ (2022) Planning, Design and Logistics of a Decision Analysis Study: The FBI/Ames Study Involving Forensic Firearms Examiners. *Forensic Sci Int: Synergy* 4(100221):1-14. Available at https://doi.org/10.1016/j.fsisyn.2022.100221
[118] Hicklin RA, Eisenhart L, Richetelli N, Miller MD, Belcastro P, Burkes TM, Parks CL, Smith MA, Buscaglia J, Peters EM (2022) Accuracy and Reliability of Forensic Handwriting Comparisons. *Proceedings of the National Academy of Sciences of the United States of America* 119(32):e2119944119. Available at https://doi.org/10.1073/pnas.2119944119
[119] Hicklin RA, Richetelli N, Emerick BL, Bever RA, Davoren JM (2023) Variation in Assessments of Suitability and Number of Contributors for DNA Mixtures. *Forensic Science International: Genetics* 65(102892):1-14. Available at https://doi.org/10.1016/j.fsigen.2023.102892

[120] Guttman B, Laamanen MT, Russell C, Darnell J, Atha C (2022) Results From a Black-Box Study for Digital Examiners (NISTIR 8412). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8412
[121] Phillips PJ, Yates AN, Hu Y, Hahn CA, Noyes E, Jackson K, Cavazos JG, Jeckeln G, Ranjan R, Sankaranarayanan S (2018) Face Recognition Accuracy of Forensic Examiners, Superrecognizers, and Face Recognition Algorithms. *Proceedings of the National Academy of Sciences of the United States of America* 115(24):6171-6176. Available at https://www.pnas.org/doi/epdf/10.1073/pnas.1721355115
[122] Koch SL, Liebowitz C, Shriver MD, Jablonski NG (2021) Microscopical Discrimination of Human Head Hairs Sharing a Mitochondrial Haplogroup. *Journal of Forensic Sciences* 66(1):56-71.
[123] Martinez-Lopez C, Trejos T, Coulson S, Goodpaster J, Igowsky K, Kuczelinis F, Mehltretter A, Pollock E, Simmross U, Weimer R (2019) Interlaboratory Evaluations of the Performance of Elemental Analytical Methods for the Forensic Analysis and Comparisons of Electrical Tapes. *Forensic Chemistry* 12:66-77. Available at https://doi.org/10.1016/j.forc.2019.01.001
[124] Prusinowski M, Brooks E, Neumann C, Trejos T (2023) Forensic Interlaboratory Evaluations of a Systematic Method for Examining, Documenting, and Interpreting Duct Tape Physical Fits. *Forensic Chemistry* 34(100487):1-13. Available at https://doi.org/10.1016/j.forc.2023.100487
[125] Lambert K, Montero S, Akmeemana A, Corzo R, Gordon G, Haase E, Jiang P, Ovide O, Prasch K, Redman K (2022) An Interlaboratory Study to Evaluate the Forensic Analysis and Interpretation of Glass Evidence. *Forensic Chemistry* 27(100378):1-12. Available at https://doi.org/10.1016/j.forc.2021.100378
[126] Hoffman T, Corzo R, Weis P, Pollock E, van Es A, Wiarda W, Stryjnik A, Dorn H, Heydon A, Hoise E (2018) An Inter-Laboratory Evaluation of LA-ICP-MS Analysis of Glass and the Use of a Database for the Interpretation of Glass Evidence. *Forensic Chemistry* 11:65-76. Available at https://doi.org/10.1016/j.forc.2018.10.001
[127] Corzo R, Hoffman T, Ernst T, Trejos T, Berman T, Coulson S, Weis P, Stryjnik A, Dorn H, Workman MS (2021) An Interlaboratory Study Evaluating the Interpretation of Forensic Glass Evidence Using Refractive Index Measurements and Elemental Composition. *Forensic Chemistry* 22(100307):1-10. Available at https://doi.org/10.1016/j.forc.2021.100307
[128] Almirall J, Arkes H, Lentini J, Mowrer F, Pawliszyn J (2017) Forensic Science Assessments: A Quality and Gap Analysis, Fire Investigation. (American Association for the Advancement of Science, Washington, D.C.). Available at https://www.aaas.org/resources/forensic-science-assessments-quality-and-gap-analysis
[129] Thompson W, Black J, Jain A, Kadane J (2017) Forensic Science Assessments: A Quality and Gap Analysis–Latent Fingerprint Examination. (American Association for the Advancement of Science, Washington, D.C.). Available at https://www.aaas.org/resources/forensic-science-assessments-quality-and-gap-analysis
[130] Butler J, Iyer H, Press R, Taylor M, Vallone P, Willis S (2021) DNA Mixture Interpretation: A NIST Scientific Foundation Review (NISTIR 8351-DRAFT). (National Institute of

Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8351-draft

[131] Lyle J, Guttman B, Butler J, Sauerwein K, Reed C, Lloyd C (2022) Digital Investigation Techniques: A NIST Scientific Foundation Review (NISTIR 8354). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8354

[132] Sauerwein K, Butler J, Reczek K, Reed C (2023) Bitemark Analysis: A NIST Scientific Foundation Review (NISTIR 8352). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8352

[133] Dror IE, Langenburg G (2019) “Cannot Decide”: The Fine Line Between Appropriate Inconclusive Determinations Versus Unjustifiably Deciding Not to Decide. *Journal of Forensic Sciences* 64(1):10-15.

