U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

The U.S. Department of Justice, Office of Justice Programs, National Institute of Justice is seeking applications for funding evaluations of criminal justice technologies. This program furthers the Department's mission by sponsoring research to provide objective, independent, evidence-based knowledge and tools to meet the challenges of crime and justice, particularly at the State and local levels.

Solicitation: Criminal Justice Technology Evaluation

Eligibility

(See "Eligibility," page 3)

Deadline

All applications are due **January 30, 2007, 11:59 p.m. eastern time.**

Contact Information

For assistance with the requirements of this solicitation, contact Brett Chapman, Social Science Analyst, Crime Control and Prevention Research Division, Office of Research and Evaluation, 202–514–2187, Brett.Chapman@usdoj.gov.

This application must be submitted through Grants.gov. For technical assistance with submitting the application, call the Grants.gov Customer Support Hotline at 1–800–518–4726.

Grants.gov Funding Opportunity No. 2007–NIJ–1415

SL# 000781

CONTENTS

Criminal Justice Technology Evaluation
CFDA No. 16.560

Overview

The National Institute of Justice (NIJ) is the research, development, and evaluation agency of the U.S. Department of Justice and a component of the Office of Justice Programs (OJP). NIJ provides objective, independent, evidence-based knowledge and tools to enhance the administration of justice and public safety. NIJ solicits proposals to inform its search for the knowledge and tools to guide policy and practice.

NIJ seeks evaluations of technologies designed to prevent and reduce crime, and improve the functioning of the criminal justice system. NIJ is specifically interested in evaluations that determine if the application of existing or newly developed technologies works to improve outcomes. Evaluations should focus on technologies that are fully functional and readily applicable in operational settings. Additionally, outcome evaluations must have well-defined, measurable indicators of success; proposals that focus on the measurable impact of technologies on reduction in crime will be given priority. Additional outcome measures may include better monitoring, tracking, or identification and apprehension of criminal suspects; effective tracking and recovery of stolen property; reduced criminal opportunities; and reduced injuries to officers, suspects, and/or bystanders.

Deadline: Registration

Registering with Grants.gov is a one-time process; however, if you are a first time registrant it could take up to several weeks to have your registration validated and confirmed and to receive your user password. Start the registration process early to prevent delays that may cause you to miss the application deadline. You must complete these three steps before you are able to register: 1) Register with Central Contractor Registry (CCR), 2) Register yourself as an Authorized Organization Representative (AOR), and 3) Be authorized as an AOR by your organization. For more information, visit www.grants.gov. **Note: Your CCR Registration must be renewed once a year. Failure to renew your CCR registration may prohibit submission of a grant application through Grants.gov.**

Deadline: Application

The due date for applying for funding under this announcement is **January 30, 2007, 11:59 p.m. eastern time.**

Eligibility

In general, NIJ is authorized to make grants to, or enter into contracts or cooperative agreements with, States (including territories), local governments (including federally-recognized Indian tribal governments that perform law enforcement functions), nonprofit organizations (including faith-based and community organizations), profit organizations, institutions of higher education, and qualified individuals. Foreign governments or organizations are not eligible to apply.

Faith-Based and Other Community Organizations: Consistent with President George W. Bush's Executive Order 13279, dated December 12, 2002, and 28 C.F.R. Part 38, it is DOJ policy that faith-based and other community organizations that statutorily qualify as eligible applicants under DOJ programs are invited and encouraged to apply for assistance awards to fund eligible grant activities. Faith-based and other community organizations will be considered for awards on the same basis as other eligible applicants and, if they receive assistance awards, will be treated on an equal basis with all other grantees in the administration of such awards. No eligible applicant or grantee will be discriminated for or against on the basis of its religious character or affiliation, religious name, or the religious composition of its board of directors or persons working in the organization.

Faith-based organizations receiving DOJ assistance awards retain their independence and do not lose or have to modify their religious identity (e.g., removing religious symbols) to receive assistance awards. DOJ grant funds, however, may not be used to fund any inherently religious activity, such as prayer or worship. Inherently religious activity is permissible, although it cannot occur during an activity funded with DOJ grant funds; rather, such religious activity must be separate in time or place from the DOJ-funded program. Further, participation in such activity by individuals receiving services must be voluntary. Programs funded by DOJ are not permitted to discriminate in the provision of services on the basis of a beneficiary's religion.

Applicants are encouraged to review the Civil Rights Compliance section under "Additional Requirements" in this announcement.

Specific Information—Criminal Justice Technology Evaluation

NIJ will consider applications that seek to evaluate applied technologies that prevent and reduce crime, and improve the functioning of the criminal justice system. In preparation for this solicitation, NIJ reviewed a wide range of technologies applicable to criminal justice. NIJ commissioned evaluation assessments of the following eight technologies and determined that they are strong candidates for a full-scale evaluation:

- Alcohol monitoring of offenders under supervised release.
- Offender tracking systems.
- Radio Frequency Identification (RFID) technologies.
- Mobile identification biometric devices.
- GPS-based Automated Vehicle Locator technologies.
- Automated License Plate Recognition systems.
- Pawnbroker databases.
- Trace detection technologies for narcotics, explosives, and other contraband.

(Appendix A presents a summary of these evaluation assessments.)

NIJ encourages applicants to consider submitting proposals in one of these technology areas. However, other appropriate technologies may also be considered for evaluation.

In all cases, proposals should provide specific evaluation information, including the name and description of the technology to be evaluated and where the technology has been implemented.

Applicants submitting proposals that involve technology areas other than those summarized in the assessments cited above must demonstrate how the technology can be evaluated. In these cases, applicants should provide specific answers to the following questions:

- What is the history of the technology?
- At what stage of implementation is the technology?
- Is the technology fully functional and readily applicable in an operational criminal justice setting?
- What outcomes are expected?
- What outcomes will be assessed?
- Is random assignment possible? If not, how will comparison groups be formed?
- Is the technology being used on a scale large enough to provide adequate statistical power to an evaluation?
- What evaluation evidence currently exists to justify an outcome evaluation?

NIJ gives special consideration to research and evaluation that entails proposing randomized experiments, regression continuity, propensity scoring, and other enhanced quasi-experimental designs that maximize validity and reliability.

What will *not* be funded:
1. Provision of training or direct service.
2. Proposals primarily to purchase equipment, materials, or supplies. (Your budget may include these items if they are necessary to conduct applied research, development, demonstration, evaluation, or analysis, but NIJ does not fund proposals that are primarily to purchase equipment.)
3. Work that will be funded under another specific solicitation.
4. Proposals to develop technology or test technology in a laboratory or other artificial setting. Development or testing of technologies is not the focus of this solicitation. NIJ only seeks proposals which will evaluate the impact of new or existing technologies in operational settings.

Cost of proposed work: NIJ anticipates that up to $1 million may become available for awards made through this solicitation. All NIJ awards are subject to the availability of appropriated funds and to any modifications or additional requirements that may be imposed by law. NIJ expects to make two to four awards depending on funds available and number of high-quality applications. If you propose a project that exceeds the amount of money that may be available for this solicitation, we recommend that you divide the project into phases, stages, or tasks so that NIJ can consider making an award for a specific portion of the work. NIJ cannot guarantee that subsequent phases, stages, or tasks will be funded. Such additional funding depends on NIJ's resources and your satisfactory completion of each phase, stage, or task. Note: Deliverables (e.g., a final report) will be required at the end of each phase, stage, or task.

A grant made by NIJ under this solicitation may account for up to 100 percent of the total cost of the project. See "Cofunding," under "What an Application Must Include."

Performance Measures

To assist in fulfilling the Department's responsibilities under the Government Performance and Results Act (GPRA), P.L. 103-62, applicants who receive funding under this solicitation must provide data that measures the results of their work. Performance measures for this solicitation are as follows:

Objective	**Performance Measures**	**Data Grantee Provides**
Develop and analyze information and data having clear implications for criminal justice policy and practice.	1. Relevance to the needs of the field as measured by whether the grantee's substantive scope did not deviate from the funded proposal or any subsequent agency modifications to the scope. 2. Quality of the research as assessed by peer reviewers. 3. Quality of management as measured by whether significant interim project milestones were achieved, final deadlines were met, and costs remained within approved limits.	1. A final report providing a comprehensive overview of the project and a detailed description of the project design, data, and methods; a full presentation of scientific findings; and a thorough discussion of the implications of the project findings for criminal justice practice and policy. 2. Quarterly financial reports, semi-annual progress reports, and a final progress report.

How to Apply

DOJ is participating in the e-Government initiative, one of 25 initiatives included in the President's Management Agenda. Part of this initiative—Grants.gov—is a "one-stop storefront" that provides a unified process for all customers of Federal grants to find funding opportunities and apply for funding.

Grants.gov Instructions: Complete instructions can be found at http://www.grants.gov/applicants/get_registered.jsp. If you experience difficulties at any point during this process, please call the Grants.gov Customer Support Hotline at 1–800–518–4726.

CFDA Number: The Catalog of Federal Domestic Assistance (CFDA) number for this solicitation is 16.560, titled "National Institute of Justice Research, Evaluation, and

Development Project Grants," and the Grants.gov funding opportunity number is 2007–NIJ–1415.

A DUNS number is required: The Office of Management and Budget requires that all businesses and nonprofit applicants for Federal funds include a DUNS (Data Universal Numeric System) number in their application for a new award or renewal of an award. Applications without a DUNS number are incomplete. A DUNS number is a unique nine-digit sequence recognized as the universal standard for identifying and keeping track of entities receiving Federal funds. The identifier is used for tracking purposes and to validate address and point of contact information. The DUNS number will be used throughout the grant life cycle. Obtaining a DUNS number is a free, simple, one-time activity. Obtain one by calling 1–866–705–5711 or by applying online at www.dnb.com/us. Individuals are exempt from this requirement.

What an Application Must Include

Standard Form 424

Program Narrative
The Program Narrative includes:

a. Abstract (not to exceed 400 words).
b. Table of contents.
c. Main body, which includes:
  - Purpose, goals, and objectives.
  - Review of relevant literature.
  - Research design and methods.
  - Implications for policy and practice.
  - Management plan and organization.
  - Dissemination strategy.

d. Appendixes (not counted against program narrative page limit) include:
  - Bibliography/References (if applicable).
  - List of key personnel (required).
  - Résumés of key personnel (required).
  - List of previous and current NIJ awards (required).
  - Letters of cooperation/support or administrative agreements from organizations collaborating in the project (if applicable).
  - Chart for timeline, research calendar, or milestones (required).
  - Other materials required by the solicitation.

Budget Detail Worksheet
Templates for filling out the Budget Detail Worksheet may be found online at www.ojp.usdoj.gov/Forms/budget_fillable.pdf, OJP Standard Forms & Instructions. If you have any questions, please contact the Office of the Comptroller's Customer Service Center at 1–800–458–0786.

Budget Narrative

Indirect Rate Agreement (if applicable)

- Applicants that do not have a federally negotiated indirect cost rate and wish to establish one, can submit a proposal to their "cognizant" Federal agency. Generally, the cognizant federal agency is the agency that provides the preponderance of direct federal funding. This can be determined by reviewing an organization's schedule of federal financial assistance. If DOJ is your cognizant federal agency, obtain information needed to submit an indirect cost rate proposal at www.ojp.usdoj.gov/oc/indirectcosts.htm.

Other Program Attachments
These include several forms, available on OJP's funding page at www.ojp.usdoj.gov/forms.htm.

Page limit: The program narrative section of your proposal must not exceed 30 double-spaced pages in 12-point font with 1-inch margins. Abstract, table of contents, charts, figures, appendixes, and government forms do not count toward the 30-page limit for the narrative section.

Cofunding: A grant made by NIJ under this solicitation may account for up to 100 percent of the total cost of the project. You must indicate whether you believe it is feasible for you to contribute cash, facilities, or services as non-Federal support for the project. Your proposal should identify generally any such contributions that you expect to make and your proposed budget should indicate in detail which items, if any, will be supported with non-Federal contributions.

Selection Criteria

Successful applicants must demonstrate the following:

Understanding of the problem and its importance.

Quality and technical merit.

1. Awareness of the state of current research or technology.
2. Soundness of methodology and analytic and technical approach.
3. Feasibility of proposed project and awareness of pitfalls.
4. Innovation and creativity (when appropriate).

Impact of the proposed project.

1. Potential for significant advances in scientific or technical understanding of the problem.
2. Potential for significant advances in the field.
3. Relevance for improving the policy and practice of criminal justice and related agencies and improving public safety, security, and quality of life.
4. Affordability and cost-effectiveness of proposed end products, when applicable (e.g., purchase price and maintenance costs for a new technology or cost of training to use the technology).
5. Perceived potential for commercialization and/or implementation of a new technology (when applicable).

Capabilities, demonstrated productivity, and experience of applicants.

1. Qualifications and experience of proposed staff.
2. Demonstrated ability of proposed staff and organization to manage the effort.
3. Adequacy of the plan to manage the project, including how various tasks are subdivided and resources are used.
4. Successful past performance on NIJ grants and contracts (when applicable).

Budget.

