

May 2020

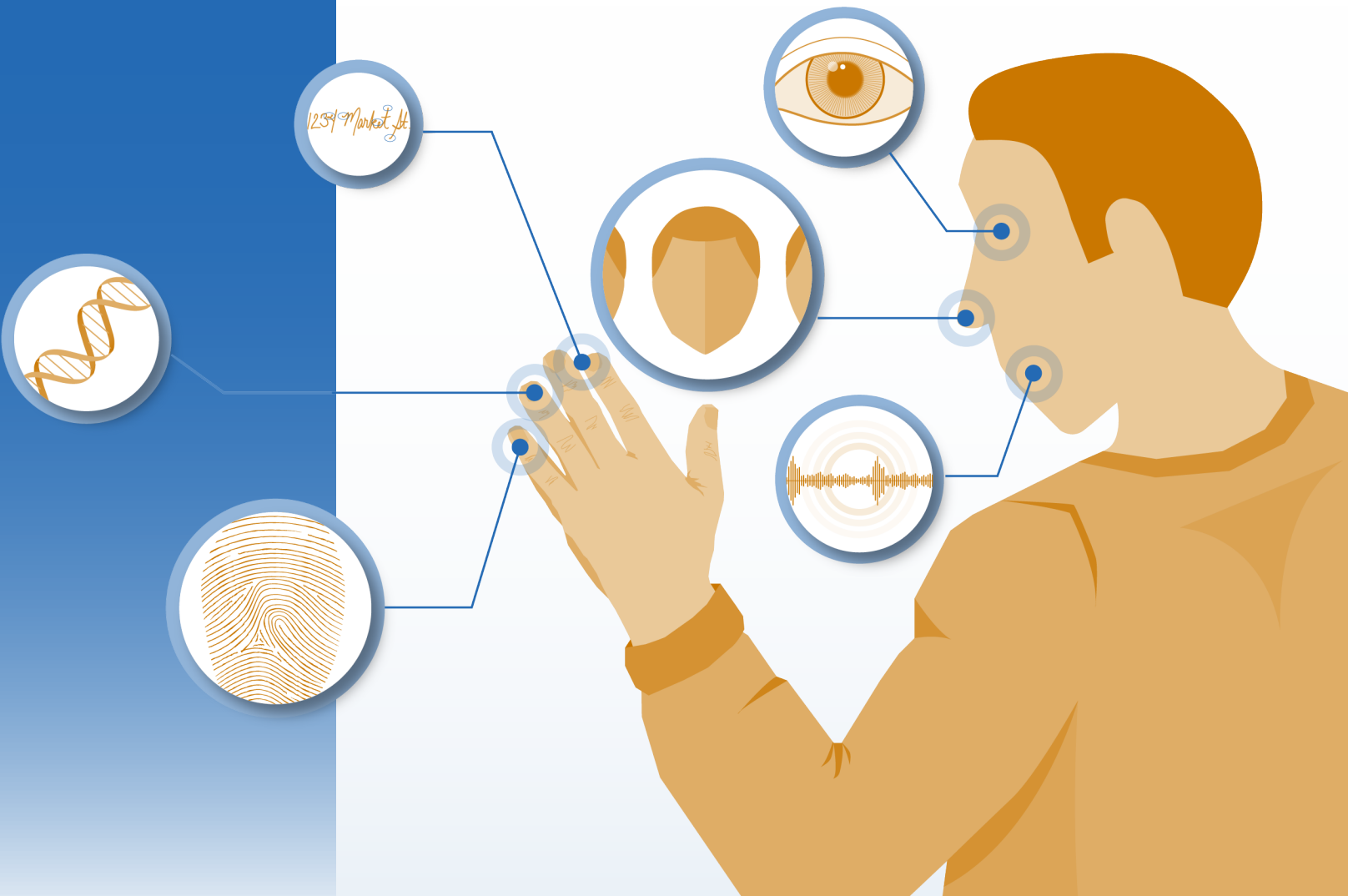
Accessible Version

TECHNOLOGY ASSESSMENT

Forensic Technology

Algorithms Used in Federal Law Enforcement

GAO-20-479SP



The cover image displays examples of evidence that federal law enforcement agencies can analyze using forensic technologies. From left to right, these images depict evidence that can be analyzed using probabilistic genotyping, latent print, handwriting recognition, face recognition, iris recognition (upper graphic), and voice recognition (lower graphic) algorithms.

GAO Highlights

Highlights of [GAO-20-479SP](#), a report to congressional requesters

May 2020

Why GAO did this study

Forensic algorithms help forensic experts partially automate the process of assessing whether or not evidence collected in an investigation may have originated from an individual, potentially increasing the speed of investigations and reducing human bias and error.

GAO was asked to conduct a technology assessment on the use of forensic algorithms in federal law enforcement. GAO is conducting this assessment in two phases. The first phase describes algorithms being used by federal law enforcement agencies and how these technologies work. The second phase will assess the approaches and challenges related to how federal law enforcement agencies apply these technologies and will identify policy options for addressing these challenges going forward.

In conducting this assessment, GAO obtained information from the National Institute of Standards and Technology, the Department of Justice, the Department of Homeland Security, and the Department of Defense; convened an interdisciplinary panel of 16 experts with assistance from the National Academies of Sciences, Engineering, and Medicine; interviewed additional stakeholders, including nonprofit groups and legal experts; and reviewed relevant literature and case law.

View [GAO-20-479SP](#). For more information, contact Karen L. Howard, PhD at (202) 512-6888, howardk@gao.gov.