[134] Dror IE, Scurich N (2020) (Mis)use of Scientific Measurements in Forensic Science. *Forensic Sci Int: Synergy* 2:333-338. Available at https://doi.org/10.1016/j.fsisyn.2020.08.006

[135] Weller TJ, Morris MD (2020) Commentary on: I. Dror, N Scurich “(Mis)use of Scientific Measurements in Forensic Science” Forensic Science International: Synergy 2020 https://doi. org/10.1016/j.fsisyn. 2020.08.006. *Forensic Sci Int: Synergy* 2:701-702. Available at https://doi.org/10.1016/j.fsisyn.2020.10.004

[136] Hofmann H, Carriquiry A, Vanderplas S (2020) Treatment of Inconclusives in the AFTE Range of Conclusions. *Law, Probability and Risk* 19(3-4):317-364.

[137] Arkes HR, Koehler JJ (2021) Inconclusives and Error Rates in Forensic Science: A Signal Detection Theory Approach. *Law, Probability and Risk* 20(3):153-168.

[138] Biedermann A, Kotsoglou KN (2021) Forensic Science and the Principle of Excluded Middle: “Inconclusive” Decisions and the Structure of Error Rate Studies. *Forensic Sci Int: Synergy* 3(100147):1-11. Available at https://doi.org/10.1016/j.fsisyn.2021.100147

[139] Dorfman AH, Valliant R (2022) Inconclusives, Errors, and Error Rates in Forensic Firearms Analysis: Three Statistical Perspectives. *Forensic Sci Int: Synergy* 5(100273):1-8. Available at https://doi.org/10.1016/j.fsisyn.2022.100273

[140] Khan K, Carriquiry A (2023) Hierarchical Bayesian Non-Response Models for Error Rates in Forensic Black-Box Studies. *Philosophical Transactions of the Royal Society A* 381(20220157):1-23. Available at https://doi.org/10.1098/rsta.2022.0157

[141] Guyll M, Madon S, Yang Y, Burd KA, Wells G (2023) Validity of Forensic Cartridge-Case Comparisons. *Proceedings of the National Academy of Sciences of the United States of America* 120(20):e2210428120.

[142] Strengthening Forensic Science: A Progress Report (2014) (White House Office of Science and Technology Policy, Washington, D.C.). Available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/forensic_science_progress_2-14-14.pdf

[143] Butler JM (2015) U.S. Initiatives to Strengthen Forensic Science & International Standards in Forensic DNA. *Forensic Science International: Genetics* 18:4-20. Available at https://doi.org/10.1016/j.fsigen.2015.06.008

[144] National Science and Technology Council Committee on Science Subcommittee on Forensic Science (2014) Strengthening the Forensic Sciences. (National Science and

Technology Council, Washington, D.C.). Available at https://www.ojp.gov/pdffiles1/NIJ/251422.pdf

[145] Subcommittee on Forensic Science Committee on Science (2015) Achieving Interoperability for Latent Fingerprint Identification in the United States. (U.S. National Science and Technology Council, Washington, D.C.). Available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/latent_fingerprint_report_may_2015.pdf

[146] Expert Working Group on Human Factors in Latent Print Analysis (2012) Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach. (National Institute of Standards and Technology and National Institute of Justice). Available at https://nvlpubs.nist.gov/nistpubs/ir/2012/NIST.IR.7842.pdf

[147] National Commission on Forensic Science National Commission on Forensic Science. Available at https://www.justice.gov/archives/ncfs

[148] National Commission on Forensic Science (2013) Charter. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/file/624216/download

[149] National Commission on Forensic Science (2017) National Commission on Forensic Science: Reflecting Back—Looking Toward the Future. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/page/file/952316/download

[150] Organization of Scientific Area Committees for Forensic Science Research and Development Needs. Available at https://www.nist.gov/organization-scientific-area-committees-forensic-science/osac-research-and-development-needs

[151] Federal Bureau of Investigation Laboratory (2015) FBI Testimony on Microscopic Hair Analysis Contained Errors in At Least 90 Percent of Cases in On-going Review. Available at https://www.fbi.gov/news/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review

[152] ABS Group (2018) Root and Cultural Cause Analysis of Report and Testimony by FBI MHCA Examiners. (ABSG Consulting, Inc. (ABS Group), Washington, D.C.). Available at https://vault.fbi.gov/root-cause-analysis-of-microscopic-hair-comparison-analysis/root-cause-analysis-of-microscopic-hair-comparison-analysis-part-01-of-01/view

[153] Committee on Appropriations (2017) *Departments of Commerce and Justice, Science, and Related Agencies Appropriations Bill, 2018. Report 155-139*. Available at https://www.congress.gov/115/crpt/srpt139/CRPT-115srpt139.pdf

[154] Expert Working Group for Human Factors in Handwriting Examination (2021) Forensic Handwriting Examination and Human Factors: Improving the Practice through a Systems Approach. (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.IR.8282r1

[155] Forensic Technology Center of Excellence (2022) *Human Factors in Forensic Science Practice* eds Eldridge H, Spellman B, Busey T (RTI International, Research Triangle Park, NC). Available at https://www.sciencedirect.com/journal/forensic-science-international-synergy/special-issue/10HK39LR2VS

