Document Title: **Evaluability Assessment of Mobile Biometric Facial Recognition Technology Pinellas County, Florida**

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this federally funded evaluability assessment available electronically.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Evaluability Assessment of Mobile Biometric Facial Recognition Technology Pinellas County, Florida

Staff Contact: Scott McCallum
Systems Analyst
Pinellas County, FL
727–453–7193

NIJ Guidance

The National Institute of Justice (NIJ) does not recommend an evaluation of mobile biometric identification in the site assessed below. NIJ remains interested, however, in evaluating the impact of this technology in other sites.

Applicants who propose to evaluate this technology (or other mobile biometric devices) are encouraged to consider the outcome variables (including efficiencies such as reduced time making identifications, increased true positive rates of identification, and decreased false positive rates of identification) and obstacles (including low base rates and unavailable or incomparable control groups) identified below. NIJ encourages applicants to identify sites where randomization is possible or where technology adoption permits pre-post comparison group designs. Finally, NIJ does not wish to pursue research on recognition decisionmaking with this solicitation.

Applicants may depart from this guidance by providing appropriate rationale.

Project Summary: The mobile biometric technology examined for this evaluation feasibility assessment was the use of facial recognition by the Pinellas County, Florida, Sheriff's Department. At present 50 patrol units are equipped with off-the-shelf digital cameras to photograph suspects, other individuals in the field who cannot provide valid identification, or those suspected of providing false identification. Although just 50 units with cameras are available across all of the shifts, other officers can and do request assistance from the cars with cameras when identification is needed. Thus, the department believes that few opportunities to photograph unknown suspects are missed.

Under those circumstances, deputies ask permission of the suspect to take his or her photograph. Local law enforcement officials report that refusals are quite rare, and besides, State law permits photographing individuals in public places. After the photograph is taken, a digital image is uploaded to the department's Mobile Identification System (MIS) via a camera docking station or USB connection through laptop computers in the patrol units. These digital images are then electronically compared to more than 1 million digital mug shots of individuals previously arrested in Pinellas County. Recently the department has partnered with neighboring jurisdictions and the State Department of Corrections to add millions of additional images to the database to improve identifications.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The MIS uses facial recognition algorithms developed by the technology vendor, Viisage, to produce rank-ordered galleries of likely matches. These picture galleries are then simultaneously displayed on the patrol laptop screen. In addition to the photographic images, demographic data are provided for arrestees shown in the gallery, including address, age, identifying features, and other personal information. The deputy uses both the visual and demographic information to attempt to identify the unknown individual. The facial recognition technology is a tool for the investigating deputy; it does not make positive match determinations on its own.

Scope of Evaluation: Several evaluation options exist: (1) a post-only outcome case study limited to Pinellas County; (2) a pre-post comparison group study in another agency that is just beginning to implement the use of mobile facial recognition technology; or (3) a true experiment involving officer or deputy recognition decision making.

Summary of Evaluability Assessment Activity: The assessment of the feasibility of evaluating mobile biometric technologies began with a literature review and a Web-based search to identify vendors of electronic biometric identification technologies. The researchers then attempted telephone interviews with 12 known biometric vendors with limited success. The researchers also interviewed technology experts at the National Law Enforcement and Corrections Technology Centers (NLECTC), and held conference calls with NIJ Program Managers from the Office of Research and Evaluation and the Office of Science and Technology.

The literature review, telephone interviews, and conference calls revealed that mobile biometric technologies are relatively new to the field of law enforcement and are used only by a handful of agencies. A variety of such technologies are available, including facial, iris, retinal, automated fingerprint identification system (AFIS), and voice recognition. The most emergent and mature technologies appear to be facial recognition and mobile AFIS. However, very little is known empirically about the effects of mobile biometric identification technology.

The Urban Institute's (UI's) initial screening identified three mature applications of mobile biometric identification technology. These were found in Pinellas County, Florida; Hennepin County, Minnesota; and San Joaquin County, California. Pinellas County began implementing facial recognition under a 2001 grant from the Office of Community Oriented Policing Services (COPS Office) and now has 50 operational units. Hennepin County began implementing mobile AFIS in 2002 and currently has 100 operational units. San Joaquin is planning to expand its limited AFIS application this fall with an additional 55 new units.