TECHNOLOGY ASSESSMENT

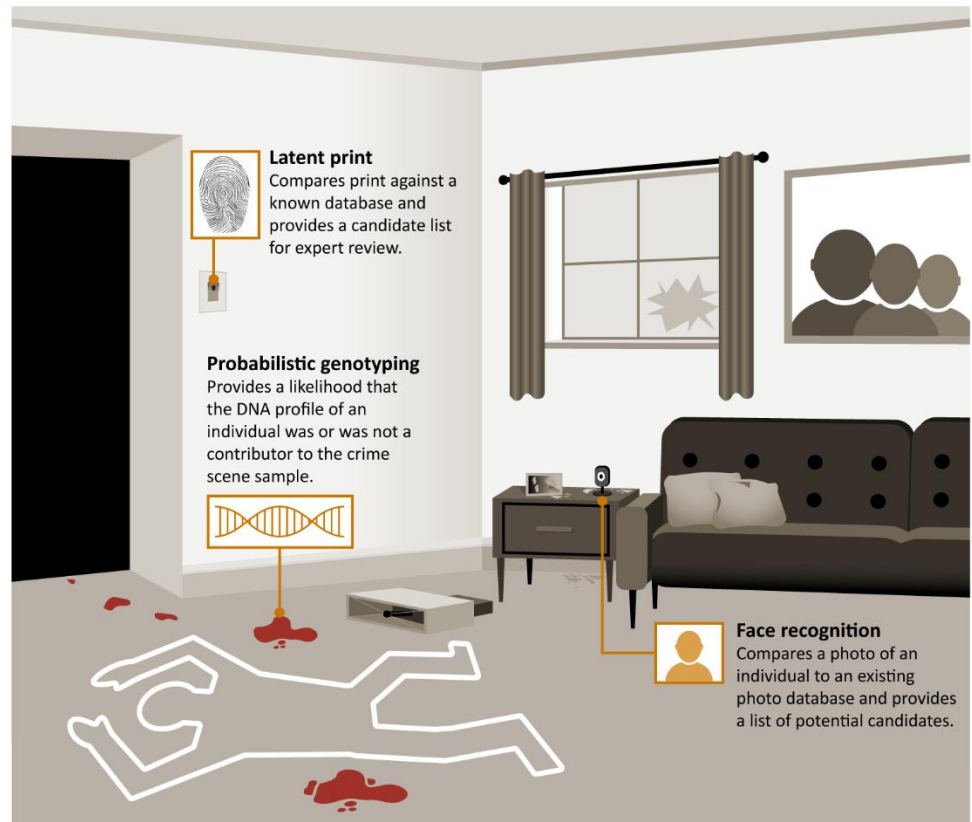
Forensic Technology

Algorithms Used in Federal Law Enforcement

What GAO found

Federal law enforcement agencies GAO reviewed are primarily using three types of forensic algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual: probabilistic genotyping, latent print (fingerprint and palm print) analysis, and face recognition. To a lesser extent, agencies also use algorithms to compare iris images, speech, and handwriting. Each type of algorithm uses different characteristics in its assessment. For example, probabilistic genotyping uses statistics to analyze biological samples found during a criminal investigation to assist in comparisons to a known DNA sample taken from a suspect, or to DNA data profiles from a database of known persons. The Federal Bureau of Investigation currently uses probabilistic genotyping and latent fingerprint algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual and face recognition to generate investigative leads. The National Institute of Standards and Technology and other organizations have developed standards to facilitate transmission of data between agencies.

Potential Uses of Forensic Algorithms to Examine Evidence from a Crime Scene



Source: GAO. | GAO-20-479SP

Latent print

Compare print against a known database and provides a candidate list for expert review.

Probabilistic genotyping

Provides a likelihood that the DNA profile of an individual was or was not a contributor to the crime scene sample.

Face recognition

Compares a photo of an individual to an existing photo database and provides a list of potential candidates.

Table of Contents

GAO Highlights3

Introduction.....1

1 Background.....3

2 Federal Law Enforcement Agencies Primarily Use Three Kinds of Forensic Algorithms.....5

 2.1 Use of probabilistic genotyping, latent print, and face recognition algorithms5

 2.2 Use of other algorithms.....6

 2.3 How forensic algorithms work.....7

 2.3.1 Probabilistic genotyping software (PGS)7

 2.3.2 Latent print analysis9

 2.3.3 Face recognition10

 2.3.4 Iris recognition.....12

 2.3.5 Voice recognition.....12

 2.3.6 Handwriting recognition.....12

3 Agencies Use Data Standards to Help Them Transmit Evidence.....13
4 Agency and Expert Comments15
Appendix I – Objective, Scope, and Methodology16
Appendix II: Expert Meeting Participation18
Appendix III: GAO Contact and Staff Acknowledgments.....19

ABBREVIATIONS

AFIS	Automated Fingerprint Identification System
AI	artificial intelligence
ANSI	American National Standards Institute
CBP	Custom and Border Protection
DFSC	Defense Forensic Science Center
DHS	Department of Homeland Security
DOD	Department of Defense
DOJ	Department of Justice
EBTS	Electronic Biometric Transmission Specification
FBI	Federal Bureau of Investigation
IEC	International Electrotechnical Commission
INCITS	InterNational Committee for Information Technology Standards
ISO	International Organization for Standardization
ITL	Information Technology Laboratory
NGI	Next Generation Identification
NIST	National Institute of Standards and Technology
OBIM	Office of Biometric Identity Management
PGS	Probabilistic Genotyping Software
TSA	Transportation Security Administration

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



441 G St. N.W.
Washington, DC 20548

Introduction

May 12, 2020

The Honorable Eddie Bernice Johnson
Chairwoman
The Honorable Frank Lucas
Ranking Member
Committee on Science, Space, and Technology
House of Representatives

The Honorable Carolyn Maloney
Chairwoman
Committee on Oversight and Reform
House of Representatives

The Honorable Mark Takano
House of Representatives

For more than a century, law enforcement agencies have examined certain types of physical evidence, such as whorls on fingerprints, to help identify suspects, solve cold cases, and find missing or exploited people. Scientific advances are now allowing forensic experts to partially automate the process of assessing the likelihood that evidence collected in a criminal investigation may have originated from an individual—a process known as forensic attribution—using forensic algorithms run on computers. Federal law enforcement agencies have adopted or are currently evaluating such algorithms to improve the speed and objectivity of their work.