1. Total cost of the project relative to the perceived benefit.
2. Appropriateness of the budget relative to the level of effort.
3. Use of existing resources to conserve costs.

Dissemination strategy.

1. Well-defined plan for the grant recipient to disseminate results to appropriate audiences, including researchers, practitioners, and policymakers.
2. Suggestions for print and electronic products NIJ might develop for practitioners and policymakers.

Relevance of the project for policy and practice:

Higher quality proposals clearly explain the practical implications of the project. They connect technical expertise with criminal justice policy and practice. To ensure that the project has strong relevance for policy and practice, some researchers and technologists collaborate with practitioners and policymakers. You may include letters showing support from practitioners, but they carry less weight than clear evidence that you understand why policymakers and practitioners would benefit from your work and how they would use it. While a partnership may affect State or local activities, it should also have broader implications for others across the country.

Review Process

NIJ is firmly committed to the competitive process in awarding grants. All proposals under this solicitation will be subjected to independent peer-review panel evaluations. External peer-review panelists consider both technical and programmatic merits. Panelists are selected based on their expertise in subject areas pertinent to the proposals.

Peer-review panelists will evaluate proposals using the criteria listed above. NIJ staff then make recommendations to the NIJ Director. The Director makes award decisions.

Reasons for rejection: NIJ may reject applications that are incomplete, do not respond to the scope of the solicitation, do not comply with format requirements, or are submitted after the deadline. No additions to the original submission are allowed.

When awards will be made: All applicants, whether they are accepted or rejected, will be notified. The review and approval process takes about 6 months. You should not propose to begin work until at least 6 months after the proposal deadline on the cover of this solicitation. Also, you should not expect to receive notification of a decision for at least 6 months after that date. Lists of awards are updated regularly on NIJ's Web site at www.ojp.usdoj.gov/nij/funding.htm.

Additional Requirements

- Civil Rights Compliance
- Confidentiality and Human Subjects Protections regulations
- Anti-Lobbying Act
- Financial and Government Audit Requirements
- National Environmental Policy Act (NEPA) compliance
- DOJ Information Technology Standards
- Single Point of Contact Review
- Non-supplanting of State or Local Funds
- Criminal Penalty for False Statements
- Compliance with Office of the Comptroller Financial Guide
- Suspension or Termination of Funding
- Nonprofit Organizations
- Government Performance and Results Act (GPRA)
- Rights in Intellectual Property

We strongly encourage you to review the information pertaining to these additional requirements prior to submitting your application. Additional information for each can be found at www.ojp.usdoj.gov/funding/otherrequirements.htm.

If your proposal is funded, you will be required to submit several reports and other materials, including:

Final substantive report: The final report should be a comprehensive overview of the project and should include a detailed description of the project design, data, and methods; a full presentation of scientific findings; and a thorough discussion of the implications of the project findings for criminal justice practice and policy. It must contain an abstract of no more than 400 words and an executive summary of no more than 2,500 words.

A draft of the final report, abstract, and executive summary must be submitted 90 days before the end date of the grant. The draft final report will be peer reviewed upon submission. The reviews will be forwarded to the principal investigator with suggestions for revisions. The author must then submit the revised final report, abstract, and executive summary by the end date of the grant. The abstract, executive summary, and final report must be submitted in both paper and electronic formats.

For program evaluation studies, the final report should include a section on measuring program performance. This section should outline the measures used to evaluate program effectiveness, modifications made to those measures as a result of the evaluation, and recommendations regarding these and other potential performance measures for similar programs. (This information will be particularly valuable to NIJ and other Federal program agencies in implementing performance measures for federally funded criminal justice programs.)

Interim reports: Grantees must submit quarterly financial reports, semi-annual progress reports, a final progress report, and, if applicable, an annual audit report in accordance with Office of Management and Budget Circular A-133. Future awards and fund drawdowns may be withheld if reports are delinquent.

Appendix A: Technology Evaluability Assessments

Introduction

The National Institute of Justice (NIJ) routinely conducts evaluability assessments as part of its evaluation planning process. The assessments serve multiple purposes. They help NIJ decide which programs or tools it should evaluate and also inform staff on the kinds of evaluation strategies likely to succeed. NIJ also provides the assessments to all potential applicants to help them make more informed decisions about whether they should write proposals.

In some previous solicitations, applicants were constrained to evaluating the programs that had been assessed. **That is not the case here.** Applicants may propose to evaluate the technologies provided here or any other public safety technology.

NIJ has included a total of eight assessments, each prefaced with some general guidance about NIJ's interests for each technology. Applicants should consider NIJ's guidance as they frame their designs but are free to propose other research questions that they believe are relevant and answerable.

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.

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.

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 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

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.

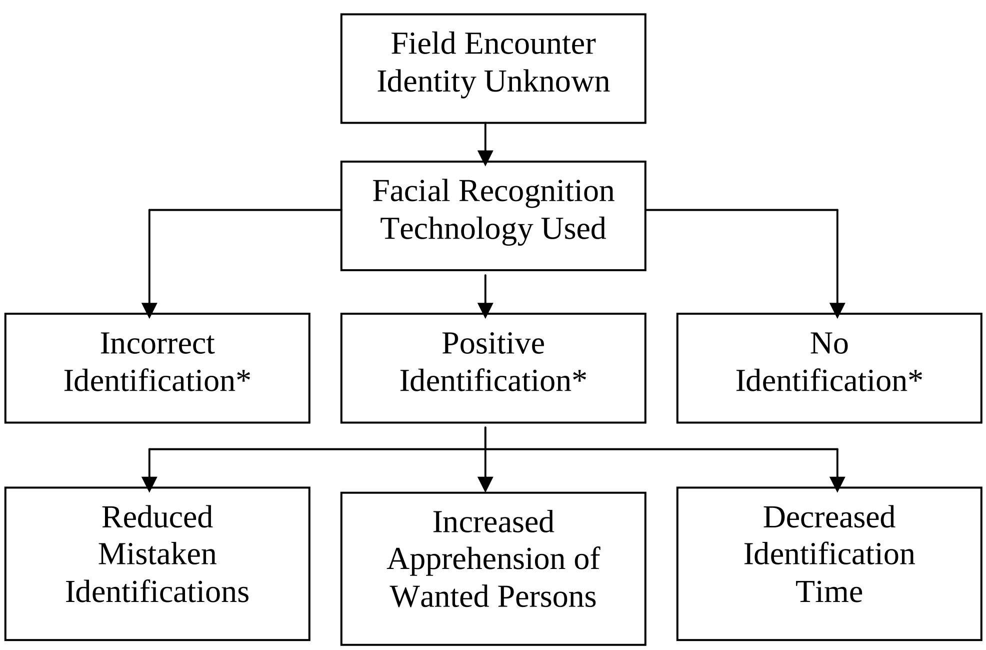
5. Program Logic Model

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.

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 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 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.

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.

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.

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

Evaluability Assessment of Global Positioning System (GPS)-Dispatching in Law Enforcement

Staff Contact: Brian Starr
Applications Developer
Sacramento Police Department
916–433–0407
bstarr@pd.cityofsacramento.org

NIJ Guidance

The National Institute of Justice (NIJ) has identified some key outcome variables and other parameters of interest for this technology, and has provided some guidance on possible evaluation designs. Applicants may depart from this guidance by providing an appropriate rationale.

The evaluability assessment identifies some key questions for outcome evaluations of these systems, and many of the methodological challenges. The assessment cites the advertised benefits of Global Positioning System (GPS)/Automated Vehicle Locator (AVL) systems at multiple places. It seems clear that graphically aided dispatch will afford differential benefits, depending on the nature of the call.

NIJ believes that these systems can be evaluated with reasonable rigor. Because of the "all or nothing" nature of implementing these systems, true experimental designs seem impractical. Pre-post designs seem feasible to implement despite possible differences in pre-post recordkeeping because dispatch systems have historically retained information about time and place of calls. Comparisons of response times and call outcomes to nearby police departments (without AVL dispatching) seem feasible as well.

NIJ sees the potential for secondary benefits for GPS/AVL dispatch and encourages applicants to consider how these benefits could be assessed during an evaluation.

- GPS/AVL systems produce positional logs that place patrol car locations for every minute of the month. Coupled with dispatch information about the time and place of calls, do these systems create new opportunities to optimize beat designs and patrol shifts?

- Training for dispatchers requires approximately18 months. Given that dispatchers now see every car and its status graphically, can training requirements be reduced significantly?

Technology Summary: New computer-aided dispatch (CAD) systems are equipped with GPS capabilities that, when paired with GPS locators in patrol cars can provide a map of the current locations of all active patrol cars. This can, in turn, be used to create automated dispatch recommendations based upon proximity.

Scope of Evaluation: An evaluation including statistical analysis of dispatches recorded in the CAD system over three time periods (before implementation, 6 months at the beginning of implementation, and thereafter) could show changes in response times to 911 calls. Prior to analysis, decisions should be made regarding the types of calls where response time is important. Furthermore, demonstrating a reduction in arrival times should be coupled with a demonstration of how a reduction in arrival times affects the quality of police responses such as by increasing arrests or reducing injury.

Summary of Evaluability Assessment Activity: Available documentation on GPS-based dispatching was reviewed and lengthy interviews were conducted with both NIJ experts and local law enforcement in Dallas, Texas; Newton, Massachusetts; Sacramento, California; and Seattle, Washington. In addition, a site visit was conducted at the Sacramento 911 Center to observe GPS dispatching in use.

Finding: *An evaluation of dispatching response times during periods with the GPS/AVL and prior to implementation may show decreased response time to emergency calls and improved outcomes to high-priority calls for police assistance.*

1. Initial Screening

Background

Within the past several years, automated vehicle locators that use global positioning satellite technology have become increasingly popular among law enforcement agencies across the country. Basic AVL technology is not new—navigational equipment that uses GPS systems are now commonplace in both commercial and military vehicles. Such systems, however, can provide special advantages to law enforcement officers, who can use them to improve the efficiency of their dispatching procedures (Teledesign Systems, Inc., 2002), to improve police tactical activities, and to increase officer safety.

Automated vehicle locators use a complex system of modems, GPS receivers, and digital communications systems that connect a vehicle to a particular base station (which, in the case of law enforcement agencies, may be a police precinct) to determine the precise location of that vehicle within a designated region (Teledesign Systems, Inc., 2002). Vehicle locator systems do this with the help of a GPS receiver that uses satellite technology to pinpoint the latitude, longitude, and speed of a vehicle. GPS receivers, which are installed in vehicles, operate by receiving information from at least thre3 (but ideally 4) of 24 satellites that orbit the earth continuously. The satellites submit information about the target vehicle to a base station, which can then estimate the location of the vehicle with an accuracy of approximately 10 meters (Taylor, 2003). By tracking the vehicle and its speed, AVL systems can track a driver's adherence to a route, estimate a particular time of arrival between a vehicle's current location and its desired endpoint, and reroute a vehicle to accommodate any roadblocks or other contingencies that may emerge (Geagan, Raad, and Lim, 2004).

What is the background/history of this technology?

Maturity (i.e., Lab prototype? Field rollout? Multiple generations/manufacturers?)
There exist multiple generations and manufacturers.

Time in the field?
Although GPS-based AVL technology has been widely available for several years, only recently have police precincts used it to provide more sophisticated proximity dispatching. The first uses of AVL technology by law enforcement agencies is not known, although several precincts have published accounts of their AVL systems online—including Sacramento, California; Collier County, Florida; and Newton, Massachusetts—each of which had AVL technology in use by 2004 (Sacramento Police Department, 2004; Baldus and Kim, date unknown; Geagan, Raad, and Lim, 2004).

Prevalence in the field? (Is site a first/early adopter?)
GPS/AVL has been built into almost all of the newly available computer-aided dispatching systems. These systems are being adopted when departments have the resources to upgrade their older systems, thus the adoption of GPS/AVL is currently scattered across the U.S. and is not yet widely prevalent.

What do we already know about technologies like these?
GPS-based AVL technology allows police precincts to refine their practice of proximity dispatching, in which particular patrol officers respond to a variety of police calls based on how near the officers are to the events in question. In the past, proximity dispatching relied on more rudimentary methods of tracking the location of patrol cars—either officers on patrol would call their dispatchers at particular time intervals to alert dispatchers of their locations, or dispatchers would send out a call through a wireless system that all officers could hear, allowing those in the nearby area to respond to the call (Baldus and Kim, date unknown). These older systems were more complicated and time consuming for both patrol officers and their dispatchers. GPS-based systems, however, allow call centers to dispatch the closest patrol units to a particular crime or incident, which can shave minutes off the amount of time it takes to arrive at a particular location (Geagan, Raad, and Lim, 2004).