[156] Spain RD, Hedge JW, Ohse D, White A (2022) The Need for Research-based Tools for Personnel Selection and Assessment in the Forensic Sciences. *Forensic Sci Int: Synergy* 4(100213):1-12. Available at https://doi.org/10.1016/j.fsisyn.2021.100213
[157] Eldridge H, Stimac J, Vanderkolk J (2022) The Benefits of Errors During Training. *Forensic Sci Int: Synergy* 4(100207):1-14. Available at https://doi.org/10.1016/j.fsisyn.2021.100207
[158] National Institute of Justice (2023) Forensic Science Research and Development Technology Working Group: Operational Requirements. (U.S. Department of Justice, Washington, D.C.). Available at https://nij.ojp.gov/topics/articles/forensic-science-research-and-development-technology-working-group-operational#latest
[159] Forensic Research Committee (2022) ASCLD Research Priorities. (American Society of Crime Laboratory Directors). Available at https://www.ascld.org/forensic-research-committee/
[160] Consortium of Forensic Science Organizations (CFSO) (2022) CFSO Research Priorities 2022-2024. (Washington, D.C.). Available at https://thecfso.org/wp-content/uploads/2022/11/CFSO-NIJ-Leave-Behind-Letter-Nov-2022.pdf
[161] National Institute of Forensic Science (NIFS) Australia New Zealand Policing Advisory Agency (ANZPAA) (2020) NIFS Research and Innovation Roadmap 2020-2025. Available at https://www.anzpaa.org.au/forensic-science-2/our-work/projects/research-and-innovation-project
[162] Netherlands Forensic Institute NFI Research and Innovation. Available at https://www.forensicinstitute.nl/research-and-innovation
[163] ENFSI Vision of the European Forensic Science Area 2030 (2021) Improving the Reliability and Validity of Forensic Science and Fostering the Implementation of Emerging Technologies. (European Network of Forensic Science Institutes (ENFSI)). Available at https://enfsi.eu/wp-content/uploads/2021/11/Vision-of-the-European-Forensic-Science-Area-2030.pdf
[164] IFSA Research and Innovation Position Statement (2021) (International Forensic Strategic Alliance (IFSA)). Available at https://www.ifsa-forensics.org/wp-content/uploads/2021/10/IFSA-Research-and-Innovation-Position-Statement.pdf
[165] 17th Interpol International Forensic Science Managers Symposium, Review Papers (2013) (Lyon, France). Available at https://www.interpol.int/content/download/13477/file/IFSMS-Review%20Papers%202013.pdf
[166] 18th Interpol International Forensic Science Managers Symposium, Review Papers (2016) (Lyon, France). Available at https://www.interpol.int/content/download/13472/file/INTERPOL%2018th%20IFSMS%20Review%20Papers-2016.pdf?inLanguage=eng-GB
[167] 19th Interpol International Forensic Science Managers Symposium, Review Papers (2019) (Lyon, France). Available at https://www.interpol.int/content/download/14458/file/Interpol%20Review%20Papers%202019.pdf
[168] 20th Interpol International Forensic Science Managers Symposium, Review Papers (2022) (International Criminal Police Organization (INTERPOL), Lyon, France). Available at

https://www.sciencedirect.com/journal/forensic-science-international-synergy/special-issue/109JJLCFP14
[169] Organization of Scientific Area Committees for Forensic Science OSAC Registry. Available at https://www.nist.gov/organization-scientific-area-committees-forensic-science/osac-registry
[170] President's Commission on Combating Drug Addiction and the Opioid Crisis (2017) The President's Commission on Combating Drug Addiction and the Opioid Crisis: Final Report. (Washington, D.C.). Available at https://trumpwhitehouse.archives.gov/sites/whitehouse.gov/files/images/Final_Report_Draft_11-15-2017.pdf
[171] Matey J, Quinn G, Grother P (2022) Forensic Iris: A Review, 2022 (NIST TN 2226). (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.TN.2226
[172] Rogers EM (1962) *Diffusion of Innovations* (Free Press, New York).
[173] Bauer MS, Kirchner J (2019) Implementation Science: What is it and Why Should I Care? *Psychiatry Research* 283(112376):1-6. Available at https://doi.org/10.1016/j.psychres.2019.04.025
[174] Eccles M, Mittman B (2006) Welcome to Implementation Science. *Implementation Science* 1(1):1-3. Available at https://doi.org/10.1186/1748-5908-1-1
[175] Swofford H, Champod C (2021) Implementation of Algorithms in Pattern & Impression Evidence: A Responsible and Practical Roadmap. *Forensic Sci Int: Synergy* 3(100142):1-18.
[176] Spallek H, Song M, Polk DE, Bekhuis T, Frantsve-Hawley J, Aravamudhan K (2010) Barriers to Implementing Evidence-based Clinical Guidelines: A Survey of Early Adopters. *Journal of Evidence Based Dental Practice* 10(4):195-206. Available at https://doi.org/10.1016/j.jebdp.2010.05.013
[177] Melnyk BM, Fineout-Overholt E (2011) *Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice* (Lippincott Williams & Wilkins, Philadelphia).
[178] Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE (2012) Knowledge Translation of Research Findings. *Implementation Science* 7(50):1-17. Available at https://doi.org/10.1186/1748-5908-7-50
[179] Waltz TJ, Powell BJ, Matthieu MM, Damschroder LJ, Chinman MJ, Smith JL, Proctor EK, Kirchner JE (2015) Use of Concept Mapping to Characterize Relationships Among Implementation Strategies and Assess Their Feasibility and Importance: Results From the Expert Recommendations for Implementing Change (ERIC) Study. *Implementation Science* 10(109):1-8. Available at https://doi.org/10.1186/s13012-015-0295-0
[180] Abu-Odah H, Said NB, Nair SC, Allsop MJ, Currow DC, Salah MS, Hamad BA, Elessi K, Alkhatib A, ElMokhallalati Y (2022) Identifying Barriers and Facilitators of Translating Research Evidence into Clinical Practice: A Systematic Review of Reviews. *Health & Social Care in the Community* 30:e3265-e3276. Available at https://doi.org/10.1111/hsc.13898
[181] Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. *Journal of Experimental Psychology: General* 144(1):114-126.