Based on the emergence of this technology, you requested that we examine the use of forensic algorithms in federal law enforcement. This technology assessment describes forensic algorithms that are being used by federal law enforcement to help assess whether or not evidence collected in a criminal investigation may have originated from an individual and how those technologies work. To address this objective, we obtained information from the Department of Commerce’s National Institute of Standards and Technology (NIST), the Department of Justice (DOJ), the Department of Homeland Security (DHS), and the Department of Defense (DOD); convened an interdisciplinary panel of 16 experts with assistance from the National Academies of Sciences, Engineering, and Medicine; conducted interviews with additional stakeholders, including nonprofit groups and legal experts; conducted literature searches; and reviewed relevant literature and case law.

We conducted our work from August 2019 to May 2020 in accordance with all sections of GAO's Quality Assurance Framework that are relevant to technology assessments. The Framework requires that we plan and perform the engagement to obtain sufficient and appropriate evidence to meet our stated objectives and to discuss any limitations to our work. We believe that the information and data obtained, and the analysis conducted, provide a reasonable basis for any findings and conclusions in this product.

1 Background

Forensic algorithms can help assess evidence through the process of expert-performed forensic attribution using what are called characteristic comparison methods. These methods, including latent print (e.g., fingerprint and palm prints) analysis and DNA analysis using probabilistic genotyping software (PGS), “attempt to determine whether an evidentiary sample (e.g., from a crime scene) is or is not associated with a potential ‘source’ sample (e.g., from a suspect), based on the presence of similar patterns, impressions, or other features in the sample and the source,” according to a 2016 report by the President’s Council of Advisors on Science and Technology.¹

There are many methods for identifying characteristics specific to an individual (e.g., prints and DNA) to assess whether or not evidence collected in a criminal investigation may have originated from an individual. Human feature comparison methods have existed circa 200 B.C., when the Chinese used prints for identification. The first known example of law enforcement use of prints was in the late 1800s in Argentina for identifying prisoners.

Each characteristic comparison method uses different attributes to assess whether an individual should be considered a person of interest or otherwise determine whether or not evidence collected in a criminal investigation may have originated from an individual. For example:

- PGS analysis assists in the interpretation of an electropherogram—a plot of DNA fragment sizes—derived from the evidence. The results can be compared to a reference profile from one or more persons of interest.
- Latent print analysis includes fingerprint and palm print analysis. Latent fingerprint analysis conducted by DOJ’s Federal Bureau of Investigation (FBI) compares the features of a partial print collected during a criminal investigation to the features of a “tenprint”—a set of all 10 fingerprints—or a set of palm prints taken from an individual under controlled conditions—from a suspect or stored in a database of known persons.
- Face recognition analysis conducted by the FBI compares a facial image of a suspect against images in a database of known persons.

Prior to the advent of these forensic algorithms, experts manually performed these characteristic comparison methods by visually comparing evidence with DNA data, stored fingerprints, or a collection of photographs. However, there were limitations to this approach, most notably the following:

- **Error and bias.** Comparisons performed manually can be subject to human error and bias, limiting accuracy of the results.

¹President’s Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity*

of Feature-Comparison Methods (Washington, D.C., September 2016).

- **Resources.** Manual comparisons can be time consuming and laborious.

In an effort to reduce such limitations, the FBI, in collaboration with NIST, sponsored research in the 1960s to automate some forensic characteristic comparison methods, such as prints. Since then, significant advances have occurred, including the incorporation of artificial intelligence (AI). Today, forensic laboratories, research groups, and commercial vendors develop new algorithms. However, bias and human error are still present given that investigators collect and select the evidence for analysis, and that analysis by humans is part of both PGS and latent print analysis. Forensic algorithms are also only as good as their source data. PGS can analyze DNA with low qualities and quantities, but these variables can have an effect on the result. In addition, there is a wide variation of quality and ways a latent print or image (such as for face recognition) can be captured. Image quality will affect the result of the forensic algorithm analysis.

During a criminal investigation, the FBI and other agencies can use forensic algorithms to generate candidates for comparisons—for example, candidate suspects whose fingerprints are consistent with those in a database. They can also be used to generate investigative leads, or both. Forensic algorithms can provide a numerical likelihood score indicating the likelihood whether or not an individual is or is not the person associated with evidence collected in the criminal investigation. An investigative lead identifies one or more individuals who are potential suspects for further investigation. Forensic algorithms do not assign guilt or innocence, rather they provide leads and potential data

to be incorporated into the investigation along with information collected from multiple other sources.

The FBI uses PGS and latent print analysis to provide candidates for comparison by an expert. For example, the FBI uses latent print analysis to help assess whether or not evidence collected in a criminal investigation may have originated from an individual if an individual's tenprint is incorporated into a database.

The FBI uses face recognition algorithms for investigative leads. For example, according to the FBI, its algorithm can identify a subset of photos from within the Interstate Photo System gallery as a candidate list. The candidate list is then reviewed by a trained examiner, and any photo determined to be a valid investigative lead is forwarded to an FBI investigator.

2 Federal Law Enforcement Agencies Primarily Use Three Kinds of Forensic Algorithms

Federal law enforcement agencies we reviewed primarily use probabilistic genotyping, latent print, and face recognition algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual. To a more limited extent, agencies also use algorithms to compare iris images, speech, and handwriting.

2.1 Use of probabilistic genotyping, latent print, and face recognition algorithms

We found that federal law enforcement agencies we reviewed use three main types of forensic algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual: probabilistic genotyping, latent print analysis, and face recognition algorithms.

DOJ uses forensic algorithms in its criminal investigations. The FBI's Laboratory Services uses PGS to assist in the interpretation of DNA evidence while investigating criminal cases. The FBI's Criminal Justice Information Services—a division that provides tools and services to law enforcement, national security, intelligence community partners, and the general public—has a repository of biometric and criminal history data known as

the Next Generation Identification (NGI) System. The FBI uses this system in combination with latent print algorithms and face recognition algorithms to determine whether evidence may have originated from an individual. In addition, DOJ's Drug Enforcement Administration uses NGI to help with identifying latent prints.