GPS-based AVL technology provides precincts with numerous capabilities, including the ability of dispatchers to continuously track the precise location of patrol cars (instead of relying on the older system of periodic "checkins"); the access of patrol officers to continuously updated navigational data that includes information on destination locations, remaining distance to destinations, and direction to destinations; and the ability to provide database administrators with periodic or real-time data updates and error capturing (Baldus and Kim, date unknown). Error capturing is especially important in its ability to allow dispatchers to enter data that records changes in traffic patterns and road blocks, each of which can stymie the ability of law enforcement to respond quickly to incidents. AVL systems also can provide officers in vehicles with aerial imagery maps that help them quickly locate addresses, intersections, or other coordinates on municipal maps (Baldus and Kim, date unknown; Teledesign Systems, Inc., 2002). AVL systems allow supervisors to improve tactics by coordinating responses that require the joint operation of multiple units. Finally, an AVL system can increase officer safety by providing the precise location of units when officers are incapacitated.

Which audience(s) would benefit from this evaluation?

- Police precincts
- Geographic information systems (GIS)-based AVL developers and manufacturers

Although NIJ is not in the business of aiding developers and manufacturers, the Sacramento police explained that technology development is slowed by developers and manufacturers misunderstanding police needs. Police and vendors would jointly benefit from an evaluation of GIS-based AVL in operation. Additionally, the Sacramento police explained that vendors, who are often excellent at engineering but not always so good at understanding police operations, largely conduct technology transfer from one agency to another. An evaluation focused on application would benefit technology transfer.

At what stage of adoption/implementation is the technology?

To observe the GPS-dispatching capabilities, we traveled to the new police dispatch center in Sacramento, California. The dispatch center has been using a system with GPS capabilities since January 2006. However, protocols on using the system have been developed more recently. All patrol car mobile data terminals have GPS capabilities. Dispatchers and supervisors (but not patrol officers) can see the positions of all cars. Automated dispatch recommendations are available to the dispatchers and must be used for all priority 1, 2, and 3 calls, such as violent crimes in progress and officers in pursuit. Officially there is full implementation, but staff is still warming to the technology and usage of automated dispatch recommendations varies.

What efficiencies or primary/secondary outcomes are expected?

Sketch the logic by which technology use should affect goals.
Is the technology well suited and appropriately specified given these goals?

The mapping technology alone will enable better localization of vehicles in the fleet. When this information is provided to dispatchers and supervisors it will lead to the following outcomes:

- *Heightened ability to coordinate activities such as perimeter closure and approach of a crime scene with multiple vehicles.*
- *Location of known endangered officers or officers who do not respond.*

In addition to these benefits from the mapping of patrol vehicle locations, automated dispatch recommendations have the potential to reduce response time to 911 calls. An outcome evaluation would presumably focus on reducing response time. A process evaluation would focus on tactical uses. An increase in officer safety is an important aspect of AVL systems, but this use is probably not amenable to either an outcome evaluation or a process evaluation because this use would be so rare.

Are there operational alternatives that could be used for comparison?

Current operational alternatives are not available for comparison within the Sacramento system. As a policy, all dispatchers are required to use the automated dispatch system for the highest priority calls. However, data from the time of adoption of the technology can be compared against data from previous years when appropriate confounding variables are taken into account.

Is the site interested in being evaluated?

The site is interested in being evaluated because little information is currently available on the benefits of this system, especially automated dispatch. If the technology is shown to be valuable, it may result in more willing acceptance among both patrol officers and dispatchers.

Is the site planning an evaluation?

The site is not currently planning an evaluation.

Data Sources

What data systems exist that would facilitate evaluation?

The computer-aided dispatch system keeps detailed records from the time of origin of a call until the a police officer arrives on the scene or the call is otherwise terminated. In a given year, the dispatch center receives between 25,000 and 40,000 high priority calls. Records from these calls can be made available to an evaluator. Further information on the outcome of these calls may be obtainable from the records management system (RMS). The RMS records case information after the dispatch center is no longer involved. Additional effort would be required to cross-match records from the RMS system to the CAD system.

What key data elements are contained in these systems?

The CAD system records several key data elements:

- *Priority of call: only high-priority calls use automated dispatch.*

- *Automated dispatch: a separate item states whether automated dispatch was used;*
- *Location of call and ID of responding vehicle.*
- *Response time from origin of call to arrival of police.*

After the CAD system closes a file it passes information to the RMS. The RMS will have additional information depending on the outcome such as whether an arrest was made.

Are there data to estimate unit costs of labor and capital?

GPS dispatching is a feature that is bundled with new CAD systems. It may be difficult to identify the cost of GPS dispatching alone; however, information on GPS units in cars and the CAD system can be identified.

Are there data for possible comparison technologies or other solutions?

Similar data is available for the CAD system before and after implementation of GPS-based automated dispatching. However, priority categories and in some cases beat numbers have changed during the same period so care must be taken to match data accordingly.

In general, how useful are the data systems to an impact evaluation?

The CAD database is an extremely rich source of data. The volume of calls the dispatch center intercepts allows for better statistical power even after breaking calls into smaller priority-based categories. Response time is recorded and is an important measure of the effectiveness of the system, even if the final outcome of the calls is not extracted from the RMS.

2. Site Visit Screening

The Intervention

Has the organization implemented a policy or training for the technology's use?

Dispatchers receive extensive training in use of the CAD system including the GPS automated dispatch. To enforce the use of the automated dispatch system, the police department set a policy requiring dispatchers to use automated dispatch recommendations for all high priority calls. The CAD system records when the automated dispatch recommendation is made.

Who are the users?

There are two sets of users for the GPS dispatching system. The first is the dispatching center, which will have access to a map of the GPS location of all *vehicles and the automated dispatching system. The second group is active patrol officers, particularly supervisors. Supervisors can access a real-time GPS map of vehicles on their mobile data terminal to better coordinate police action.*

Who/what are the targets?

The targets, i.e., items that are tracked, consist of all marked and unmarked police vehicles. Vehicles are color coded on the map according to status (active on-call vs. available). GPS capability could be used in police radios to further localize officers who are not in their patrol cars, but Sacramento does not have this capability at this time.

Who/what gets excluded as a user or target?

Police vehicles that are not being used or that have been taken out of the fleet for maintenance do not broadcast their position. The GPS antennae may be sabotaged or obscured easily, but thus far Sacramento has had no problems with this behavior. The antennae may be obscured from satellite in other ways if, for example, the police car is on the lower levels of a parking garage. However, using latest location information, the vehicle could still be located quickly.

Have the characteristics of the user or target population changed over time?

The system has been in place for less than a year, so the user and target populations have not changed significantly. Currently only supervisors and dispatchers have access to the map of all vehicles, but regular patrol officers may have access to this map in the future.

What values/outcomes do users see/envision in the technology?

There are two components of the technology that each have their own benefits: the GPS map of all active vehicles and the automated dispatch system. The prime value that users see in the technology is in the GPS map of vehicle locations. With this map, supervisors are better able to coordinate group movements of vehicles such as perimeter closures. In addition to this there is a benefit to officer safety. Before the GPS system was in place, dispatchers and supervisors knew the location of a vehicle only when an officer radioed in to respond to a call. With the GPS map, an officer's car can be located at any time, whether or not the officer is able to respond over the radio. While the officer may be outside of the car, the location of the car gives a much more localized point to begin a search.

The automated dispatch system has the potential to reduce response times to emergency calls. This potential has not been tested in Sacramento, although other systems (notably Ottawa, Canada) have done response time testing. It is unclear to what extent the automated system will provide useful information to a well-trained dispatcher. It should be noted that the automated dispatch system is *only useful if it provides better assignments than a well-trained dispatcher. That is, if the automated dispatch produces the same recommendation as a trained dispatcher, there will be no observed benefit in response times.*

What are the limitations/obstacles to using the technology?

Patrol cars must have GPS capability installed and the software to visualize the vehicle map in order to provide a benefit. All Sacramento police cars have GPS antennae and only supervisors' cars have access to the map. Beyond the hardware requirements, the GPS system adds no burden to the patrolmen.

The automated dispatch system is still being refined and has some limitations. Currently the automated dispatch system provides recommendations based solely upon the latitude and longitude of vehicles. It is not linked into a driving time estimator, so it will always recommend the geographically *closest vehicle, which is not necessarily the same as the vehicle which can arrive fastest. The system also does not take into account finer distinctions in patrol car status. Any patrol car may be reassigned if it is not already involved in a high priority call. Thus, an officer involved in a low priority call may be assigned over another available officer one minute further away.*

What outcomes could be assessed? Using what measures?

The benefit of the GPS vehicle map to supervisors would be difficult to assess. Quantitative measures such as number of successful arrests are affected by many variables outside of successful perimeter containment. The benefit to supervisors in planning and the benefit in terms of officer safety may only have anecdotal evidence.

It seems reasonable to hypothesize that incorporating GPS/AVL technology into a CAD system would reduce response times for priority calls. An evaluator could test this hypothesis using CAD data. Several observations may be useful:

1. *Response time is relatively unimportant for some calls for service. In fact, Sacramento dispatchers refuse to even dispatch a police car in response to some instances (such as traffic accidents) that do not involve injury. Response time is only relevant for priority calls. An evaluator should be able to determine if response time has decrease for those response calls.*

2. *However, quantifying the* importance *of a decrease in response times should be part of an evaluation. For example, the probability of an injury from a domestic dispute may decrease as response time decreases; the probability of apprehending a suspect in response to a silent alarm may increase with a decrease in response time. In*

contrast, the prospect of clearing a home burglary probably has nothing to do with the speed at which police arrive at the scene.

3. *The benefit of a GPS automated dispatching system probably varies with environmental conditions. During times of the day when calls for assistance are relatively light, computer aided dispatching may contribute little to improving performance. The improvement may be limited to periods when use of police responses is most intense. Such periods may be situational and unanticipated; for example, there may be spells during which police resources are under unusually high demand, and at such times, the GPS automated dispatching may be most important.*

4. *As noted, supervisors have access to the AVL location system in Sacramento, and reportedly they use the system for tactical operations such as closing perimeters. We are unsure of the extent of such use. An evaluation could benefit from a qualitative description of the types of problems that cause police supervisors to coordinate the actions of multiple police cars. The description would include an account for the objectives of that coordination, such as capturing a suspect known to be on foot in the area. Once such events are described, they become in theory countable. If countable, then an evaluator could judge whether or not the use of GPS had improved responses.*

The point of these comments is that the benefit of an AVL system may be situational. Quite possibly, a CAD system that is not equipped with GPS-based AVL technology could perform as well as a CAD system that is equipped with GPS-based AVL technology for routine operations. The value of the GPS-based AVL technology may be limited to certain situations. An evaluation based on mean response time may mislead regarding the value of a GPS-based AVL when extreme demands are placed on the system.

Designing a Study

Are there other operational environments for which the technology is well-suited?

What are the constraints in such environments?
A call for assistance might result in several possible public responses: dispatch of police, dispatch of ambulances, and the dispatch of fire-fighting equipment. There are important differences across these three responses. Police cars are typically on patrol, and a vehicle location system provides the advantage of tracking all on-duty cars so that the car that is the nearest to a requested call for service can be dispatched. Emergency equipment is more likely to be stationary, but nevertheless, ambulances and firefighting equipment are often dispatched

while returning from a service call, and additionally equipment could be diverted from lower-priority calls to higher-priority calls.

In Sacramento, the call center receptionists receive calls for service and make separate referrals to police and fire/emergency equipment dispatchers. We did not investigate the latter use of GPS-based AVL technology.

Do the technology "events" permit randomly generated applications of the technology?

If not, can comparison samples be formed? With what difficulties?
An evaluation based on random assignment seems unlikely. An evaluator would have to find a way to randomize both calls for service and police responses, using GPS-assisted dispatching for the "treatment" group and routine dispatching (CAD without GPS-assisted dispatching) for the "control group." This seems impractical. A quasi-experiment is practical. There are three principal research questions.

One question is whether or not the outcome of a call for assistance is affected by police response time. The answer is not straightforward, however. There are several considerations:

1. *The importance of quick response time probably varies with the nature of the call for service. For some service calls, the outcome might be judged by the arrest of a suspect, the identification of witnesses, or by securing evidence that would otherwise disappear. For other service calls, the outcome might be judged by preventing injury to a victim. For still other types of calls, the outcome might be judged by preventing injury to the police officers. An evaluator will need to develop a typology and specify a meaningful outcome for every element of that typology.*

2. *The relationship between response time and outcome is likely to be nonlinear. For domestic abuse calls, reducing the response time from 6 minutes to 4 minutes may make a material difference with respect to injury. In contrast, for domestic abuse, reducing the response time from 20 minutes to 18 minutes may be immaterial.*

The above points are speculative, because with few exceptions, these topics have not been researched. Nevertheless, without understanding how response time is related to outcomes, there is little or no basis for placing a value on the ability of GPS-assisted dispatching to response time.