[182] Dietvorst BJ, Simmons JP, Massey C (2018) Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms if They Can (Even Slightly) Modify Them. *Management Science* 64(3):1155-1170.

[183] Logg JM, Minson JA, Moore DA (2019) Algorithm Appreciation: People Prefer Algorithmic to Human Judgment. *Organizational Behavior and Human Decision Processes* 151:90-103.

[184] Reisman D, Schultz J, Crawford K, Whittaker M (2018) Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability. *AI Now Institute*:1-22. Available at https://ainowinstitute.org/aiareport2018.pdf

[185] National Institute of Standards and Technology (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0). (U.S. Department of Commerce, Gaithersburg, MD). Available at https://doi.org/10.6028/NIST.AI.100-1

[186] Executive Office of the President (2023) Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. (Washington, D.C.). Available at https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

[187] Summary Report of the G7 Multistakeholder Conference on Artificial Intelligence (2018) (Innovation, Science and Economic Development Canada, Montreal, Canada). Available at https://www.ic.gc.ca/eic/site/133.nsf/vwapj/EN_MSC-Final-Report.pdf/$file/EN_MSC-Final-Report.pdf

[188] National Artificial Intelligence Initiative Act of 2020 (P.L. 116-283) (2021) Available at https://www.congress.gov/116/plaws/publ283/PLAW-116publ283.pdf

[189] Dupont B, Stevens Y, Westermann H, Joyce M (2018) Artificial Intelligence in the Context of Crime and Criminal Justice: A Report for the Korean Institute of Criminology, International Centre for Comparative Criminology. (Université de Montréal, Montréal). Available at https://web.archive.org/web/20220401132103/https://www.cicc-iccc.org/public/media/files/prod/publication_files/Artificial-Intelligence-in-the-Context-of-Crime-and-Criminal-Justice_KICICCC_2019.pdf

[190] Toward Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement (2020) (United Nations Interregional Crime and Justice Research Institute (UNICRI) and The International Criminal Police Organization (INTERPOL)). Available at https://unicri.it/towards-responsible-artificial-intelligence-innovation

[191] Artificial Intelligence and Robotics for Law Enforcement (2019) (United Nations Interregional Crime and Justice Institute (UNICRI) and The International Criminal Police Organization (INTERPOL)). Available at https://unicri.it/artificial-intelligence-and-robotics-law-enforcement

[192] Garrett BL, Rudin C (2023) Interpretable Algorithmic Forensics. *Proceedings of the National Academy of Sciences of the United States of America* 120(41):e2301842120. Available at https://doi.org/10.1073/pnas.2301842120

[193] Noiret S, Lumetzberger J, Kampel M (2021) Bias and Fairness in Computer Vision Applications of the Criminal Justice System. *2021 IEEE Symposium Series on*

Computational Intelligence (SSCI) (Institute of Electrical and Electronics Engineers (IEEE)), pp 1-8. Available at https://arxiv.org/pdf/2208.03209.pdf

[194] Piquado T, Matthies CF, Strang L, Anderson JM (2019) Forensic Familial and Moderate Stringency DNA Searches: Policies and Practices in the United States, England, and Wales. (RAND, Santa Monica). Available at https://www.rand.org/pubs/research_reports/RR3209.html

[195] Algorithms Strengthen Forensic Analysis, but Several Factors Can Affect Outcomes (Technology Assessment: Forensic Technology GAO-21-435SP) (2021) (Government Accountability Office, Washington, D.C.). Available at https://www.gao.gov/assets/gao-21-435sp.pdf

[196] Duke University School of Law Court Rules. Available at https://law.duke.edu/lib/research-guides/court-rules

[197] Fed. R. Evid. Rule 401

[198] Fed. R. Evid. Rule 403

[199] Fed. R. Evid. Rule 702

[200] Committee Notes on Rule (2000 Amendment) *Fed R Evid Rule 702*. Available at https://www.law.cornell.edu/rules/fre/rule_702

[201] General Electric Co. v. Joiner (522 U.S. 136 (1997)).

[202] Claar v. Burlington N.R.R. (29 F.3d 499 (1994)).

[203] Compare Ambrosini v. Labarraque (101 F.3d 129 (D.C.Cir. 1996)).