On the basis of the screening information compiled, UI and NIJ mutually decided that Pinellas County, Florida, would be the location for a further site visit screening. Pinellas County was also selected as a site for global positioning system (GPS) offender tracking site screening, which is discussed in a separate assessment report.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Finding: A scientifically rigorous outcome evaluation of the Pinellas County application would be difficult. Current deployment precludes randomization, and naturally occurring comparison groups are not present. It may be possible to conduct a pre-post comparison group study in another agency that is just beginning biometric implementation. However, an opportunity exists to implement a randomized laboratory-type experiment of deputy decision-making using this technology.

1. Brief Literature Review

What do we already know about projects like these? Would this evaluation add to what we know?

State and local law enforcement agencies have a critical need for accurate mobile identification of individuals. When officers encounter persons unknown to them they may need to ascertain whether those persons have outstanding warrants, have suspended or revoked driving privileges, are gang members, have been reported missing, or may be dangerous based upon past behaviors or a criminal record. Until relatively recently, the only means available in the field to meet these needs was to rely on identification carried by potential suspects, such as driver's licenses. Unfortunately, the police often encounter individuals without identification cards or with falsified ones.

In order to solve these problems in the field, mobile biometric technologies have recently been designed for police use. These include facial, AFIS, iris, retinal, and voice recognition, among other technologies. According to the literature, the most emergent of these mobile law enforcement solutions to date, have been facial recognition and mobile AFIS.

Recent tests of prototype mobile facial recognition biometrics have included their use by the Los Angeles Police Department's gang task force and by patrol deputies with the Pinellas County Sheriff's Office. In Los Angeles, a handheld 1.3 megapixel Neven-Vision Mobile Identifier is used to take a digital photograph of a suspect and compare it to an existing database of similar images from the field (Trask, 2006). The manufacturer reports a 99-percent positive identification rate for its technology with the first or second comparison photograph yielding a positive match (Siuru, 2006).

Pinellas County uses off-the-shelf digital cameras to take photographs in the field and submit them for match wirelessly from their patrol cars. Comparisons against booking photographs are reportedly returned within 30 seconds (Simon, 2005). During the first 3 years of use officers were able to identify 53 wanted felons who gave them false names. They were also able to correctly identify 57 individuals who were suspected of having warrants, but in fact did not (Simon, 2005).

Recent examples of uses of mobile AFIS technology come from Ontario, California, and Hennepin County, Minnesota. In California, a handheld fingerprint scanner was used in the field more than 3700 times during the first 6 months of 2003, resulting in successful identification of 816 individuals and detention of 164 of them. In Minnesota, deputies

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

used the system "679 times, identifying 110 individuals and detaining 37." (NLECTC, 2004). Comments from law enforcement officials about this technology were positive in both jurisdictions. Other benefits cited were lightweight and easy-to-use scanners, increased information sharing, and increased cross-jurisdictional information sharing (Justice Technology Information Network, 2005; NLETC, 2004).

Little empirical evidence of outcome effectiveness exists for mobile biometric technologies, either facial recognition or AFIS.

What audience would benefit from this evaluation?

The primary beneficiaries would be law enforcement policymakers, administrators, and investigators. An evaluation would also contribute significantly to empirical knowledge about the use of technology to aid in decision-making and to improve efficiency and effectiveness outcomes, which would benefit the research community. Federal funding agencies would also find the results of an evaluation useful for policy and program development.

2. Level of Site Cooperation

Pinellas County voiced a willingness to cooperate in an evaluation.

Is there local interest in being evaluated?

Pinellas County also voiced an interest in being evaluated.

Is there a local evaluation?

There has been no formal evaluation to date and none is currently planned.

3. Background History

Implementation of this technology began in 2001 with the acquisition of a COPS Office grant. It is estimated that to date more than $7 million has been invested in this facial recognition software and hardware (including patrol-unit laptop computers). However, add-on unit costs are relatively small at $1,500 per new camera and software license. Current plans are to expand the use of this technology to include six other law enforcement agencies in the region as facial recognition technology partners. Support for this expansion is reportedly coming from a Department of Defense earmark appropriation.

4. Program Design

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Target Population

The target population is suspects or other individuals encountered by law enforcement officers in the field who cannot be identified or who present false identification.

Project Goals and Objectives

The goals of the use of this technology are to improve the efficiency and effectiveness of field identification of unknown persons. The objectives are to: 1) increase the apprehension of wanted persons; 2) decrease the amount of time required to identify unknown persons; and 3) reduce the number of mistaken identifications, thereby clearing innocent persons thought to be wanted.