DHS uses forensic algorithms to support homeland security missions and criminal investigations. For example, the Office of Biometric Identity Management (OBIM) uses forensic algorithms in its Automated Biometric Identification System called IDENT for face recognition and latent print analysis.² The Transportation Security Administration (TSA) uses FBI and OBIM biometric fingerprint algorithms as a part of its civil background investigations.³ TSA is also testing the U.S. Custom and Border Protection (CBP) biometric facial algorithms to verify passenger identities, for passengers who have opted into the TSA biometric testing program. Since 2017, TSA has conducted a series of pilot tests—some in partnership with CBP—to assess the feasibility of using face recognition to automate traveler identity verification at airport security checkpoints.

Finally, DOD also uses forensic algorithms for criminal investigations. Within DOD, the Defense Forensic Science Center (DFSC)

²The Automated Biometric Identification System is the central DHS-wide system for storage and processing of biometric and associated biographic information for national security; law enforcement; immigration and border management; intelligence; background investigations for national security positions and certain positions of public trust; and associated

testing, training, management reporting, planning and analysis, or other administrative uses.

³TSA is not primarily a law enforcement agency, but does have a law enforcement component (the Law Enforcement/Federal Air Marshal Service).

performs forensic analyses, using latent print analysis and probabilistic genotyping algorithms. An official from a unit of DFSC told us that the agency also submits evidence to other agencies for forensic database searching support, as well as developing its own internal algorithms. For example, DFSC has its own software package for latent print analysis, which can be used to provide statistical support for results of manual comparisons. DFSC recently made this available as open source software. For DNA analysis, DFSC uses software to assist with the separation of mixed DNA profiles—those that contain DNA from more than one individual—and a separate program to assist with certain calculations, such as inferring the biological sex of an individual based on evidence collected during a criminal investigation.

2.2 Use of other algorithms

To a lesser extent, federal law enforcement agencies we reviewed also use other algorithms to assess whether or not evidence collected in a criminal investigation may have originated from an individual, such as algorithms for comparing iris images, voice recordings, and handwriting.

- **Iris recognition algorithms.** Iris recognition algorithms compare images of an individual’s iris to a database of iris images. DHS’s OBIM uses iris methods as part of its IDENT system. FBI officials said that the agency has a pilot program to develop

iris matching algorithms, which it will soon incorporate into the NGI System as the National Iris System.

- **Voice recognition algorithms.** Officials with the U.S. Secret Service told us that it has the ability to compare a recording of an unknown speaker with one or more recordings of known speakers to help investigators identify the unknown speaker. OBIM is also exploring the use of automatic voice recognition algorithms.
- **Handwriting recognition algorithms.** U.S. Secret Service officials said their agency can use a computer algorithm to compare manually collected digital measurements of handwriting characteristics to previously collected measurements, some of which may be attributed to a known author.

In addition to these algorithms that are in use or being tested, agencies are researching and developing additional algorithms they may use in the future. Our expert meeting participants identified gait analysis and genetically variant peptide analysis algorithms as methods being researched.⁴ According to NIST publications and FBI officials, the two agencies previously collaborated on research on image-based tattoo recognition algorithms.⁵

⁴A peptide is a molecule consisting of two or more amino acids. Peptides are smaller than proteins, which are also chains of amino acids.

⁵NIST, *Tattoo Recognition Technology—Challenge (Tatt-C): Outcomes and Recommendations (Revision 1.0)*, NISTIR 8078

(Washington, D.C.: September 2016) and NIST, *Tattoo Recognition Technology—Evaluation (Tatt-E): Performance of Tattoo Identification Algorithms*, NISTIR 8232 (Washington, D.C.: October 2018).

2.3 How forensic algorithms work

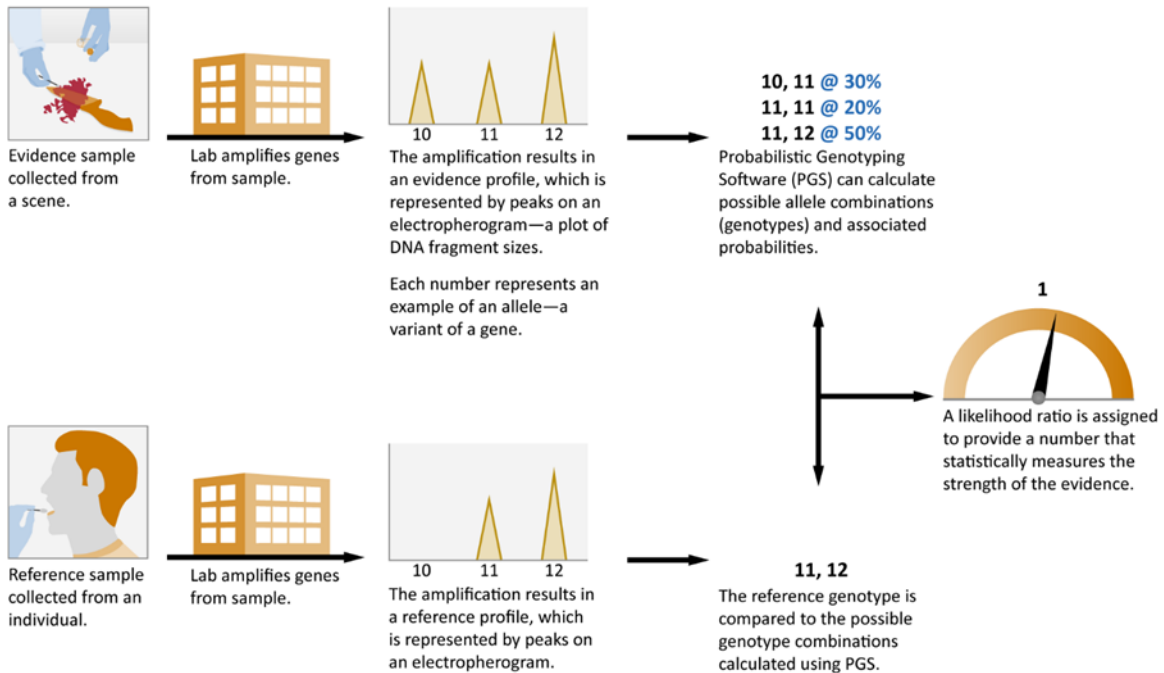
2.3.1 Probabilistic genotyping software (PGS)