Having established that response time is worth reducing, the second question is whether GPS-assisted dispatching can reduce response time. This question *might be answered with a simple design that divides time into three periods: pre-GPS-assisted dispatching, the first 6 months of using GPS-assisted dispatching, and thereafter. The 6-month time span is an estimate of time required to achieve a steady-state use of the new system. The true time might be longer or shorter. Again, there are some complications:*

1. *This is a pre-post design, and for it to be valid, we must assume that nothing besides the implementation of the GPS-assisted CAD system has changed. That assumption may be difficult to justify. At the least, we would require that the county had the same number of patrol cars and police and calls for service. If that assumption is unsustainable, then the evaluator must be prepared to introduce statistical controls.*

2. *Identifying statistical controls is not trivial. At the least, an evaluator would have to create a control variable that represents the availability of police responses when a call for service was received. When assessing the improvement of response times under the pre-GPS-assisted and post-GPS-assisted dispatching, the evaluator must be careful to compare response times under similar conditions.*[1] *Additionally, it would be prudent to control for day of the week, time of the day, and weather—variables that affect travel times—when conducting the comparison.*

3. *Dispatchers in Sacramento felt that they made decisions that were as good (and often better) than rule-driven dispatching based on event/car location. If that is true, the utility of a GPS-assisted dispatching may be greatest when inordinate demands are placed on the police, because in such cases dispatchers may benefit most from automating their responses and the utility of a rule-driven system may be most apparent when police responses are most strained. This suggests that the control variable (discussed above) might be used as an explanatory variable—that is, the "treatment effect" would depend on the relative demands placed on the system.*

4. *Reducing the* average *response time is probably not a useful metric. A GPS-assisted CAD system is likely to change the entire distribution of response times. If the relationship between response time and*

[1] Developing a suitable measure would be a major activity. At the time that the nth call for service is received by a dispatcher, the evaluator would have to determine how many vehicles could be dispatched to field the call, and how many vehicles could be diverted from lower-priority business. A more sophisticated variable could be constructed given deeper understanding of dispatch procedures and data availability (especially the likely proximity of the vehicle to the call for service).

outcomes is nonlinear, then detecting a change in this distribution is especially important.

Sacramento police tell us that supervisors use the GPS system to inform tactical decisions. An evaluator might attempt to identify events where the GPS system has been used for this purpose. Qualitative analysis might be most informative. If events can be classified, then one might compare the outcomes pre-GPS and post-GPS availability.

How many times would the technology be applied in one year?

The use of the GPS-assisted CAD would be continuous. The use of the GPS system for tactical decisionmaking would be episodic, but we are uncertain of the frequency of use.

Will modest but statistically significant effect sizes be detectable given sample sizes?

Because we lack knowledge about the relationship between response time and outcomes, we have no good way of defining "modest." Given the volume of calls for service (25,000–40,000 high priority calls), it is highly likely that an evaluator could detect a reduction in mean response time of less than one minute and even a reduction of a fraction of a minute. (We do not know the average responsive time in Sacramento, but we presume that it is roughly 6–8 minutes, so a 1-minute reduction would be large.) This is misleading, however. A reduction in response time is only important if it has a material effect on the outcome of the call for assistance, and this may be true for only a small proportion of calls for assistance.

How many units—if any—would have to be procured for an evaluation?

None if the evaluation were done in Sacramento.

What does a control/comparison group receive?

Routine patrol.

What kinds of data elements are available from existing data sources?

The dispatch data provide response times starting from when a call for assistance was received by the CAD center and ending with the arrival of a police car at the scene. (The arrival time may be more complicated than this if the first car is not sufficient to deal with the problem.) The problem is that an evaluator would need to know more about the call. What was it about? What was its resolution? An evaluator would need to extract this information from the RMS.

What specific input, process, and outcome measures would they support?

The principal measure is the time from receipt of a call for assistance until a police unit arrives on the scene. Other measures would be important:

1. *For domestic violence cases, and for other assaults, an outcome measure might be the use of an ambulance to transport a victim to the hospital. Obviously this measure misses many elements of harm that might be suffered by a victim, but we would expect that an effective police response would tend to reduce serious episodes of personal harm.*

2. *For some property crimes, a measure of success would be the arrest of a suspect; an additional measure would be the identification of a witness.*

How complete are data records? (Attain samples if possible)

Although we did not acquire records from the dispatch center, we understand that they are complete. We did not independently review the RMS data.

Can user and/or target populations be followed over time?

This question is irrelevant for this evaluation.

Can the dosage of technology used be identified?

This question is irrelevant for this evaluation.

Can data systems help diagnose implementation problems?

This seems unlikely. In theory, it might be possible to observe the frequency with which dispatchers follow the rules for assigning police cars, because the dispatchers indicate whether or not they have followed the rules. We were told that the dispatchers have developed work-around procedures.

What threats to a sound evaluation are most likely to occur?

The evaluation rests importantly on an assumed equivalency between the pre- and post- periods. "Equivalency" means that police responses during the two periods would provide the same response times in the absence of GPS-assisted CAD. There is not apparent way to test this assumption.

What changes is the site director willing to make to support the evaluation?

No changes appear to be necessary.

3. Overall

Would you recommend that the technology be evaluated? Why or why not?

An evaluator should be able to answer these questions: Does response time matter? For what types of calls for assistance? What is the relationship between response time and outcomes? The answers to these questions are at the heart of a cost-benefit analysis of any innovation that purports to reduce response time.

An evaluator will have more difficulty answering these questions: Does a GPS-assisted CAD system reduce response times? Under what conditions? The problem is a familiar one in quasi-experiments: How can we know the counterfactual when the GPS-assisted CAD system is universally employed? An evaluator would have to assume the equivalency of the pre and post periods, and this may be difficult if policing in Sacramento has changed in other ways.

Nevertheless, this is not an outrageous assumption because the pre and post periods are not distant in time. Furthermore, an evaluator could check for trends, and the evaluator could introduce controls for other police resources.

It is more difficult to see how an evaluator could perform an outcome evaluation of the tactical use of GPS. Furthermore, given that GPS is new in Sacramento, we would expect the tactical use to evolve with police experience. A descriptive study of tactical use would be useful.

What type of evaluation design would you recommend?

This is necessarily a pre-post test design. Statistical modeling seems necessary for answering the questions about the relationship between response time and outcomes and about the relationship between GPS-assisted CAD and response time.

References

Baldus, G., and J. Kim. (Date unknown). *Integrating GIS and Real-Time Vehicle Tracking for County Emergency Response*, Naples, FL: Collier County Government.

Geagan, T., M. Raad, and E. Lim. (2004). "Case Study—Use of GIS for Automated Vehicle Location (AVL)—Newton Police Department." San Diego, CA: ESRI User Conference, 2004.

Sacramento Police Department. (2004). *Annual Report 2004.*

Taylor, S. (2003). "Automatic Vehicle Locator," *Pocket GPS World*. Retrieved September 18, 2006, from http://pocketgpsworld.com/avl.php.

Teledesign Systems, Inc. (2002). “Automatic Vehicle Location.” Retrieved September 18, 2006, from http://www.teledesignsystems.com/automatic_vehicle_locator.html.

GPS Offender Monitoring and Tracking
Pinellas County, Florida/Marion County, Illinois

Staff Contact:	Scott McCallum	Brian Barton
	Systems Analyst	Director
	Pinellas County, FL	Marion County Community Corrections
	727–464–6415	708–341–9361

NIJ Guidance

The National Institute of Justice (NIJ) does not recommend an evaluation of global positioning system (GPS) offender monitoring and tracking in the sites assessed below. We remain interested, however, in evaluating the impact of this technology in other sites.

Applicants who propose to evaluate this technology are encouraged to consider the outcome variables (including supervision compliance, reoffending while on supervision, and postsupervision recidivism, as well as potential cost savings from reduced incarceration) and obstacles (including incomplete data, unavailable or incomparable control groups, and insufficient experimental group size) identified below.

Applicants may depart from this guidance by providing an appropriate rationale.

Project Summary: GPS offender monitoring and tracking technology is designed for use with both sentenced offender and pretrial populations. Two field implementations were examined for this feasibility assessment. The first was in Pinellas County, Florida. Select nonviolent offenders there are sentenced directly by the court to an alternative sentencing program. Under this program they are tracked electronically to insure adherence to conditions of their sentence and supervised by deputies from the sheriff’s department. This program began in 2003 and currently has 253 clients participating. The second was in Marion County, Illinois. The primary use there is for offenders sentenced to home detention and a group of pretrial defendants. This program began in 1999 primarily for domestic violence cases and currently has 219 clients participating.

Scope of Evaluation: An outcome evaluation of GPS offender monitoring and tracking is not recommended based upon this feasibility assessment.

Summary of Evaluability Assessment Activity: The assessment of the feasibility of evaluating GPS offender tracking technologies began with a review of the literature and a web-based search to identify vendors that perform electronic tracking of offenders under community supervision. Telephone interviews were then attempted with eight known electronic-monitoring vendors, although this met with limited success. Interviews of technology experts at the National Law Enforcement and Corrections Technology Centers (NLECTC) were also conducted, as were conference calls with NIJ Program Managers from the Office of Research and Evaluation and the Office of Science and Technology. A conference call with NIJ and Mitretek, which is undertaking an electronic monitoring implementation study for NIJ, was also conducted.

The literature review, telephone interviews, and conference calls revealed that GPS electronic monitoring of offenders in the community, although quite widespread, is a relatively new application in the corrections arena. Very little empirical evidence exists regarding the effects of GPS technology. The handful of studies conducted to date suggests that users have encountered technical problems with some applications. Still, early detection of risky behaviors has been cited as a positive outcome in several jurisdictions. In addition, reduced technical violations, reoffending, and absconding have been noted as observable outcomes of electronic monitoring applications.

The initial screening by Urban Institute (UI) identified eight mature applications of GPS offender tracking technology. These were found in Marion County, Illinois; U.S. Pre-Trial Services in the Central District of California; New Mexico Department of Corrections; the City and County of Denver (Colorado); Oakland County; Michigan, Community Corrections; Court Supervision and Offender Services Agency, Washington, D.C.; Texas Department of Criminal Justice Services; and Pinellas County, Florida, Alternative Sentencing Program.

On the basis of the screening information compiled, NIJ and UI mutually decided on October 2, 2006, that Pinellas County, Florida, would be the location for a further site visit screening.

In addition, Marion County was separately chosen for an evaluability assessment of the Secure Continuous Remote Alcohol Monitor (SCRAM). Because a site screening was already planned there, this site was also selected for a supplemental screening of its GPS offender-tracking program. Findings from those site visit interviews are therefore also provided in this assessment report.

1. Brief Literature Review

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

According to an April 2006 survey, 22 States are currently using GPS monitoring systems (ICAOS, 2006). GPS is most commonly used to track sex offenders, but some states are using GPS to monitor other high-risk offenders. For example, New Jersey and California are contemplating GPS monitoring in domestic violence cases; Delaware uses GPS to track movements of juveniles under house arrest; and Pasco County, Florida, is using GPS for pretrial inmates to reduce jail overcrowding (Perlman, 2005).

Despite its prevalence in the field, GPS monitoring has not been the subject of much formal evaluation. However, a recent study in Maryland found that staff training in the use of the technology was inadequate, the system often emitted false readings, hardware failed repeatedly (with vendor response times often taking 2–3 days), and batteries routinely died. Nonetheless, researchers determined that GPS aided in the early detection of risky behaviors before offenders committed new crimes (MTOP, 2004). Several other

California assessments reinforce these findings (Perlman, 2005). Another recent study (Padget, Bales, and Blomberg, 2006) found this technology resulted in significantly reduced technical violations, reoffending, and absconding.

What audience would benefit from this evaluation?

The primary beneficiaries would be corrections, probation, and parole policymakers and practitioners, as well as judges and court administrators. An evaluation would also contribute significantly to the field of empirical knowledge about using technology for offender and pretrial monitoring 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

Both Pinellas and Marion Counties voiced a willingness to cooperate in an evaluation.

Is there a local evaluation?

There have been no formal evaluation in either county to date and none is currently planned.

3. Background History

The Pinellas County Sheriff's Department originally implemented GPS technology in 2003. An alternative sentencing program using GPS emerged from an in-house work release program that was in place at that time. Two hundred fifty-three sentenced offenders currently participate. Each participant is affixed with an ankle bracelet and is required to carry a portable GPS transmitter when away from home. When at home, the mobile unit is normally plugged into its charger and passive transmitter. Offenders are typically sentenced to this program by the court for periods ranging from 10 days to 1 year. Violent and sex offenders are excluded from participation in this program. In addition to those sentenced to this program, the sheriff's department recruits participants from its regular jail population based upon written eligibility requirements. Approximately one-third of current participants are recruited from the jail population and a current waiting list of 50-75 exists.

A sheriff's deputy is assigned to each participant; each deputy typically has a caseload of 40–45 offenders and must visit each offender every 7–10 days. Offenders must reimburse the department $6per day for equipment costs. An active monitoring alarm system notifies the supervising deputies if the bracelets are tampered with or the GPS transmitter is more than 30 feet away from the offender. Passive transmission of the daily GPS-tracked whereabouts of each offender is automatically made daily from the GPS unit when docked in the charging unit. The vendor, ProTech, produces a comprehensive report on the movements of each offender each morning. The alternative sentencing team reviews this report daily for violations. In addition, the vendor's software can display maps of offender locations in 15-second intervals and track movements of offenders over the previous 24 hours throughout the county and beyond. These maps are regularly overlaid on maps of crime in the county for the previous day to see if offenders may have been in the area of a crime when it was reported to have occurred.