[204] Sheehan v. Daily Racing Form, Inc. (104 F.3d 940, 942 (7th Cir. 1997)).

[205] Kumho Tire Co. v. Carmichael (526 U.S. 137 (1999)).

[206] Moore v. Ashland Chemical, Inc. (151 F.3d 269 (5th Cir. 1998)).

[207] Sterling v. Velsicol Chem. Corp. (855 F.2d 1188 (6th Cir. 1988)).

[208] Cole SA, Edmond G (2015) Science Without Precedent: The Impact of the National Research Council Report on the Admissibility and Use of Forensic Science Evidence in the United States. *Br J Am Leg Studies* 4:585-617.

[209] National Center on Forensics Post-PCAST Court Decisions Assessing the Admissibility of Forensic Science Evidence. Available at https://nij.ojp.gov/program/national-center-forensics/post-pcast-court-decisions-assessing-admissibility-forensic-science-evidence

[210] Schiltz P (2022) *Memorandum to Honorable John D. Bates, Chair, Standing Committee on Rules of Practice and Procedure RE: Report of the Advisory Committee on Evidence Rules*. Available at https://www.uscourts.gov/file/45210/download

[211] Judicial Conference Committee on Rules of Practice and Procedure (Standing Committee) (2022) Summary of the Report of the Judicial Conference Committee on Rules of Practice and Procedure, Agenda E-19 (Summary): Rules. (Judicial Conference). Available at https://www.uscourts.gov/file/45210/download

[212] Roberts J (2023) *Letter to Honorable Kevin McCarthy, Speaker, United States House of Representatives*. Available at https://www.supremecourt.gov/orders/courtorders/frev23_5468.pdf

[213] Roberts J (2023) *Letter to Honorable Kamala D. Harris, President, United States Senate*. Available at https://www.supremecourt.gov/orders/courtorders/frev23_5468.pdf

[214] Committee Notes on Rule (2023 Amendment) *Fed R Evid Rule 702*. Available at https://www.law.cornell.edu/rules/fre/rule_702

[215] U.S. Const. amend. V
[216] U.S. Const. amend. XIV
[217] Due Process Available at https://www.law.cornell.edu/wex/due_process
[218] U.S. Const. amend. VI
[219] Melendez-Diaz v. Massachusetts (557 U.S. 305 (2009)).
[220] Bullcoming v. New Mexico (564 U.S. 647 (2011)).
[221] Smith v. Arizona (22-899 (U.S. Mar 16, 2023)). Available at https://www.supremecourt.gov/docket/docketfiles/html/public/22-899.html
[222] Kwong K (2017) The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence. *Harv JL & Tech* 31:275-301.
[223] Cino JG (2017) Deploying the Secret Police: The Use of Algorithms in the Criminal Justice System. *Ga St UL Rev* 34(4):1073-1102.
[224] Imwinkelried EJ (2016) Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques. *DePaul L Rev* 66:97-132.
[225] Nutter PW (2019) Machine Learning Evidence: Admissibility and Weight. *U Pa J Const L* 21(3):919-958. Available at https://scholarship.law.upenn.edu/jcl/vol21/iss3/8
[226] Goodison S, Davis R, Jackson B (2015) Digital Evidence and the U.S. Criminal Justice System: Identifying Technology and Other Needs to More Effectively Acquire and Utilize Digital Evidence. (RAND Corporation, Santa Monica). Available at https://www.ojp.gov/pdffiles1/nij/grants/248770.pdf
[227] Novak M (2020) Digital Evidence in Criminal Cases Before the U.S. Courts of Appeal: Trend and Issues for Consideration. *Journal of Digital Forensics, Security and Law* 14(4):1-42. Available at https://commons.erau.edu/cgi/viewcontent.cgi?article=1609&context=jdfsl
[228] Miller CM (2023) A Survey of Prosecutors and Investigators Using Digital Evidence: A Starting Point. *Forensic Sci Int: Synergy* 6(100296):1-22. Available at https://doi.org/10.1016/j.fsisyn.2022.100296
[229] Cole SA (2018) Who Will Regulate American Forensic Science. *Seton Hall L Rev* 48(3):563-581. Available at https://scholarship.shu.edu/shlr/vol48/iss3/2/
[230] Butler JM (2017) Recent Activities in the United States Involving the National Commission of Forensic Science and the Organization of Scientific Area Committees for Forensic Science. *Aust J Forensic Sci* 49(5):526-540. Available at https://doi.org/10.1080/00450618.2016.1243153
[231] Organization of Scientific Area Committees for Forensic Science Support for Forensic Science Standards Development & Implementation. Available at https://www.nist.gov/organization-scientific-area-committees-forensic-science/support-forensic-science-standards
[232] Timmermans S, Epstein S (2010) A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annu Rev Sociol* 36:69-89.
[233] Morrison GS, Neumann C, Geoghegan PH (2020) *Vacuous Standards—Subversion of the OSAC Standards-Development Process* (Elsevier), Vol. 2. Available at https://doi.org/10.1016/j.fsisyn.2020.06.005

[234] Mohammed LA, McKiel M, Aschheim K, Berg G, Davis LA (2021) Response to Vacuous Standards Subversion of the OSAC Standards Development Process. *Forensic Sci Int: Synergy* 3(100145):1-2. Available at https://doi.org/10.1016/j.fsisyn.2021.100145