PGS may be used to evaluate single source or complex DNA evidence.⁶ PGS posits two competing hypotheses: A) the DNA evidence is the result of contributions by a person of interest and other unknown, unrelated individuals and B) the DNA evidence is the result of contributions by unknown individuals. It provides a likelihood that the observed data resulted from each scenario, giving a likelihood ratio of hypothesis A versus hypothesis B. PGS is more effective than

traditional DNA analysis when the DNA is from two or more individuals or when DNA from some or all of the contributors is present in low quantities. Unlike conventional approaches, PGS can attach a number that statistically measures the strength of the evidence when a DNA sample is from an unknown number of contributors where it is possible that some of the DNA from one or more contributors failed to be detected.

The first step in DNA analysis usually involves the extraction of genetic material from both the evidence and reference samples (see fig. 1). Commercially available kits are then used to repeatedly copy specific regions of human

Figure 1: How probabilistic genotyping software works



Source: GAO. | GAO-20-479SP

⁶In this section we provide a generalized description of how forensic algorithms work for PGS, latent prints, face, iris, voice,

and handwriting recognition. However, each algorithm and software package may differ depending on the developer.

Data for Figure 1: How probabilistic genotyping software works

- 1a. Evidence sample collected from a scene.
- 1b. Lab amplifies genes from sample.
- 1c. The amplification results in an evidence profile, which is represented by peaks on an electropherogram – a plot of DNA fragment sizes. Each number represents an example of an allele – a variant of the gene.
- 1d. Probabilistic Genotyping Software (PGS) can calculate possible allele combinations (genotypes) and associated probabilities.
- 2a. Reference sample collected from an individual.
- 2b. Lab amplifies genes from sample.
- 2c. The amplification results in a reference profile, which is represented by peaks on an electropherogram.
- 2d. The reference genotyping is compared to the possible genotype combinations calculated using PGS.
3. A likelihood ratio is assigned to the statistical strength of the evidence, whether the individual is a contributor or non-contributor to the evidence sample.

DNA that are likely to differ in lengths across individuals. These amplified pieces of DNA are separated by size. The resulting mix of fragment lengths represents a profile, also known as a genotype. The profile is normally represented as a series of peaks on a graph known as an electropherogram. Investigators generate profiles from the crime scene evidence sample. Separately, they may also generate profiles from samples taken from one or more persons of interest called a reference sample. This process is used in multiple types of DNA analysis and is not unique to PGS.

What distinguishes PGS is the steps that follow. In the first of these, investigators use a mathematical model encoded in PGS software to estimate the likelihoods associated with

two competing hypotheses, such as hypotheses A and B described above.

PGS mathematically compares the crime scene profile with many hypothetical profiles based on various possible genotype combinations. This process allows the software to assess the relative likelihood that various genotype combinations contributed to the crime scene sample. It also allows the software to separate out genotypes of individual contributors. The first steps of PGS do not use a genotype from an individual in question.

The most sophisticated PGS software models examine many variables simultaneously and can be very computationally intensive. They often do this through a computer simulation that considers a large number of contributor

combinations of, for example, two-, three-, and four-person mixtures. Using a set of parameters and mathematical modeling of the data, the computer estimates the likelihood that each of these combinations best explains the results. For those hypotheses that posit a contribution from a specific individual, the software will simulate large numbers (often hundreds of thousands) of possible states of those variables and return an estimate of the probability that the test results from the evidence sample would appear as they did if the individual had contributed to it. PGS software will then perform the same simulation using the same model to estimate the probability that the test results from the evidence sample would appear as they did if a different theory of the case was correct.

These probability estimates are highly dependent on the models and their variables. However, the ratio of the likelihoods of observing the data under two alternative hypotheses for a single evidence sample can be helpful if they have been estimated by software using the same models and variables. If the ratio of the likelihood of hypothesis A to that of hypothesis B is greater than 1, the test results are more consistent with hypothesis A. Likelihood ratios of less than one suggest that the test results are more consistent with B. (A likelihood ratio is not the probability that the individual's DNA is actually contained in a DNA mixture.)

Once the algorithms have determined the weighted optimal combinations of contributors to the mixture (independent of