Marion County began using GPS in 1999, primarily for domestic violence cases and specifically in response to a particularly high-profile domestic violence case that resulted in a homicide. Shortly thereafter, the legislature became interested in GPS after another high-profile case during which a parolee on electronic monitoring managed to sneak next door and kill his neighbor without triggering an alert on his electronic monitoring unit. This prompted the passage of a law requiring all violent offenders on home detention to be tracked by GPS. In addition, all sex offenders are required to be on parole and GPS-monitoring for life. While not all judges follow these laws, most do, which eliminates any opportunity for a meaningful control group among the sentenced population. Among the pre-trial population on GPS, judges do have discretion, but many are unwilling to make a bail decision without the security of GPS to back them up.

4. Program Design

Target Population

The population for Pinellas County is 253 nonviolent offenders sentenced by the court to the alternative sentencing program or recruited from a similar population currently incarcerated in the county jail. In Marion County, 219 clients (133 pretrial and 64 sentenced) were reported as currently on GPS tracking. The majority of offenders were sentenced for either a felony D offense or a class A misdemeanor (violent or sex offenses).

Project Goals and Objectives

The basic outcome logic of this technology is that nonviolent offenders can be supervised in the community through electronic monitoring of their movements using GPS. The primary outcome suggested is a reduction or minimization of jail overcrowding. In addition, supervision costs using this technology are perceived by program staff to be much lower than incarceration. Theoretically, this technology may also reduce technical and criminal offenses during the period of supervision and reduce longer term recidivism. However, in Pinellas, at least, these last outcomes were not emphasized.

The goals of the use of this technology are to provide a safe and secure alternative to incarceration through electronic GPS monitoring. The objectives are to: 1) reduce jail overcrowding; 2) decrease supervision costs; and 3) prevent reoffending while under supervision (through detection of technical and criminal violations).

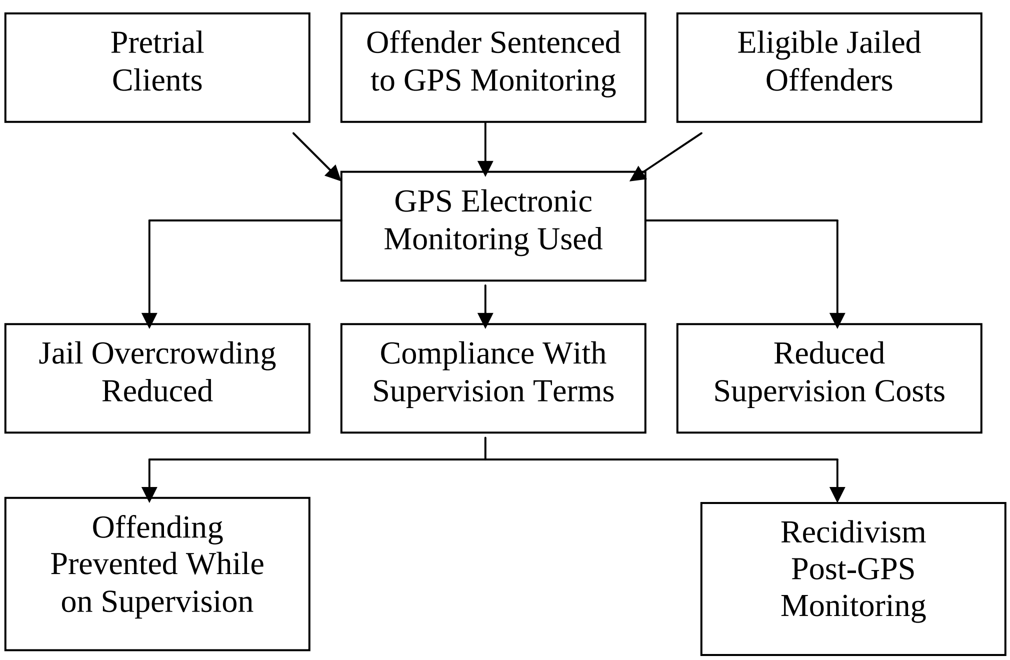
5. Program Logic Model

Describe the logic that connects project activities to project goals.

The basic technology logic model is presented in exhibit 1. As can be seen in this model, GPS monitoring can be implemented with several populations. Pretrial clients released on bail or their own recognizance are one possible application. Sentenced offenders are another population, and those monitored by GPS tracking technology can be directly sentenced as an alternative to incarceration. Alternatively, at least as is the case in Pinellas County, clients can be recruited from low-risk offenders currently incarcerated, much as they might be for work release.

Three primary intermediate outcomes were noted during background research and onsite screenings. First, the use of this technology is an alternative to incarceration and can therefore reduce jail populations and overcrowding. Second, by supervising clients in the community and requiring them to contribute to the costs of this program, supervision costs will be reduced, particularly in comparison to incarceration in the local jail. Third, this technology will increase the likelihood that clients will comply with the conditions of their supervision. For example, they will be deterred from breaking curfews or frequenting locations where they are not supposed to be. As a result, it may be hypothesized that reoffending will be prevented while under electronic monitoring. Theoretically, long-term recidivism might be reduced should this technology have its desired effect as well. It should be noted, however, that officials in Pinellas County emphasized repeatedly that they held no such belief and that neither rehabilitation nor reduced recidivism was considered as likely long-term outcomes.

Exhibit 1. GPS Monitoring Logic Model



Is the logic supportable by empirical evidence?

No empirical evidence was offered in support of the logic model. Anecdotal evidence provided by deputies and program administrators suggests that the vast majority of those participating return to the criminal justice system. One official characterized the system as "a revolving door," although neither jurisdiction reported actually tracking recidivism systematically.

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

A question exists about the logic of this technology reducing overcrowding. At the time of our visit, the jail in Pinellas County housed almost 300 offenders over capacity, which suggests that the extent of overcrowding may be the most appropriate outcome measure. In addition, it may be that there is a "widening of the net" phenomenon at work with this technology. That is, it may be that these clients might have been sentenced to other community programs or released on bail or personal recognizance anyway. This may just be a supervision tool for those on community release or community sentence and not really a technology that could affect some of the longer term outcomes hypothesized.

6. Implementation Issues

Is the project being implemented as planned?

Yes, according to field interviews in Pinellas County. In fact, it is anticipated that the program will expand in the near future. Similarly, in Marion County GPS use appears to be running very smoothly. There were some vendor problems there initially, but the latest GPS technology employed in Marion County sends alerts to a central monitoring unit, which is more efficient than the officer pager system previously employed. Moreover, Marion County GPS clients are now equipped with small personal digital assistant (PDA)-size units rather than the larger more cumbersome units that were originally used.

Describe staffing.

In Pinellas County 6 deputy sheriffs are responsible for caseloads of approximately 45–50 clients each. They are supported by five alternative sentencing support staffers (not including information technology support) that screen and report cases for the deputies and department administrators. Marion County currently has 24 officers responsible for the county's entire home detention caseload of 1,700 clients assigned to one or a combination of the following technologies: electronic monitoring through radio frequency, GPS, and SCRAM.

Describe the stability of the project over time

The implementation of this technology is mature in both jurisdictions. Initial implementation problems arose with technical issues surrounding use and alerts, but these have been largely overcome.

What aspects of the project could be evaluated for outcome?

An impact evaluation would need to explore whether GPS changes client behaviors, making them less likely to commit new crimes while on supervision or awaiting trial and more likely to follow all their terms of supervision. The outcome measures, therefore, would be changes in the rate at which GPS clients commit new crimes and the rate at which they are returned to jail for failing to comply with other terms of supervision or pretrial release.

The most rigorous impact evaluation design would be a randomized controlled trial (RCT), whereby candidates for GPS are randomly assigned either to be on GPS or to receive an alternative sanction (most likely electronic monitoring). It is highly unlikely, however, that judges would agree to random assignment, as that would require them to relinquish judicial discretion.

An interrupted time series design is another method often employed in impact evaluations. However, the way in which GPS was implemented does not suit itself to this evaluation method. For example, in Marion County, GPS was implemented with a very small number of clients in 1999, and those numbers grew slowly until recent years. It is unlikely that the impact of GPS would be so great that one would observe an aggregate effect over time, even if one existed.

The third approach would be to select a matched comparison group among those clients who were not assigned GPS. This would require the identification of characteristics of GPS clients and selecting non-GPS clients who have those same traits (age, race, criminal history, current offense, etc.). Both groups would be tracked over time to compare outcomes and determine if statistically significant differences exist between treatment and comparison groups. However, there are other design issues that would need to be overcome. For example, since neither county maintains recidivism data, it is difficult to determine the appropriate sample sizes that would yield enough statistical power to identify a treatment effect if one exists. Researchers would have to pull historical data by hand to identify the recidivism rate, which could be very time consuming.

What would the outcome measures be?

Outcome measures would be jail population trends, compliance with supervision terms, costs, reoffending on supervision, and postmonitoring recidivism.

How could an appropriate comparison group be created?

Creation of appropriate comparison groups was noted as a serious issue for evaluation in both jurisdictions. Random assignment of sentences by judges was viewed as not feasible and naturally occurring samples of similar comparison groups do not exist.

Are the sample sizes statistically significant?

The population of clients is quite small in both jurisdictions and sampling would not be required.

Is random assignment possible?

This did not appear to be feasible in either jurisdiction.

Recommended Approach

An implementation process evaluation may be warranted. There appear to be a number of important implementation lessons learned that would benefit other policymakers and practitioners considering the use of this technology in the future. However, current evaluation design options are not rigorous enough to produce sound outcome findings. Furthermore, the populations are quite small, making detection of effect sizes difficult, even if comparison groups could be identified. Therefore an outcome evaluation of the application of this technology is not presently recommended.

Alternative Approach

N/A

What strengths and weaknesses do the designs have?

N/A

How long in duration would the evaluation be?

N/A

What would be the estimated cost?

N/A

What aspects of the project make an evaluation more difficult?

Obstacles are described above. The inability to create a similar comparison group and the limited number of participants are the most serious constraints for a successful implementation of a rigorous outcome evaluation of GPS technology at present.

7. Measurement Model

If design challenges could be overcome, the measurement model would correspond to those outcomes described in the logic model and outcome measures section above.

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 arrest and booking data (including digital mug shots) back to 1994, calls for service, incident reports (records management system), and computer-aided dispatch records. Also available are case-specific violation data and location data for each participant over their entire period of participation in the program. These are all maintained by the department itself and access was reportedly not controlled by any of the systems' vendors, as is sometimes the case in other law enforcement agencies. Furthermore, the department voiced support for participating in an evaluation and a willingness to share its internal data with researchers.

Marion County maintains extensive electronic data on clients on GPS, including demographic information, current offense, criminal history, risk level, drug testing dates and results (if applicable), and violations of terms of supervision. This database, however, is case based and does not allow for the creation of reports that aggregate data across the entire client base. Nonetheless, the data exist and could be extracted manually in order to track outcomes and identify characteristics of those on GPS in order to create propensity scores for identifying a comparison group.

Can services delivered be identified?

Delivery of services is not an element of this technology application per se. However, supervision and service delivery are also provided to those participating in these monitoring programs. Whether these services can be systematically identified and tracked was not explored during the site screening since the emphasis was on evaluation of the application of the technology itself.

Can target population be tracked over time?

The current population can be tracked over time. However, there appears to be significant variation in the amount of time any individual client might participate in the program (i.e., dosage). This can range from 10 days to a year for misdemeanants or much longer for felons.

Would an evaluation have to generate new or additional data?

Although extensive case-based databases exist, a substantial amount of extraction would be required for evaluation data analysis purposes. In addition, recidivism and service provision data would need to be generated and collected for research purposes.

9. Summary Remarks

Recommendations for Evaluation

Due to the current inability to create comparison groups, a relatively small number of participants, and current data limitations, possible outcome evaluation designs would be necessarily quite weak. Therefore, at present the GPS offender tracking technology application, as currently implemented in Pinellas and Marion County, is not recommended for evaluation.

References

Interstate Commission for Adult Offender Supervision (ICAOS). (2006). *GPS Update Survey*.

Maryland Transitional Offender Program (MTOP). (2004). *Comprehensive Approaches to Sex Offender Management,* Final Report, Grant # 2001–WP–BX–0022.

McKay, Jim. (2006.) "Electronic Tether," *Government Technology*. Retrieved August 17, 2006, from http://www.govtech.net/magazine/story.php?id=98310&issue=2:2006.

Padget, Kathy G., William D. Bales, and Thomas G. Blomberg. (2006). "Under Surveillance: An Empirical Test of the Effectiveness of and Consequences of Electronic Monitoring," *Criminology and Public Policy* 5(1): 61–91.

Perlman. Ellen. (2005). "Where Are They Now? States and Localities Are Using GPS to Put Moving Targets on the Map," *Governing Magazine* (October) .

Reza, J.D. (2004). "Do You Know Where Your Offenders Are?" *Law Enforcement Technology* 31(6): 118–123.

Tewey, John. F. (2005). "Task Force to Study Criminal Offender Monitoring by Global Positioning Systems: Final Report to the Governor and General Assembly," Annapolis, MD.