[235] International Organization for Standardization (ISO) ISO Technical Committee (TC272) Forensic Sciences. Available at https://www.iso.org/committee/4395817.html

[236] International Organization for Standardization (ISO) (2016) *ISO 18385:2016 Minimizing the Risk of Human DNA Contamination in Products used to Collect, Store and Analyze Biological Material for Forensic Purposes—Requirements*. Available at https://www.iso.org/standard/62341.html

[237] International Organization for Standardization (ISO) Forensic Science (07.140). Available at https://www.iso.org/ics/07.140/x/

[238] Bureau of Justice Statistics (2016) Publicly Funded Forensic Crime Laboratories: Quality Assurance Practices, 2014. (U.S. Department of Justice, Washington, D.C.). Available at https://bjs.ojp.gov/content/pub/pdf/pffclqap14.pdf

[239] International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC) *17025:2017 General Requirements for the Competence of Testing and Calibration Laboratories* (International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC)). Available at https://www.iso.org/standard/66912.html

[240] International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC) *17020:2012 Requirements for the Operation of Various Types of Bodies Performing Inspection*. Available at https://www.iso.org/standard/52994.html

[241] ANSI National Accreditation Board Forensic Accreditation. Available at https://anab.ansi.org/en/forensic-accreditation

[242] American Society of Crime Laboratory Directors (ASCLD) Our History. Available at https://www.ascld.org/our-history

[243] ANSI National Accreditation Board (2016) ANAB and ASCLD/LAB Merge Forensic Operations. Available at https://anab.ansi.org/latest-news/anab-and-ascldlab-merge-operations

[244] Forensic Technology Center of Excellence (2022) In-Brief: The American Society of Crime Laboratory Directors Accreditation Initiative: Successes, Challenges, and Future Directions (In-Brief). (RTI International, Research Triangle Park, NC). Available at https://forensiccoe.org/private/63077170df6b7

[245] National Commission on Forensic Science (2015) Recommendation to the Attorney General: Universal Accreditation. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/file/477851/download

[246] Lynch L (2015) *Memorandum for the Heads of Department Components: Recommendations of the National Commission on Forensic Science Announcement for NCFS Meeting Eight* (Office of the Attorney General, Washington, D.C.).

[247] National Commission on Forensic Science (2017) Recommendation to the Attorney General: Accreditation of Digital and Multimedia Evidence Forensic Service Providers. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/page/file/934401/download

[248] ANSI National Accreditation Board (2023) Accreditation Requirements for Forensic Testing and Calibration (AR 3125). Available at https://anab.qualtraxcloud.com/ShowDocument.aspx?ID=12371

[249] ANSI National Accreditation Board (2023) Accreditation Requirements for Forensic Inspection Bodies (AR 3120). Available at https://anab.qualtraxcloud.com/Showdocument.aspx?ID=14476

[250] National Commission on Forensic Science (2016) Views of the Commission: Certification of Forensic Science Practitioners. (U.S. Department of Justice, Washington, D.C.).

[251] Melbourn H, Smith G, McFarland J, Rogers M, Wieland K, DeWilde D, Lighthart S, Quinn M, Baxter A, Quarino L (2019) Mandatory Certification of Forensic Science Practitioners in the United States: A supportive Perspective. *Forensic Sci Int: Synergy* 1:161-169. Available at https://doi.org/10.1016/j.fsisyn.2019.08.001

[252] National Commission on Forensic Science (2016) Views of the Commission: Proficiency Testing in Forensic Science. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/page/file/1004461/download

[253] National Commission on Forensic Science (2016) Recommendation to the Attorney General: Proficiency Testing. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/page/file/905566/download

[254] Koertner AJ, Swofford HJ (2018) Comparison of Latent Print Proficiency Tests with Latent Prints Obtained in Routine Casework Using Automated and Objective Quality Metrics. *J For Ident* 68(3):379-388.

[255] Max B, Cavise J, Gutierrez RE (2019) Assessing Latent Print Proficiency Tests: Lofty Aims, Straightforward Samples, and the Implications of Nonexpert Performance. *J For Ident* 69(3):281-298.

[256] Kelley S, Gardner BO, Murrie DC, Pan KDH, Kafadar K (2020) How Do Latent Print Examiners Perceive Proficiency Testing? An Analysis of Examiner Perceptions, Performance, and Print Quality. *Science & Justice* 60(2):120-127. Available at https://doi.org/10.1016/j.scijus.2019.11.002

[257] Cowan EJ, Koppl R (2011) An Experimental Study of Blind Proficiency Tests in Forensic Science. *Rev Austrian Econ* 24(3):251-271. Available at https://doi.org/10.1007/s11138-010-0130-4

[258] Gardner BO, Neuman M (2023) Perceptions of Blind Proficiency Testing Among Latent Print Examiners. *Science & Justice* 63(2):200-205. Available at https://doi.org/10.1016/j.scijus.2022.12.005

[259] Pierce ML, Cook LJ (2020) Development and Implementation of an Effective Blind Proficiency Testing Program. *J Forensic Sci* 65(3):809-814.