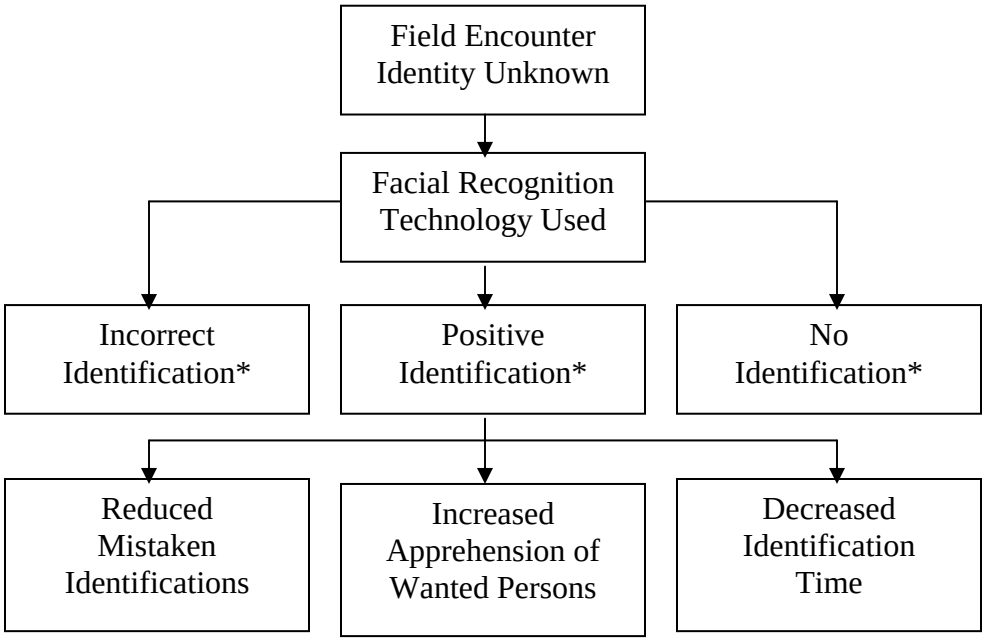
5. Program Logic Model

Describe the logic that connects project activities to project goals.

Exhibit 1 presents the basic technology logic model.

As this logic model shows, users of facial recognition technology hypothesize that it results in three primary effects that are consistent with program goals and objectives: 1) less time spent identifying unknown individuals; 2) fewer mistaken releases of those wanted; and 3) increased apprehension of suspects, particularly those with outstanding warrants.

Exhibit 1. Facial Recognition Logic Model



* Intermediate decisionmaking outcomes.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

However, careful consideration of this logic model shows intermediate outcomes of the employment of this technology as well. A deputy or officer must decide, based on the photograph array and background information provided by the technology, whether a match exists or not. The hypothesized positive efficiency and effectiveness outcomes are contingent on valid and reliable identification decisions by an individual deputy or officer in the field. It is assumed that technology utilization results in positive identification decisions, but this assumption has not been empirically tested.

Is the logic supportable by empirical evidence?

The only empirical evidence at present includes descriptive findings of the numbers of individuals apprehended or released for warrants following the use of this technology. Other outcomes have not yet been documented, nor have current findings been compared to other identification approaches in the field.

Are there apparent contradictions or conflicts between certain activities and the outcome expected?

The use of this technology as a tool for law enforcement to improve the efficiency and effectiveness of the identification of unknown persons appears logical. However, these outcomes are dependent on individual officer decision making when comparing suspects to a simultaneous mug shot array. Previous research on the use of mug shots for lineup identification suggests that simultaneous aggregate arrays may lead to more identification errors than sequential comparisons. Whether or not this is the case with this application has not been investigated, nor has the fundamental assumption that patrol officers can make accurate and consistent identification decisions using this technology been tested. In the absence of such knowledge, attribution of efficiency and effectiveness to use of the technology could be questioned.

6. Implementation Issues

Is the project being implemented as planned?

Yes, according to field interviews. However, the assessment team did not secure the original 2001 program plans.

Describe staffing.

Currently 50 units are deployed across all patrol shifts, and officers receive 4 hours training in the use of this technology. It is estimated that between 100 and 150 uses take place each month with approximately 15 successful identifications resulting.

Describe the stability of the project over time.

The implementation of this technology is mature. Current plans are to expand the use of this technology to include six other law enforcement agencies in the region as facial

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

recognition partners. Support for this expansion is reportedly coming from a Department of Defense earmark appropriation.

What aspects of the project could be evaluated for outcome?