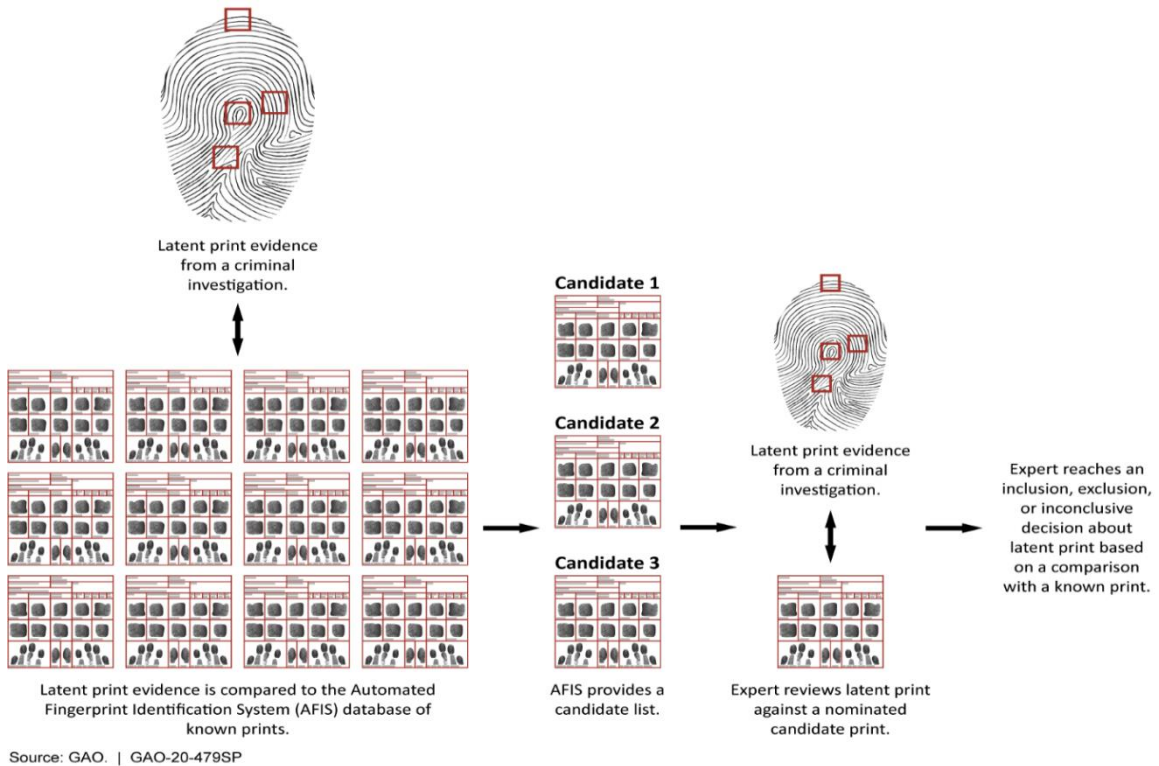
the profile from the person of interest), investigators compare the known profile of the individual to the weighted combinations to determine if the individual can be explained as being a contributor or non-contributor to the mixture. The result is a number that statistically measures the strength of the evidence, which can be taken into consideration by investigators.

2.3.2 Latent print analysis

A latent print can be a partial or incomplete print left on a surface and then recovered during a criminal investigation. It may be smudged or distorted. In latent fingerprint analysis, investigators compare a latent print with a tenprint—a set of prints from all 10 of an individual's fingers, taken under controlled circumstances. Additionally, investigators can compare a latent palm print with known palm

prints—a set of four or six prints of known palm data. The latent print is digitally scanned, its details or minutiae marked by a human examiner, and the scan is uploaded into the Automated Fingerprint Identification System (AFIS), which uses multiple algorithms to analyze the print. The algorithms can improve image quality and read the many minutiae specific to a fingerprint or palm print. The algorithms also compare the layout of minutiae detected in the latent print to those found in a tenprint and palm print database of known individuals. This comparison provides a list of individuals who may be the source of the latent print found during an investigation. An expert independently compares this list of

Figure 2: How latent print analysis works



Data for Figure 2: How latent print analysis works

1. Latent print evidence from criminal investigation.
2. Latent print evidence is compared to the Automated Fingerprint Identification System (AFIS) database of known prints.
3. AFIS provides candidate list.
4. Expert reviews latent print against a nominated candidate print.
5. Expert reaches an inclusion, exclusion, or inconclusive decision about latent print based on a comparison with a known print.

Note: Squares indicate print minutiae.

candidates and, based on their own judgment, reaches an identification, exclusion, or inconclusive decision (see fig. 2).

2.3.3 Face recognition

If an image of an unknown individual associated with a criminal investigation is available, face recognition could compare it

against a database of images of known persons. For example, the image of a known individual is captured under controlled conditions. During comparison to photos of known individuals, a probe photograph (a photo of an unknown individual) is compared against the photos of known individuals in the NGI System that were obtained in controlled conditions. The separate enrollment and matching phases usually depend on multiple

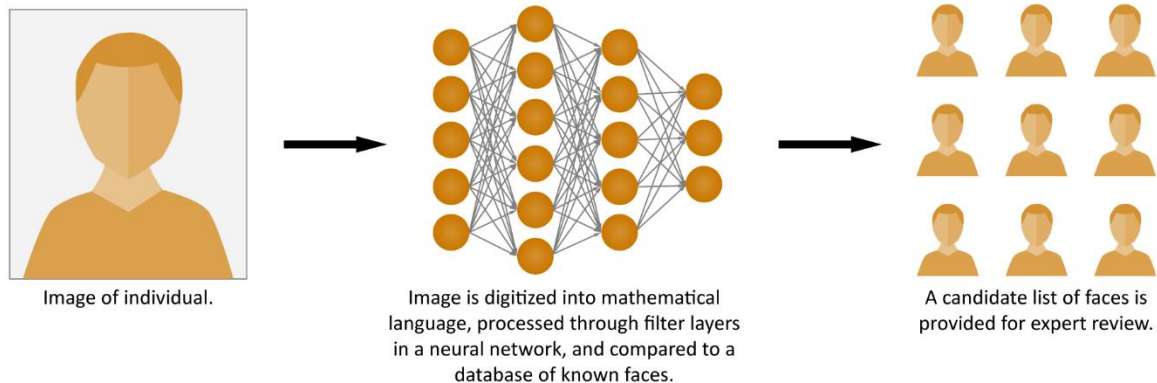
algorithms. For example, in enrollment an initial algorithm will detect the face in the probe photo and orient it. A second algorithm will then analyze the entire set of pixels across the image to generate a mathematical representation of the face. This mathematical representation of the face is called a “template.” A matching algorithm is then used to compare the probe template to an entire gallery database of known templates.

This process may use an AI technology known as convolutional neural networks. A probe photograph of an individual is digitized into a mathematical language that forms a template. A program runs this information through several algorithms and compares it to a database of known facial images. This results in a candidate list of faces from the database, with a ranking from most to least similar to the probe photograph.