[260] Hundl C, Neuman M, Rairden A, Rearden P, Stout P (2020) Implementation of a Blind Quality Control Program in a Forensic Laboratory. *J Forensic Sci* 65(3):815-822.

[261] Neuman M, Hundl C, Grimaldi A, Eudaley D, Stein D, Stout P (2022) Blind Testing in Firearms: Preliminary Results From a Blind Quality Control Program. *J Forensic Sci* 67(3):964-974.

[262] Mejia R, Cuellar M, Salyards J (2020) Implementing Blind Proficiency Testing in Forensic Laboratories: Motivation, Obstacles, and Recommendations. *Forensic Sci Int: Synergy* 2:293-298. Available at https://doi.org/10.1016/j.fsisyn.2020.09.002

[263] Forensic Technology Center of Excellence (2016) State Forensic Science Commissions. (National Institute of Justice, Washington, D.C.). Available at https://nij.ojp.gov/library/publications/state-forensic-science-commissions-final-report

[264] Forensic Technology Center of Excellence (2022) Forensic Science State Commissions and Oversight Bodies—A 2022 Update. (National Institute of Justice, Washington, D.C.). Available at https://nij.ojp.gov/library/publications/forensic-science-state-commissions-and-oversight-bodies-2022-update

[265] Bell S, Sah S, Albright TD, Gates Jr SJ, Denton MB, Casadevall A (2018) A Call for More Science in Forensic Science. *Proceedings of the National Academy of Sciences of the United States of America* 115(18):4541-4544. Available at www.pnas.org/cgi/doi/10.1073/pnas.1712161115

[266] National Association of Forensic Science Boards (2023) Available at https://www.nafsb.org/

[267] Texas Forensic Science Commission (2023) Commission to Host National Association of Forensic Science Boards Inaugural Meeting in Austin November 16-17, 2023. Available at https://www.txcourts.gov/fsc/news/commission-to-host-national-association-of-forensic-science-boards-inaugural-meeting-in-austin-november-16-17-2023/

[268] The Consortium of Forensic Science Organizations Available at https://thecfso.org/

[269] State of Texas Forensic Science Commission. Available at https://www.txcourts.gov/fsc/about-us/

[270] State of New York Commission on Forensic Science. Available at https://www.criminaljustice.ny.gov/forensic/aboutofs.htm

[271] State of Illinois Forensic Science Commission. Available at https://isp.illinois.gov/Forensics/ForensicScienceCommission

[272] Bureau of Justice Statistics (2018) Census of State and Local Law Enforcement Agencies - Statistical Tables, 2018. (U.S. Department of Justice, Washington, D.C.). Available at https://bjs.ojp.gov/library/publications/census-state-and-local-law-enforcement-agencies-2018-statistical-tables

[273] Organization of Scientific Area Committees for Forensic Science OSAC Registry Implementers. Available at https://www.nist.gov/organization-scientific-area-committees-forensic-science/osac-registry-implementers

[274] Bureau of Justice Statistics Forensic Sciences. Available at https://bjs.ojp.gov/topics/forensic-sciences

[275] Speaker P Foresight. *West Virginia University John Chambers College of Business and Economics* . Available at https://business.wvu.edu/research-outreach/forensic-business-studies/foresight

[276] National Institute of Justice Funding & Awards. Available at https://nij.ojp.gov/funding

[277] National Science Foundation Center for Advanced Research in Forensic Science (CARFS). Available at https://iucrc.nsf.gov/centers/center-for-advanced-research-in-forensic-science/

[278] National Institute of Standards and Technology Forensic Science Center of Excellence Frequently Asked Questions. Available at https://www.nist.gov/coe/forensic-science-center-excellence-frequently-asked-questions

[279] Iowa State University (2015) Iowa State to be Home to a New, $20 Million National Center for Forensic Science. Available at https://www.news.iastate.edu/news/2015/05/25/forensicsciencecenter

[280] Iowa State University (2020) Iowa State University CSAFE Receives $20 Million Extension. Available at https://www.research.iastate.edu/news/iowa-state-university-csafe-receives-20-million-extension

[281] House of Lords Science and Technology Select Committee (2019) Forensic Science and the Criminal Justice System: A Blueprint for Change. *3rd Report of Session 2017-19*. Available at https://publications.parliament.uk/pa/ld201719/ldselect/ldsctech/333/333.pdf

[282] Wilcox AM, NicDaeid N (2018) Jurors' Perceptions of Forensic Science Expert Witnesses: Experience, Qualifications, Testimony Style and Credibility. *Forensic Science International* 291:100-108.