Although the focus of this evaluability assessment has been on the feasibility of rigorous outcome designs, it was apparent during our screenings and site visit that a case study process evaluation also has potential merit. Pinellas County has been involved in the implementation and adaptation of facial recognition technology for 5 years; the lessons learned during this period would be of significant value to law enforcement policymakers and practitioners who are considering similar field applications in the future. Of particular value would be a detailed exploration into implementation costs and their relationship to perceived and actual benefits. As noted earlier, estimated implementation costs for a relatively limited field application of this technology reportedly have exceeded $7 million. If similar startup costs are likely to be required, it seems unlikely that it will be adopted elsewhere in the absence of sizable external funding, no matter how positive the results of any outcome evaluation. Alternatively, the partnerships and collaborative efforts of this jurisdiction's experience may suggest implementation successes and funding alternatives for other law enforcement policymakers and practitioners.

There are several outcome designs worth considering for an evaluation of biometric facial recognition technology for law enforcement. If the focus of an evaluation will be on the perceived efficiency and effectiveness outcomes as expressed by local stakeholders in Pinellas County, then the most feasible design appears to be a post-only case study design. This is due in large part to the maturity of implementation at this site and the inability to introduce randomization. Baseline measurement of efficiency and effectiveness before implementation would be similarly extremely difficult, if not impossible. A post-only comparison group design would be theoretically feasible using another sheriff's department as a comparison group.

An alternative and more rigorous outcome evaluation design that also appears feasible is a pre-post comparison group design in a site that is currently planning to implement this solution. Under this design, baseline preimplementation measurements could be made on the hypothesized outcomes. Two design options exist for comparison areas. One design would restrict implementation to randomly selected areas of the jurisdiction (precincts or districts, for example). Those areas would become the experimental areas and the other areas would be controls. Another alternative for this design would be implementing the technology throughout a department, which would be the experimental department, and using a similar department without mobile biometric technology as a control. Under both options, pre-post and longitudinal time series outcomes could be compared. Based upon the experience of Pinellas County, it does not appear that random assignment in the field would be possible.

The final option would be to focus on what are hypothesized to be the intermediate outcomes. As described in the outcome measures section and shown in the basic logic model (exhibit 1), these outcomes would be the accuracy and consistency of individual

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

officer identification decisions. Not only would this option provide valuable outcome information for policymakers and practitioners considering the use of this technology, but it would be an extremely valuable contribution to the growing and hotly debated field of research about suspect identification generally.

What would the outcome measures be?

Efficiency and effectiveness outcome measures would include mistaken identifications, apprehension of wanted persons, and time required to identify unknown persons. Intermediate outcome measures include accuracy and consistency of officer identification decision-making using this technology. These might include identifications made under differing experimental conditions such as type of array presentation (simultaneous or sequential, known matches and fillers).

How could an appropriate comparison group be created?

Naturally occurring comparison events appear to occur very infrequently given backup deployment of equipped units in Pinellas County. But, in an efficiency and effectiveness study, internal comparison groups could be created in a startup agency by restricting use of the facial recognition technology to selected geographic areas within the agency's jurisdiction and using other areas within the jurisdiction as comparison groups. Alternatively, another similar law enforcement agency could be recruited for comparison purposes. In a study of intermediate decision making outcomes, experimental and control events could be created randomly under laboratory conditions.

Are the sample sizes statistically significant?

Given the relatively infrequent uses of this technology (50–100 per month) and small number of positive identifications (15 per month), sampling of events would not be required. However, the relatively low number of positive identifications per month suggests that detection of outcome effect sizes will be somewhat difficult.

An experiment focusing on intermediate decisionmaking outcomes is also feasible using the entire population of events; sampling would not be required. A laboratory experiment could be implemented under controlled circumstances to generate a number of test events far in excess of the actual number of field events to maximize the detection of the effects of technology and other factors on intermediate decisionmaking.

Is random assignment possible?

Not for a post-only or comparison group efficiency and effectiveness outcome design. However, for an experimental study to evaluate officer identification decisions, suspect images, filler images, and display methods (simultaneous vs. sequential) could be randomly generated under laboratory conditions.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Recommended Approach

It is recommended that NIJ support a study of officer/deputy identification decisionmaking using facial recognition technology. This is important, given that the current outcome logic model assumes, without known empirical support, that officers and deputies can make correct identification matches using the technology as currently deployed. In addition, such a study would be a significant contribution to the field of knowledge about identification of individuals using photographic evidence. Finally, this design is the most rigorous approach and the one least likely to result in findings subject to alternative explanations.