Convolutional neural networks may use multiple layers to analyze templates. After the

image is filtered through the layers, the resulting mathematical patterns are compared with those extracted similarly from face images in a known database. This comparison method does not use facial features (e.g., eye distance or nose size), but rather uses mathematical aspects of a digitized image. This comparison generates a “similarity score” which is specific to individual algorithms. Once the probe photo has been compared against the entire database, the system will present to the user a candidate list of photos ranked from highest similarity score to lowest. If the algorithm identifies multiple likely candidates, it will generate a list of best-matched photos. The length of this candidate list is determined by the system operator, but typically is between 20 and 100. In contrast, the system could return no candidates if no database photos are found to be sufficiently similar to the probe photo (see fig. 3)

Figure 3: How face recognition works



Source: GAO. | GAO-20-479SP

Data for Figure 3: How face recognition works

1. Image of individual.

2. Image is digitized into mathematical language, processed through filter layers in a neural network, and compared to a database of known faces.
3. A candidate list of faces is provided for expert review.

2.3.4 Iris recognition

Iris recognition compares an iris image associated with an investigation or a person of interest to a database of known iris image patterns. The iris recognition software uses handcrafted algorithms (as opposed to AI) to convert the digital image into mathematical patterns of the digitized iris, known as an IrisCode. The mathematical patterns of the IrisCode are compared to other IrisCodes of known individuals. Using statistical comparisons, the algorithms determine whether the two things being compared are likely to be from the same or different individuals.

2.3.5 Voice recognition

A voice sample that is associated with an investigation can be isolated and analyzed by voice recognition software. In a forensic case, the voice sample from the investigation and a known voice sample are provided to software that uses forensic algorithms to find abstract, short-term features. These abstract features can be put through further layers of processing and then compared to produce a numeric score giving the similarity between the investigative and known samples. The automatic system will output a likelihood ratio (i.e., the likelihood of observing the measured similarity between speech samples assuming that they were spoken by the same speaker or different speakers). These results can be fragile, in the sense of being dependent on confusing factors such as type of microphone, background noise, and transmission channel.

2.3.6 Handwriting recognition

As with latent prints, handwriting samples associated with an investigation can be collected and digitally scanned. The handwriting samples are uploaded into software that uses forensic algorithms to perform digital measurements of the handwriting features that have been manually marked by an expert. Comparisons can be made between evidence and either a known or unknown writing sample. An expert then reviews the results.

3 Agencies Use Data Standards to Help Them Transmit Evidence

We found four standards that agencies use to facilitate the transmission of data between agencies for examination by PGS, latent print, and face recognition algorithms. In our review, we found one international standard, one U.S. standard, and two standards specific to a federal agency (see table 1).⁷

The international standard was developed jointly by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) to enable the interoperability and data interchange among biometric applications and systems. It includes guidance for fingerprints, facial images, and DNA data (used for PGS).

NIST developed a standard for prints, facial images, and DNA data. The ANSI/NIST standards were developed for federal agencies to specify a common format for data exchange across jurisdictional lines or between dissimilar systems made by different manufacturers. According to NIST officials, these standards were developed with criminal justice in mind.

The FBI developed a standard for electronically encoding and transmitting biometric image, identification, and arrest data known as the Electronic Biometric Transmission Specification (EBTS). This standard, based on the ANSI/NIST-ITL 1-2011, Update: 2015 standard, applies to the FBI's database of biometric and criminal history information (NGI System) and helps ensure that the data format for prints and facial images matches that of the NGI System. Similarly, DOD developed the EBTS, based on the ANSI/NIST-ITL 1-2011, Update: 2015 standard, to interface with DOD's biometric database.

⁷According to the International Organization for Standardization and the International Electrotechnical Commission, standards are established by consensus and

approved by a recognized body that provides, for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context.

Table 1: Data standards for probabilistic genotyping software, latent prints, and face recognition algorithms

Standard	Jurisdiction	Relevant algorithm	Standard developer
INCITS/ISO/IEC 19794 (parts to be superseded by parts of the ISO/IEC 39794 series)	International	Probabilistic genotyping software, latent prints, face recognition	ISO/IEC
ANSI/NIST-ITL 1-2011, Update: 2015	National	Probabilistic genotyping software, latent prints, face recognition	ANSI/NIST
Electronic Biometric Transmission Specification	Agency	Latent prints, face recognition	FBI
Electronic Biometric Transmission Specification	Agency	Latent prints, face recognition	DOD

Source: GAO analysis of Federal Bureau of Investigation (FBI), National Institute of Standards and Technology (NIST), Department of Defense (DOD) and International Organization for Standardization (ISO) documents. | GAO-20-479SP

Legend: INCITS = InterNational Committee for Information Technology Standards, ISO = International Organization for Standardization, IEC = International Electrotechnical Commission, ANSI = American National Standards Institute, ITL = Information Technology Laboratory

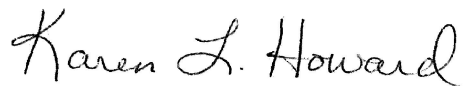
4 Agency and Expert Comments

We provided a draft of this report to the Attorney General of the Department of Justice and the Secretaries of the Departments of Homeland Security, Defense, and Commerce with a request for technical comments. We incorporated agency comments into this report as appropriate.

We invited the 16 participants from our meeting of experts to review our draft report. Among these participants, 7 provided technical comments, which we incorporated as appropriate.

We are sending copies of the report to the appropriate congressional committees, relevant federal agencies, and other interested parties. In addition, the report will be available at no charge on the GAO website at <https://www.gao.gov>.

If you or your staff have any questions concerning this report, please contact me at (202) 512-6888 or howardk@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Key contributors to this report are listed in appendix III.



Karen L. Howard, PhD
Director
Science, Technology Assessment, and Analytics

Appendix I – Objective, Scope, and Methodology

This technology assessment describes forensic algorithms that are being used by federal law enforcement to help associate evidence with civilian individuals and how these technologies work.

To address this research objective, we conducted interviews with relevant federal agencies, including federal law enforcement agencies; convened an interdisciplinary panel of 16 experts with assistance from the National Academies of Sciences, Engineering, and Medicine; conducted interviews with additional stakeholders, including nonprofit groups and legal experts; conducted a literature search; and reviewed relevant literature and case law.