Evaluability Assessment of License Plate Reader Technology

Staff Contact: Mark Bateson
Technology Consultant
Sacramento Police Department
(916) 765-3030
mbateson@pd.cityofsacramento.org

NIJ Guidance

The National Institute of Justice (NIJ) has identified some key outcome variables and other parameters of interest for this technology, and has provided some guidance on possible evaluation designs. Applicants may depart from this guidance by providing appropriate rationale.

Technology Summary: LPR (license plate reader or license plate recognition) software is now available that can automate detection of license plates associated with stolen vehicles and other crimes. A mobile camera system mounted on police patrol vehicles recognizes plates in real time, compares them against a database of suspect vehicles, and alerts the officer to any matches.

Scope of Evaluation: The evaluation would entail a randomized assignment of days during which LPR-enhanced and traditional police cars patrol in areas prone to vehicle theft.

Summary of Evaluability Assessment Activity: Documents and evaluations of license plate reader technology were collected from case studies in both Europe and the United States. In addition, experts from NIJ and Appian Technologies were interviewed as well as local law enforcement in the following districts currently using the technology: Pinellas County, Florida; Seattle, Washington; Anne Arundel County, Maryland; Lancaster, Pennsylvania; Mesa, Arizona; and Sacramento, California. A site visit was conducted in Sacramento, California, with members of their vehicle theft unit.

Finding: *License plate reader technology is well suited to a randomized experiment. Benefits in terms of recovery rate, time to recovery, and possibly arrest rate could be measured quantitatively.*

1. Initial Screening

Background

Within the past several years, license plate recognition systems have been piloted extensively throughout Great Britain and, to a lesser extent, in other countries throughout Europe. The promising results of these pilots have led U.S. law enforcement precincts and State highway patrols to consider the possible benefits of LPR technology within the

United States. Using license plate recognition technology for other applications is widespread—in many developed countries, including the U.S., for example, plate readers are frequently used to help monitor electronic toll collection networks. Plate readers have also been used as an intelligence-gathering tool and in surveillance operations (Author unknown, date unknown).[2] The pilot studies in Great Britain, however, opened up the possibility of using plate-reader technology as a law enforcement tool with the potential of proactively addressing criminality—especially vehicular crime, including auto-theft (PA Consulting Group, 2004).

License plate recognition technology that is capable of "reading" plate numbers uses a complicated system of algorithms, cameras, databases, and police intelligence to be successful. Foremost, plate recognition technology requires the use of infrared cameras with optical character recognition software. These cameras can be attached to police cars or other mobile units (along highways or other frequently passed roads); or used as hand-held units that police officers can take to a variety of locations throughout a jurisdiction; or placed in fixed locations (along overpasses, for example) connected to closed-circuit televisions (CCTVs) (Ohio State Highway Patrol, 2005; PA Consulting Group, 2004; Civica Platescan, date unknown). Cameras equipped with plate recognition software are capable of recording the license plate numbers of vehicles driving at high speeds. Reports on the uppermost vehicle speeds possible at which plate scanning software can still record accurate results have varied from 65 mph to as fast as 100 mph (Ohio State Highway Patrol, 2005; PA Consulting Group, 2004). The accuracy of plate readers at various speeds is contingent on a variety of factors, including camera quality, weather conditions, and the existence of common obstructions (like dirt and general plate wear-and-tear) that can obstruct the camera's view.

To read plate numbers accurately, infrared cameras use software with a number of algorithms to identify license plate characters. These algorithms include: Plate Localization (or Image Acquisition), in which a camera identifies a license plate; Plate Extraction and Normalization, in which a camera detects the dimensions of a plate, by compensating for any skewing, adjusting for brightness and contrast, and filtering out any unwanted objects; Character Segmentation, in which a license plate sequence is segmented into individual characters; and Character Recognition, in which the segmented characters are matched to a template of letters and numbers (Kwasnicka and Wawrzyniak, 2002; Parker and Federl, 1996; Valliappan, Sumari, and Kamarulhaili, 2004; Wikipedia, 2006). Typically the software is geared to read plates of a specific State, and even within a State, the software may have limitations when reading atypical plates such as vanity plates. Once a license plate's characters are identified, the information can be sent to relevant databases within a particular jurisdiction to perform background checks on the vehicle. When a vehicle's license plate is flagged, a patrol officer on site can pursue appropriate action, which usually includes stopping the vehicle's driver for questioning, or requesting assistance with the recovery of a stolen but unoccupied vehicle.

---

[2] Closed-circuit television cameras (CCTVs)—a precursor to license plate readers—are common in the U.S. and are used by law enforcement as a surveillance tool that helps provide security within a variety of public venues.

What is the background/history of this technology?

Maturity (i.e., Lab prototype? Field rollout? Multiple generations/manufacturers?)
License plate reader technology has been used in Europe since the 1990s, but its use in the United States is significantly more recent. Several departments told us that they had implemented the technology only within the last year, often with only one vehicle in use. The exceptions are departments such as Sacramento Police Department, which has been testing and using the technology for nearly 3 years, and the Los Angeles Police Department, which has 36 vehicles equipped with license plate reader technology. Because police departments often have unique software systems, significant postproduct development has been necessary to incorporate LPR in an easily usable interface into a typical patrol vehicle.

Furthermore, the *application* of the technology could be improved by better use of extant computerized data. For example, the technology is used to enforce motor vehicle laws, including the use of LPR to identify and boot cars with outstanding warrants. However, that process does not communicate data about stolen cars to police. As another example, the LPR is not yet linked with the State's list of stolen vehicles, so that the list must be downloaded into the system once per day rather than being downloaded as stolen vehicle reports are received. These are not limitations of LPR technology; they are current limitations to information flows in Sacramento. LPR remains an emergent technology for law enforcement purposes.

Time in the field?
Although license plate recognition technology has been available to law enforcement since the 1990s, only within the past 5 years have police agencies begun to use such technology as a tool of crime reduction and prevention. A number of jurisdictions in the United States (including those in Ohio, Florida, Washington, D.C., and California, among others) have implemented pilot tests of their own in the past couple of years.

Prevalence in the field? (Is site a first/early adopter?)
The precise number of municipalities in the United States that use license plate recognition technology as a tool of criminal law enforcement is not known. There is reason to believe, however, that the number is relatively small. LPR systems are expensive and only departments that have high auto crime rates have been receptive to the growing number of LPR vendors in the United States. Much of the literature on license plate recognition technology was published within the past 2 years, and most of it references pilot tests in the United Kingdom that were conducted within the past 5 years.

What do we already know about technologies like these?
The evaluations of license plate readers in the United Kingdom yielded several positive results. During the United Kingdom's 6-month pilot, police officers used plate reader technology to recover £2.75 million of stolen vehicles and goods, and seize more than £100,000 of drugs.[3] In addition to the vehicles recovered during the pilot, eight vehicles

[3] Other recent studies conducted in Great Britain suggest a link between vehicle theft and other serious crime (see Chenery, Henshaw, and Pease, 1999, as cited in PA Consulting Group, 2004).

were recorded as stolen by plate readers, but were not recovered because police officers were interrogating other drivers at the time of detection (PA Consulting Group, 2004). Teams of patrol officers who used plate readers also achieved an arrest rate that was 10 times the national average, although according to the pilot's evaluators, "it will be essential to know the outcome of arrests made by intercept teams relative to conventional policing"—for example, the number of arrests made during the pilot that go to court and the number of defendants who are convicted compared to the national average (PA Consulting Group, 2004).

Within the United States, the success of license plate recognition technology has been, at times, less apparent. While preliminary pilots in the U.S. have resulted in some positive outcomes, these pilots have also revealed some of the contingencies that software developers must contend with in the U.S.[4] Because the size and shape of license plate characters can vary from State to State, plate reader technology can frequently misread license plate numbers. In addition, plate readers frequently register false alarms by matching a license plate number to the plate of another State. According to the Office of Law Enforcement Technology Commercialization (2004), "currently available [plate recognition] systems do not distinguish between States. Therefore, if a system encounters a string of letters and numbers that are wanted in one State on another State's plate, the system will alert. Common vanity plates such as "HELLO" or "GOODBYE" are especially susceptible to this problem.[5] Other inaccuracies may arise due to common obstructions to license plates, like trailer hitches, ice and snow, and vanity plate covers, which are still legal in a number of jurisdictions (McFadden, 2004).

What could an evaluation of this technology add to current knowledge?
An evaluation of the benefits of this technology could potentially show recovered property equal or greater in value to the cost of the system both in terms of materials and training. If benefits such as these can be proven, then precincts that show high auto theft rates would be more likely to invest in this technology to assist recovery and ongoing investigations.

Moreover, both the technology and law enforcement applications of the technology are emergent. The uses of LPR technology to identify stolen cars and drivers with outstanding warrants are apparent; so, too, is the use of LPR for enforcing motor vehicle violations. During our site visit to Sacramento, officials discussed the use of LPR technology to monitor the flow of cars at crimes scenes, and they noted the potential for using LPR technology to passively monitor the movement of vehicles associated with suspicious drivers including terrorist suspects.

---

[4] A preliminary field test of plate recognition software conducted by the Washington Area Vehicle Enforcement Unit recovered 8 cars, found 12 stolen plates, and made 3 arrests in a single shift (McFadden, 2004).

[5] A study conducted by the Ohio State Highway Patrol (2004) concurred with the findings of the McFadden, noting, "standardizing license plates across states would greatly enhance the performance of automatic plate reader technology." In addition, OSHP mentioned, "easier state recognition on the plates would also improve the usefulness of LPR technology…."

Which audience(s) would benefit from this evaluation?

- Police precincts
- Highway authorities

What could they do with the findings?
If the evaluation shows real benefit to the departments and potential increases in revenue from tickets and violations or decreases in court costs, police departments could use this data to gain funding for purchases of LPR systems.

At what stage of adoption/implementation is the technology?

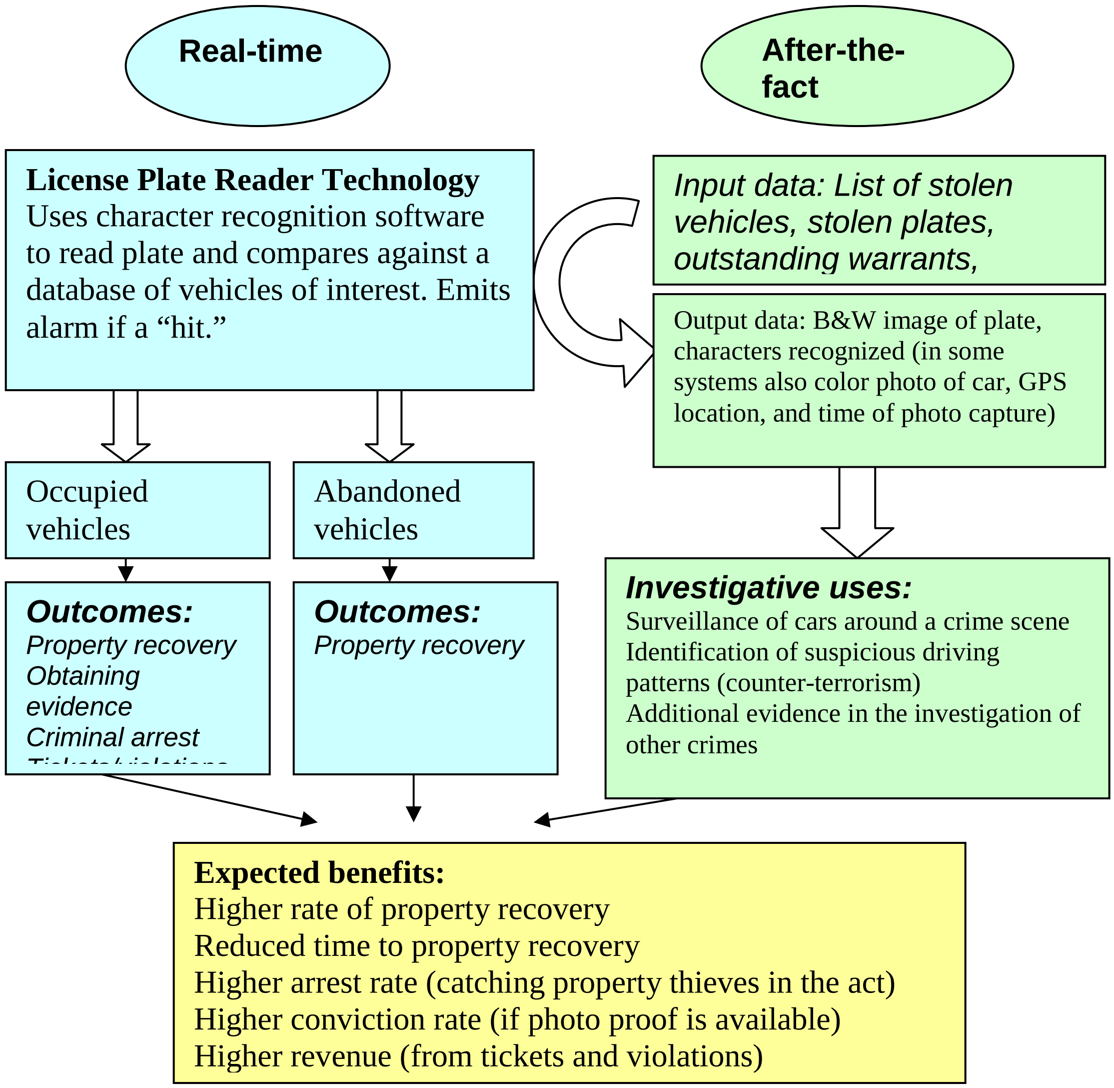
Sacramento Police Department was one of the first field testers of license plate technology and has been testing and using systems for nearly 3 years. However, they continue to work with vendors on improving the user interface and optimizing capabilities. Currently they have only one car and one surveillance van outfitted with license plate reading technology. Few individuals have been trained in its use and the vehicle is taken out approximately once a week. Because the technology is still changing rapidly, documentation has not yet been written for a standard protocol. When fully implemented, the Sacramento Police Department hopes to have several cars (up to six) collecting plate information on both shifts (total of 20 hours), 7 days a week.