Alternative Approach

An alternative approach would be to implement a post-only case study to inform future research in this area. In addition, a pre-post comparison group design is feasible in another jurisdiction just beginning implementation of facial recognition technology. Unfortunately, the initial assessment screening did not identify any prospective agencies that might be considering the implementation of facial recognition technology, although others, such as San Joaquin County, are considering other biometric approaches such as AFIS.

What strengths and weaknesses do the designs have?

The primary strength of a decisionmaking design would be its scientific rigor. The major weakness would be implementation costs and labor intensity associated with maintenance of the design over time.

The efficiency and effectiveness designs suffer from the typical threats to validity associated with preexperimental approaches. Their primary strength is the generation of knowledge on which to base future research efforts in an area where very little is known from a social science perspective.

How long in duration would the evaluation be?

It is estimated that a randomized officer decisionmaking evaluation could be accomplished within 18 months. Approximately the same duration would be required for a post-only case study of efficiency and effectiveness in Pinellas County, Florida. A pre-post comparison group study of efficiency and effectiveness in a jurisdiction just beginning the implementation of facial recognition technology would likely take an additional 6 months or more. This would be primarily due to the extra time required to identify and recruit a comparison area or agency and to collect and analyze new data.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

What would be the estimated cost?

An experimental laboratory study of decisionmaking would be the most expensive, likely in excess of $375,000. This is because this design would require extensive site-based control and monitoring of experimental conditions, not to mention likely vendor programming costs to be able to generate mug shot comparisons in both simultaneous and sequential arrays. A pre-post comparison group study of a new application is estimated to require $325,000–350,000 because of new data collection requirements and startup recruitment costs associated with use of a comparison agency. The least costly would be the post-only case study of Pinellas County. Some new data would need to be collected, but current data systems are quite good. Estimated costs for this approach would be in the $175,000–200,000 range.

What aspects of the project make an evaluation more difficult?

The major challenges of an experimental study of officer decision-making would be modification of the existing technology, maintenance of the actual experimental design on site, and securing officer/deputy time to participate in the experiment white on duty or on overtime. For the pre-post comparison group design, site recruitment, data access, and gaining buy-in for an evaluation, particularly from the comparison agency, could present obstacles. A post-only case study would be the least challenging, but would still require agency and researcher data collection demands.

7. Measurement Model

The efficiency and effectiveness outcomes and intermediate decision making outcome measures are summarized in the logic model (exhibit 1). These include accuracy and consistency of officer or deputy identification decisions that result in the apprehension of wanted suspects, release of suspects not wanted, and decreased identification time.

8. Data

Comment on the quality and availability of project-generated data to support these measures.

The Pinellas County Sheriff’s Department maintains comprehensive and sophisticated electronic databases that can be used for evaluation purposes. These include outcomes of field technology use, arrest and booking data (including digital mug shots) back to 1994, calls for service, incident reports (records management system), and computer-aided dispatch. The department maintains all these data itself, and none of the systems’ vendors control access, as is sometimes the case in other law enforcement agencies.

Can services delivered be identified?

Delivery of services is not an element of this technology application.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Can target population be tracked over time?

The current population of technology use events can be tracked over time.

Would an evaluation have to generate new or additional data?

Regardless of the design employed, new additional data would have to be collected. New data collection would be most burdensome for the decisionmaking experimental design and least demanding for a post-only efficiency/effectiveness design.

10. Summary Remarks

Recommendations for Evaluation

It is recommended that an intermediate outcome decisionmaking evaluation be considered. Knowing how well this technology helps in making correct decisions can not only contribute to knowledge about the effectiveness of this particular technology but also to the broader research on identification of unknown persons more generally.

References

Justice Technology Information Network. (2005). "NEC Witness Serious Success With Automated Fingerprint Identification System Deployment for Tacoma Police Department," *Business Wire* (January 10). Retrieved August 29, 2006, from http://www.nlectc.org/justnetnews/01132005.html.

National Law Enforcement and Corrections Technology Centers (2004). "IBIS: Fingering the Felon," *TECHbeat* (Spring).

Simon, Sam. (2005). "About Face Recognition," *Law Enforcement Technology* 32(7)(July).

Siuru, Bill. (2006) "The Mobile Identifier—Instantaneous Mug Shots on Demand," *Police and Security News* 22(3)(June).

Trask, Colin. (2006). "From Handprints to Heartbeats: Acceptance Grows for Biometrics," *Kiosk Marketplace*. Retrieved August 29, 2006 from http://www.kioskmarketplace.com/article.php?id=15580