We met with or obtained information from the following federal agencies:

- Department of Justice: Federal Bureau of Investigation, Drug Enforcement Administration, Office of Justice Programs
- Department of Homeland Security: Office of Biometric Identity Management, U.S. Secret Service, Transportation Security Administration, U.S. Customs and Border Protection, U.S. Immigration and Customs Enforcement

- Department of Defense: Defense Forensic Science Center, Defense Forensics and Biometrics Agency, Naval Criminal Investigative Service
- Department of Commerce: National Institute of Standards and Technology

We focused our review on automated or partially automated computer-software-based algorithms used for analyzing forensic evidence during a law enforcement investigation to help establish a link between an individual and collected evidence. This excluded from our scope certain algorithms used by law enforcement, such as ballistics algorithms and digital forensics algorithms, because we found that they may not link an individual to a criminal investigation. Further, we focused our review on algorithms used by federal law enforcement agencies—as opposed to state or local agencies—and on civilian criminal law enforcement.

To conduct the expert meeting, we collaborated with the National Academies to convene a 1½-day meeting of 16 experts on forensic algorithms used by federal law enforcement. We worked with the National Academies' staff to identify experts from a range of stakeholder groups, including federal agencies, academia, and industry. We evaluated the experts for any conflicts of interest.⁸ A conflict of interest was considered to be any current financial or other interest (such as an organizational position) that might

⁸This meeting of experts was planned and convened with the assistance of the National Academy of Sciences to better ensure that a breadth of expertise was brought to bear in its preparation. However all final decisions regarding meeting

substance and expert participation are the responsibility of GAO. Any conclusions and recommendations in GAO reports are solely those of the GAO.

conflict with the service of an individual because it could (1) impair objectivity or (2) create an unfair competitive advantage for any person or organization. The 16 experts were determined to be free of reported conflicts of interest, except those that were outside the scope of the forum or where the overall design of our panel and methodology was sufficient to address them, and the group as a whole was determined to not have any inappropriate biases. (See app. II for a list of these experts and their affiliations.) The comments of these experts generally represented the views of the experts themselves and not the agency, university, or company with which they were affiliated, and are not generalizable to the views of others in the field.

We divided the meeting into five moderated discussion sessions based on key questions we provided on the following topics: (1) overview of forensic algorithms and their operational use; (2) characterizing the accuracy of forensic algorithms; (3) strengths and limitations of forensic algorithms; (4) key issues affecting usage of forensic algorithms; and (5) policy options relevant to the use of forensic algorithms. For sessions two through five, the discussion focused on latent prints, probabilistic genotyping, and face recognition

algorithms. We reported on findings from session one in this technology assessment, and we plan to report on findings from the other sessions in a forthcoming technology assessment. The meeting was transcribed to ensure that we accurately captured the experts' statements. After the meeting, we reviewed the transcripts to characterize their responses and to inform our understanding of forensic algorithms. Following the meeting, we continued to seek the experts' advice to clarify and expand on what we had heard. Consistent with GAO's Quality Assurance Framework, we provided the experts with a draft of our report and solicited their feedback, which we incorporated as appropriate.

We conducted our work from August 2019 through May 2020, in accordance with all sections of GAO's Quality Assurance Framework that are relevant to technology assessments. The Framework requires that we plan and perform the engagement to obtain sufficient and appropriate evidence to meet our stated objectives and to discuss any limitations to our work. We believe that the information and data obtained, and the analysis conducted, provide a reasonable basis for any findings and conclusions in this product.

Appendix II: Expert Meeting Participation

We collaborated with the National Academies of Sciences, Engineering, and Medicine to convene a 1½-day meeting of 16 experts on forensic algorithms used in federal law enforcement. The meeting was held on January 15-16, 2020 in Washington, D.C. Many of these experts provided us with additional assistance throughout our work, including sending additional information for our review or reviewing our draft report for technical accuracy. The experts who participated in this meeting are listed below.

Sarah Chu

Senior Advisor on Forensic Science Policy
Innocence Project

Michael Coble

Associate Director of the Center for Human
Identification
University of North Texas Health Science
Center

Robert English

Special Counsel, Science and Technology
Branch
Federal Bureau of Investigation

Tamara Giwa

Attorney, Assistant Federal Defender
Federal Defenders of New York

Patrick Grother

Scientist, Information Technology Laboratory,
Information Access Division, Image Group
National Institute of Standards and
Technology

William Guthrie

Division Chief, Statistical Engineering Division
National Institute of Standards and
Technology

Karen Kafadar

Commonwealth Professor and Chair of
Statistics
University of Virginia

Dan E. Krane

Professor and Interim Dean
Wright State University

James Loudermilk

Senior Director, Innovation and Customer
Solutions
IDEMIA National Security Solutions

Anne May

Biometric Support Center Program Manager,
Office of Biometric Identity Management
Department of Homeland Security

Mark Perlin

Chief Scientific and Executive Officer
Cybergenetics

Peter M. Vallone

Scientist, Biomolecular Measurement
Division
National Institute of Standards
and Technology

Kit Walsh

Senior Staff Attorney
Electronic Frontier
Foundation

James L. Wayman

Editor-in-Chief
IET Biometrics Journal

Rebecca Wexler

Assistant Professor
University of California, Berkeley School of
Law

Michael Yates

Senior Technical Advisor on Biometrics,
Science and Technology Branch
Federal Bureau of Investigation

Appendix III: GAO Contact and Staff Acknowledgments

GAO contact

Karen L. Howard, PhD (202) 512-6888 or howardk@gao.gov

Staff acknowledgments

In addition to the contact named above, Sushil Sharma (Assistant Director), Allen Chan (Analyst-in-Charge), Mariel Alper, Nora Adkins, Virginia Chanley, Hayden Huang, Eliot Fletcher, Anika McMillon, Eleni Orphanides, and Ben Shouse made key contributions to this report.

(103767)

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<https://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <https://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [Twitter](#), and [YouTube](#).

Subscribe to our [RSS Feeds](#) or [E-mail Updates](#).

Listen to our [Podcasts](#) and read [The Watchblog](#).

Visit GAO on the web at <https://www.gao.gov>.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact: Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548