What efficiencies or primary/secondary outcomes are expected?

Sketch the logic by which technology use should affect goals (see exhibit 1).

Is the technology well suited and appropriately specified given these goals?

Exhibit 1. LPR Technology Logic Model



Are there operational alternatives that could be used for comparison?

The operational alternatives to LPR are sending officers out to run plate information on a plate-by-plate basis for suspicious vehicles (clearly abandoned vehicles). Generally a police department is only aware of vehicles stolen in its area and may not recognize plates as stolen if the vehicle in question was stolen in another area of California.

Is the site interested in being evaluated?

The Sacramento Police Department would be happy to be part of an evaluation of this technology.

Is the site planning an evaluation?

No evaluation is yet planned.

Data Sources

What data systems exist that would facilitate evaluation?

There are two types of data that would facilitate evaluation: the data generated by the LPR system itself and the CrimeConnect database. (See the discussion below.) Additional information on outcomes can be gleaned from the Sacramento Police records management system (RMS).

What key data elements are contained in these systems?

The Platescan reader (the LPR system used in Sacramento) automatically records the number of plates read in each session and the number of positive hits. It does not distinguish between true and false positive hits, but license plates are routinely double-checked by the officer on duty and false positives will be noted. No outcome information is contained in Platescan, so for additional information on the number of recovered cars, the value of recovered cars, and the number of arrests, one would have to go through the Sacramento RMS system to trace each case.

In addition to the LPR reader, there is the database CrimeConnect containing all images of license plates, cars, and global positioning system (GPS) position and time of photo for each vehicle. (The system shows the reading made by the LPR, a picture of the plate, and a picture of the car.) This database is used for further investigation of cases. Usage of this system could be fairly simply tracked to show how often the vehicle theft group uses the database for investigative purposes. At the time of our visit, usage was limited by staff availability.

Are there data to estimate unit costs of labor and capital?

The specific capital expenditure for each system is between $20,000 and $25,000. Beyond that there are costs for maintenance and support with the vendor as well as for a technology consultant to manage the equipment and address local user concerns. Training costs are unknown as the system is still very much in flux and no final training protocol has been developed. In general, costs for this system have not been separated from other costs to the department.

Are there data for possible comparison technologies or other solutions?

In the absence of an LPR system, stolen cars are likely to be recovered during traffic stops resulting from other enforcement activity and as a consequence of citizens reporting abandoned vehicles. Thus, the "other solution" is traditional enforcement. As we discuss later, the use of this technology to recover stolen cars could be evaluated using a random assignment design. The "treatment group" would be an area patrolled using LPR technology; the "control group" would be an area patrolled by traditional policing.

Going beyond automobile theft recovery to examine the role of LPR as a crime-solving tool, we are less certain about suitable comparison groups. An evaluator would have to investigate how LPR was used in crime investigation settings, the type of data that it generated, and how those data were used. An outcome analysis would probably be premature prior to understanding the process.

In general, how useful are the data systems to an impact evaluation?

The data automatically generated by the Platescan system will be very useful in an impact evaluation, but is not a complete picture. In order to do an impact evaluation, data on additional outcome measures will need to be generated through searches of cases in the RMS system. Furthermore, if the evaluation extended to investigating prosecution and conviction rates, an evaluator would have to match arrests with prosecutions and court data.

2. Site Visit Screening

The Intervention

Has the organization implemented a policy or training for the technology's use?

The Sacramento Police Department has trained a few users within the department in the use of the LPR-enabled vehicle. A protocol for the real-time capture of license plates calls for different responses depending on whether the vehicle is occupied or unoccupied and whether the police officer is available. The protocol follows these guidelines:

- *Alarm sounds*
- *Is officer occupied with another call?*
    - *If yes, he or she radios the information to another officer for confirmation and follow-up*
    - *If no, he or she confirms plate information*
- *Is vehicle occupied?*
    - *If yes, officer attempts to pull over vehicle*
    - *If recently occupied, officer may wait for return of driver*
    - *If abandoned, officer calls tow truck and completes paperwork*

Who are the users?

The vehicle theft unit currently manages the system, though the database can be accessed for investigative purposes in other crimes. Officers in the vehicle theft unit drive the test vehicle and surveillance van.

Who/what are the targets?

The license plate reader system used by Sacramento has been tailored to read standard California auto license plates. Three target databases are currently loaded into the system: stolen vehicles, stolen license plates, and a relatively small group of vehicles associated with other crimes. All of these represent only California vehicles. The system has the capability to use other databases including parking violations and unlicensed drivers, but these have not yet been implemented.

Who/what gets excluded as a user or target?

The technology currently recognizes about one in three or four plates as the unit is being driven past a series of cars. This does not reflect the technological barriers to character recognition on a visible plate, but rather the number of plates that are obscured or otherwise unreadable. It has difficulty reading plates that are:

- *Bent*
- *Dirty or obscured (for instance parked too close to another car for a full plate read)*
- *Covered (illegally) with reflective material*
- *Positioned high (as on an SUV) or at a high angle to the camera*
- *Older plates*
- *On cars traveling at high speeds*

It also has difficulty reading plates that have stacked characters (some State license plate designs, handicapped license plates).

Stolen cars that may not be identified are those without plates, or with borrowed or altered plates. Cars without plates will not appear on the system but are in *obvious violation of the law for other reasons. Borrowed plates may allow a stolen car to pass unnoticed as long as those plates have not been reported stolen. Altered plates may also pose a difficulty, but California plates have raised letters and thus are more difficult to alter.*

In addition to the technological difficulty of reading plates, the system will not recognize stolen autos from other States because only California stolen vehicle and stolen plate information is entered into the system. The system sometimes reads a partial plate, allowing an investigator to do wildcard searches to match those partially observed plates with a list of plates used for crime detention.

Have the characteristics of the user or target population changed over time?

The characteristics of the target population have changed and continue to change over time. The system is designed to read current California plates (within the last 25 years). These will make up a larger proportion of the total California plates observed as time goes on. Sacramento continues to work with state and local authorities to gain access to other databases of both input and output data. Input data include data on outstanding warrants for arrest while output data include the database of license plates gathered by the Sacramento Parking Authority that may prove useful in investigative efforts.

In the future, the users are likely to change as well. The department may enlist the help of retired officers to drive the vehicle, and observe and report any hits to an active officer. Alternately, the officer driving the vehicle may use a community service officer. These approaches free police from completing paperwork on abandoned vehicles and being out-of-service during the time required for towing.

What values/outcomes do users see/envision in the technology?

The current system is used in a patrol car to identify stolen vehicles and in a surveillance van to support long-term investigations. The Sacramento Police Department envisions other uses, such as using the LPR to track vehicle travel patterns across the county. Presumably this would work by identifying certain cars as worthy of being tracked, and then recording the geographic position of "hits" to establish driving patterns of suspicious individuals. Artificial intelligence system might be used to identify patterns worthy of enforcement attention. As they characterized the technology: LPR can be used to collect data; the police then need to develop the means to analyze those data to provide information of use to enforcement.

What are the limitations/obstacles to using the technology?

The benefits from real-time license plate recognition are only as good as the database used. For the police to identify stolen vehicles, the LPR system must have fresh and accurate information. Currently the data regarding stolen cars, plates, and outstanding warrants are downloaded before using the outfitted patrol *car, so that the data become dated as the day progresses. If the system were implemented more widely, multiple downloads in a day would enhance the officer's ability to identify stolen vehicles quickly and ideally while still occupied. (This data update could happen by driving a car through a "hot spot" that would allow an automatic download.) Obtaining access to data requires developing partnerships within the state. Some of these partnerships have already been developed, but access to additional databases will have to be worked out on an individual basis.*

The investigative benefits to the system would be significantly enhanced with more widespread use of the technology. The investigative database currently contains some 50,000 plate images, many of which are duplicates. (The system is intended to include duplicates, as one use of the system is to record multiple occurrences of the same car.) As the system is used more frequently and as other patrol cars are outfitted with the system, the database will be greatly increased. This will make it more likely that a vehicle of interest has been viewed and tagged with a location.

It should be noted that some popular systems capture neither the GPS location nor a color photo of the vehicle. These simple systems are valuable for real-time recognition of stolen vehicles but are much less useful for investigation purposes. The records for these systems are typically not kept and may be purged from the system after 2 weeks or less.

What outcomes could be assessed? Using what measures?

Using LPR to identify and recover stolen cars provides the clearest measure of outcomes. Cars are stolen for two generic purposes. The first motivation is to either sell the car or to strip its parts and market those parts for resale. It seems unlikely that an LPR would be effective at identifying cars stolen for commercial theft purposes, because the thief has an incentive to move the stolen car from the point of the theft to a hidden location before the owner files a stolen car report. Thereafter, there is little reason for the thief to drive the car into an area where it might be detected by the LPR.

The second motivation is for a thief to acquire a car for instrumental purposes. These may range from joyriding by teenagers to using the car as a temporary conveyance to or from a crime. Cars stolen for such purposes are usually recovered, although the car may have been damaged.

It seems unlikely that a LPR would increase the eventual recovery rates, but nevertheless, a LPR could have three distinct advantages:

1. *Because an LPR operates in real-time, the LPR system may increase the probability that the police detect a stolen vehicle that is* occupied*. The occupant may be the thief, or at least, the occupant may be someone linked to the thief. Thus some form of clearance rate (arrest, prosecution,*

conviction) would provide a metric of the outcome from using an LPR system.

2. *The LPR might cause the earlier recovery of stolen cars. Even if all cars stolen for instrumental purposes were eventually recovered, the time from theft to recovery may be material and costly to the car's owner. A simple metric is the length of time from theft to recovery, which might be monetized by using the dollar-cost of the rental of an equivalent car.*

3. *The LPR might reduce the damage to a stolen car. Stolen cars are often vandalized. For example, opportunistic thieves may remove the tires from an abandoned car. Hence, reducing the time until recovery might reduce the damage. We are uncertain that the police would record the damage amount, but owners could estimate those damages if asked. Otherwise, one might rely on a proxy estimate of damage as a function of time until recovery using insurance claims to estimate. This would require the cooperation of insurance companies.*

These outcomes are measurable and, as we discuss below, an evaluator could use a strong research design to evaluate the effectiveness of LPR at reducing the cost of automobile theft and, perhaps, at increasing the capture rate of automobile thieves.

As noted above, the LPR system might be used in the investigation of major crimes. At the least, this use could be converted to a counting exercise—how frequently was the system queried to detect driving patterns useful for an investigation? How often did these queries yield useful information? Since these queries must be done at a central source, they would be countable. Then an evaluator might track how those results were used in an investigation and in support of a prosecution. The Sacramento Police Department also uses an LPR-equipped van to monitor traffic at crime scenes. This latter use of LPR technology is unlikely to lend itself to random assignment experiments or even to quantification. Although the police might be convinced to use random assignment of the van to major crimes, we suspect that they would demand to assign the van to crimes where the van's utility would be the greatest. In either case, we expect that there would be a fairly small number of such events, so that statistical analysis would be precluded. An evaluator might better use a qualitative plan to study the use of a van equipped with LPR technology to augment other investigation resources.

Designing a Study

Are there other operational environments for which the technology is well suited?

What are the constraints in such environments?
The technology is also well-suited for use in fixed locations, as has been shown in European usage of the technology. These sites could be placed on major

traffic arteries to alert the police to stolen vehicles entering or leaving the city. However, these systems are of little use unless paired with law enforcement backup. The great benefit of the mobile systems in place in Sacramento is the ability for the alert to be immediately acted upon.

Do the technology "events" permit randomly generated applications of the technology?