[283] National Commission on Forensic Science (2016) Views of the Commission: Technical Merit Evaluation of Forensic Science Methods and Practices. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/archives/ncfs/file/881796/download

[284] Roux C (2019) Forensic Science Isn't "Reliable" or "Unreliable": It Depends on the Questions You're Trying to Answer. *The Conversation*. Available at https://phys.org/news/2019-09-forensic-science-isnt-reliable-unreliable.amp

[285] Ainsworth R (2023) Forensic Odontology Response to the NIST Report and CIFS Video. *American Academy of Forensic Sciences* . Available at https://www.aafs.org/article/forensic-odontology-response-nist-report-and-cifs-video

[286] Bjerkhoel A, Plourd HCJ (2023) The Death Row Case of Ray Krone, the Beginning of the End of Bite Mark Evidence in the United States. *Journal of the California Dental Association* 51(1):2210330. Available at https://doi.org/10.1080/19424396.2023.2210330

[287] National Commission on Forensic Science (2015) Views of the Commission: Scientific Literature in Support of Forensic Science and Practice. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/ncfs/file/786591/download

[288] National Commission on Forensic Science (2016) Views of the Commission: Regarding Identifying and Evaluating Literature that Supports the Basic Principles of a Forensic Science Method or Forensic Science Discipline. (U.S. Department of Justice, Washington, D.C.). Available at https://www.justice.gov/ncfs/file/839716/download

[289] Burns PB, Rohrich RJ, Chung KC (2011) The Levels of Evidence and Their Role in Evidence-Based Medicine. *Plastic and reconstructive surgery* 128(1):305-310. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3124652/#!po=15.5172

[290] United Kingdom Forensic Science Regulator (2020) Validation (FSR-G-201, Issue 2). Available at https://www.gov.uk/government/publications/forensic-science-providers-validation

[291] United Kingdom Forensic Science Regulator (2020) Method Validation in Digital Forensics (FSR-G-218, Issue 2). Available at https://www.gov.uk/government/publications/method-validation-in-digital-forensics

[292] Human Factors Committee (2020) Human Factors in Validation and Performance Testing of Forensic Science (OSAC Technical Series 0004). Available at https://doi.org/10.29325/OSAC.TC.0004

[293] Trejos T, Koch S, Mehltretter A (2020) Scientific Foundations and Current State of Trace Evidence—A Review. *Forensic Chemistry* 18(100223):1-22. Available at https://doi.org/10.1016/j.forc.2020.100223

[294] Speaker Recognition Subcommittee of the Organization of Scientific Area Committees for Forensic Science (2021) Essential Scientific Literature for Human-Supervised Automatic Approaches to Forensic Speaker Recognition. Available at https://www.nist.gov/system/files/documents/2021/06/09/SR%20Essential%20Literature%2020210415_FINAL.pdf

[295] Taylor D, Biedermann A, Samie L, Pun K-M, Hicks T, Champod C (2017) Helping to Distinguish Primary from Secondary Transfer Events for Trace DNA. *Forensic Science International: Genetics* 28:155-177.

[296] Scientific Working Group for the Analysis of Seized Drugs Available at https://www.swgdrug.org/

[297] Rabin B, Swanson K, Glasgow R, Ford B, Huebschmann A, Marsh R, Tabak R, Brownson R, Malone S, Padek M Dissemination and Implementation Models in Health Research and Practice Webtool. Available at https://dissemination-implementation.org/

[298] American Society of Crime Laboratory Directors (ASCLD) Forensic Research Committee (FRC). Available at https://www.ascld.org/forensic-research-committee/

[299] Gamette MJ, Wickenheiser RA (2023) Establishment of the National Technology Validation and Implementation Collaborative (NTVIC) and Forensic Investigative Genetic Genealogy Technology Validation Working Group (FIGG-TVWG). *Forensic Sci Int: Synergy* 6(100317):1-3. Available at https://doi.org/10.1016/j.fsisyn.2023.100317

[300] Morrissey J, Stodter A, Sherratt F, Cole MD (2023) Partnership Between Academics and Practitioners–Addressing the Challenges in Forensic Science. *Science & Justice* 63(1):74-82. Available at https://doi.org/10.1016/j.scijus.2022.11.005

[301] Kleinman MS, Mold JW (2009) Defining the Components of the Research Pipeline. *Clinical and Translational Science* 2(4):312-314. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5407485/pdf/CTS-2-312.pdf

[302] GAO (2020) Technology Readiness Assessment Guide: Best Practices for Evaluating the Readiness of Technology for Use in Acquisition Programs and Projects (GAO-20-48G). (U.S. Government Accountability Office, Washington, D.C.). Available at https://www.gao.gov/assets/gao-20-48g.pdf

[303] United Kingdom (2021) Forensic Science Regulator Act of 2021. https://www.legislation.gov.uk/ukpga/2021/14/contents/enacted.

[304] DNA Identification Act of 1994 (P.L. 103-322) (1994) Available at https://www.congress.gov/103/statute/STATUTE-108/STATUTE-108-Pg1796.pdf

[305] ANSI National Accreditation Board ISO/IEC 17025 Forensic Testing Laboratory Accreditation. Available at https://anab.ansi.org/forensic-accreditation/fbi-qas

[306] ASTM International Forensic Science Standards. Available at https://www.astm.org/products-services/standards-and-publications/standards/forensic-science-standards.html

[307] Wilson-Wilde L (2018) The International Development of Forensic Science Standards—A Review. *For Sci Int* 288:1-9. Available at https://doi.org/10.1016/j.forsciint.2018.04.009

[308] Czyryca C (2015) A Proficiency Test Provider's Perspective: Collaborative Testing Services, Inc. (Presentation for National Commission on Forensic Science on behalf of The Accreditation and Proficiency Testing Subcommittee). (Washington, D.C.). Available at http://www.justice.gov/ncfs/file/761061/download

[309] 74 OK Statute (§ 74-150.37) (2014)

[310] D.C. Code § 5–1501.06 (2012)