If not, can comparison samples be formed? With what difficulties?
Random design experimental studies seem practical for evaluating the use of LPR to reduce the cost of automobile theft. One approach would be to allow the police to select an area to be patrolled and then to randomly select a day to use LPR technology to patrol that area. The same area would be patrolled routinely (without LPR technology) during the previous 6 days. The one-day patrol would be the experimental period; the 6-day patrol would be the control period.[6] *Outcome measures would be cars recovered during the experimental period and cars recovered during the control period (prorated to a daily recovery basis). The principal metrics are the average length of time required to recover cars during the experimental period compared with the control period and the average damage to the car during the experimental period compared with the control period.*

Two elements of this design require discussion. First, the experimental period must follow the control period rather than the reverse. Reversing this order would bias the treatment effect. Specifically, suppose that the LPR technology does not necessarily increase the recovery rate but that it does reduce the time to recovery and the damage done to recovered cars. Although the eventual recovery rate would be the same whether or not LPR technology were used, the LPR technology will cause more stolen cars to be recovered on a day when LPR is used than on a day when it is not used. Thus, there would be fewer stolen cars on the street in the 6 days that follow the day when LPR is used, so an evaluator would not want to use that subsequent 6-day period as a control period. Second, we assume that the patrol areas should be stratified, and the experiment would pertain to a selected stratum. This is not a necessary step, but it does provide some assurance that the experiment would be limited to those patrol areas where stolen cars (occupied or unoccupied) are prevalent. Furthermore, the stratification assures that an experimental condition will not be repeated too quickly in a stratum, because a quick repetition will tend to reduce the recovery rate. [7]

[6] There is nothing special about the length of the control period. A longer period increases the sample size. Give that the recovery rate would be smaller during the control period, we would want the control period to be longer to provide a sufficient sample.
[7] Suppose that a patrol area receives LPR patrol on day 1. The effectiveness of another LPR patrol would be less on day 2 than it would on day 3; it would be less on day 3 than it would on day 4; and so on. This follows because a patrol on day 1 would remove cars that otherwise would be identified on day 2, while a new patrol on day 3 would provide some opportunity for stolen cars to be replenished. There is no need to

There is a second approach. If the patrol car equipped with LPR were only used on one day per week, then that car might be used to simulate patrol on other days. Specifically, the patrol car could be driven by a retired police officer or by a community-service officer who would simply record stolen cars as they are identified by the LPR system. (Recording of both is automatic. In this variation of the random assignment, the retired officer/community service officer would simply be collecting data rather than performing enforcement.) If the retired officer/community service officer reported those stolen cars when detected, this would be little different in theory from having an on-duty officer driving the LPR-equipped car.

If the retired officer or community-service officer did not report the stolen car to on-duty police[8], there are two problems. The first problem is that we would not have a metric for how long is required to recover a stolen car detected by the LPR system. This is a minor problem, however, because we can estimate that time based on the average time observed for the real-world use of the LPR system. The second problem is that the resulting data will not provide any information about the damage to the car that resulted from the delay between theft and detection. However, this can be estimated by studying damage as a function of recovery time. If the retired officer or community service officer reported the stolen car, these concerns would be moot.

Finally, we note that the value of using LPR to locate stolen cars depends on the frequency with which LPR is used to patrol an area. The more frequent the patrol, the greater the value, because LPR causes cars to be recovered more quickly than they would otherwise be recovered. Of course, the more frequent the patrol, the higher the cost. Presumably the cost of patrolling with LPR increases linearly while the benefit increases at a decreasing rate. The breakeven point can be inferred from the experimental evidence.

A random assignment experiment seems unnecessary for determining whether or not LPR increases the frequency with which an officer detects an occupied stolen vehicle. Even without random assignment, one could observe the frequency at which a routine patrol car identifies an occupied stolen car and the frequency with which a LPR-equipped patrol car identifies an occupied stolen car. Nevertheless, the same randomized design as was suggested for estimating how LPR reduces recovery time and reduces vehicle damage would apply to judging whether LPR-equipped vehicles identify more occupied vehicles per patrol period than non-LPR-equipped vehicles.

---

test how the effect diminishes with the frequency of patrol, because this can be inferred from the length of time that a car had been reported as stolen at the time that it was recovered.

[8] The Sacramento Police Department considered the prospect of hiring retired police or using community service officers to simply drive the LPR-equipped car. This would allow on-duty officers to perform regular police work until called by the surrogate officers. If Sacramento adopted this procedure, even for an experimental period, then the random assignment evaluation could be done with this configuration. From an evaluation standpoint, this would provide the same data as having on-duty officers drive the LPR-equipped car.

When researching the issue of identifying occupied vehicles, however, we suggest that the evaluator also perform a qualitative process analysis of what happens to people who are observed occupying stolen vehicles. How often are they prosecuted? Convicted? How often are their cases dismissed (or no charges filed) because they were in fact authorized to drive the vehicle?

We recommend the use of qualitative methods to investigate the use of LPR for collecting evidence from crime scenes and for monitoring the driving patterns of suspicious persons.

How many times would the technology be applied in 1 year?

The technology would be applied multiple times. When fully implemented, the application would be continuous.

Will modest but statistically significant effect sizes be detectable given sample sizes?

Although we are uncertain of estimates, it appears that a LPR patrol car will discover 1–2 stolen cars per shift. (We posed this question during our site visit to Sacramento.) If an LPR patrol car were to patrol on one shift per week, then the LPR car would detect 52–104 stolen cars per year. Patrols during the longer control periods might detect about the same number of cars. The question is whether or not the LPR patrol would identify stolen cars sooner and with less damage. Given the expense of an LPR system, only large effects are likely to be of interest. However, we cannot tell what is large and small, so any power calculations would be suspect.

However, an experimental condition with as few as 52 observations is small, and we seek ways to expand this number. If the LPR-equipped car could be driven 7 days per week, then the sample would be between 7x52=364 for one shift or 7x104=728 for two shifts each day. The size of the control sample would be about the same. This would certainly appear to be a sufficient sample to detect even a modest treatment effect. If not, the experiment might add an additional LPR-equipped car to the study.

We know that discovering an occupied stolen vehicle is a relatively rare event, but unless it is very rare, a sample of 364–728 stolen cars for the LPR-equipped and non-LPR-equipped recovered cars should provide sufficient power to judge whether or not LPR-equipped cars have a material effect on increasing the arrests and prosecutions of car thieves.

How many units—if any—would have to be procured for an evaluation?

If a single LPR-equipped car were available for seven shifts per week, the sample size would be adequate, and there would be no need for additional equipment. However, we are uncertain that the car could be put into service for that period because of unavoidable equipment failures. The cost of the LPR *reader is about $20,000–$25,000, so the purchase of one unit would not be a prohibitive expense. This presumes that the agency would pay for the patrol car, and as noted, someone has to drive this car: using retired officers or community service officers to drive the LPR-equipped car would require payment that is less than the cost of a patrol officer. If they were acting as data collectors (rather than as police adjuncts), their cost might be paid by a grant.*

What does a control/comparison group receive?

The control group would receive routine patrolling.

What kinds of data elements are available from existing data sources?

Existing data elements would list stolen cars, when they were reported as stolen, and when and where they recovered. We are uncertain if the reports include estimates of damage. We would also know when the LPR-equipped car was in service and where it patrolled. And, we would expect the records to record arrests, while prosecution records would tell about prosecutions and convictions.

What specific input, process, and outcome measures would they support?

We are uncertain that the data would provide estimates of damage. This part of the evaluation might have to be based on insurance claims by year, model, and time until recovery. Various insurance institutes collect these data and could probably be persuaded to provide estimates. There might be a data processing cost.

How complete are data records? (Attain samples if possible)

We have not collected samples, but the data needs are minimal.

Can user and/or target populations be followed over time?

Yes, a license plate can be followed over time (this is in fact one potential application suggested by the Sacramento police), although this does not seem to be an important element of the research design.

Can the dosage of technology used be identified?

Yes, the dosage (frequency of patrol with a LPR-equipped car) is an observable metric. We think that the more important question is whether this study could provide an estimate of the relationship between dosage and benefit. We have sketched how that relationship could be established with a simple mathematical model.

Can data systems help diagnose implementation problems?

Yes. A problem with the LPR technology is that the technology cannot always read a plate. This failure rate can be estimated for abandoned cars by examining the period following *patrol by the LPR-equipped car. Cars that were stolen before the LPR patrol, abandoned in the area patrolled by the LPR-equipped patrol, yet not detected would represent a potential failure to detect a plate.*

There is a potential problem. The car may have been stolen before the LPR patrol but abandoned (in the LPR patrol area) after the LPR patrol. But if that were true, then there would be a pattern to the cars missed by the LPR patrol: Those stolen most recently would have a higher rate of being missed by the LPR patrol, and one could adjust the statistics based on that observed pattern.

The CrimeConnect data could be reviewed to identify the rate at which the system reads partial plates.

What threats to a sound evaluation are most likely to occur?

We see no reason to suppose that LPR equipment will increase the rate at which stolen cars are recovered. We do expect that LPR equipment will reduce the time until recovery; we do expect that LPR equipment will reduce damage to recovered vehicles, and we do expect that the use of LPR technology will increase the apprehension rates of car thieves. Measuring the time until recovery is straightforward; measuring the damage to recovered vehicles is more demanding. We suspect, however, that the principal benefit from LPR-equipped patrols comes from reducing damage. Furthermore, the rate at which LPR-equipped patrol vehicles apprehend suspects from occupied vehicles is observable and supports an inference that LPR technology increases the rate at which car thieves are identified driving cars that they have stolen.

What changes is the site director willing to make to support the evaluation?

The evaluation is not disruptive of operations provided the Sacramento police would be willing to randomly assign LPR-equipped patrol cars to patrol on some days and not on others. An alternative plan of using retired officers or community service officers as data collectors would reduce the burden.

3. Overall

Would you recommend that the technology be evaluated? Why or why not?

Cars, especially those that are attractive to thieves, are remarkably expensive. The cost to the victim of an automobile theft is commensurate. Although theft for profit (for resale and for chop-shops) imposes the largest per-unit losses, the inconvenience of waiting for a stolen car to be recovered and the expenses of repairing a seriously damaged car are consequential. A technology that promises *to reduce recovery time and to reduce the damage from car theft has the potential to be cost-effective. Increasing the rate at which thieves are apprehended is also beneficial.*[9]

Restricting an evaluation to the benefit from reducing the cost of automobile theft would miss an important point, however. LPR is an emerging technology. Its use to recover stolen cars is obvious and an evaluation would be straightforward. But as an emerging technology, the LPR is in an early stage as a crime tool. The Sacramento Police Department believes that LPR can be used as a more general crime-fighting tool. If that proves to be true, then the cost of LPR technology would fall, increasing its value as a means for reducing the costs of automobile theft. We do not see how this emergent aspect of the technology would be evaluated experimentally, but a study of reducing the cost of automobile theft should include a process analysis of LPR as an emergent technology.

Furthermore, LPR has other extant applications. It is used to identify cars with outstanding tickets, leading to booting of the cars of offending owners. The sharing of equipment and data could both reduce the cost of equipment used for a single application and increase the effectiveness of that application. This sharing is not practical at this time in Sacramento, but consideration of sharing should be part of a process analysis.

What type of evaluation design would you recommend?

As noted above, random assignment—with or without expanding patrols—is recommended. Evaluating this technology lends itself readily to random assignment, and given the benefits of random assignment, it should be used. Qualitative analysis should be used to extend the evaluation of emerging uses of LPR technology.

References

Author unknown. Date unknown. "Denying criminals use of the roads: experience with Automatic License Plate Recognition and its impact on volume crime, terrorism and drug trafficking," *The Innovation Groups*, retrieved November 3, 2006, from http://www.ig.org/eCommunitySummit/UK,%20Association%20of%20Chief%20Police%20Officers.pdf.

Civica Platescan. Date unknown. "Automatic License Plate Recognition: How Is It Used?" Retrieved September 11, 2006, from http://www.platescan.com/used/default.asp.

[9] The eventual benefit would come from increased deterrence.

Kwasnicka, H., and B. Wawrzyniak, 2002. "License plate localization and recognition in camera pictures," *Artificial Intelligence Methods*, November 13–15, 2002, Gliwice, Poland.

McFadden, T. 2005. "Automatic License Plate Recognition Systems Emerge as a Law Enforcement Tool," *OLETC Today*, April 2005.

Ohio State Highway Patrol, Planning Services Section—Research and Development Unit. 2005. Automatic Plate Reader Technology.

PA Consulting Group. 2004. *Engaging Criminality—denying criminals use of the roads*, London, United Kingdom.

Parker, J.R., and P. Federl, 1996. "An Approach to License Plate Recognition," University of Calgary: Laboratory for Computer Vision. Retrieved September 11, 2006, from http://pages.cpsc.ucalgary.ca/~federl/Publications/LicencePlate1996/licence-plate-1996.pdf.

Valliappan, R., P. Sumari, and H. Kamarulhaili, 2004. "An Analysis of Traffic Monitoring by License Plate Recognition Method," *Enformatika*, 3: 80–82.

Wikipedia, the free encyclopedia. 2006. "Automatic number plate recognition." Retrieved September 11, 2006, from http://en.wikipedia.org/wiki/Automatic_number_plate_recognition.

Evaluability Assessment of Pawnbroker Databases

Staff Contact: William Rhodes
Abt Associates
617–349–2731
bill_rhodes@abtassoc.com

Myfanwy Callahan
Abt Associates
617–520–3056
myfanwy_callahan@abtassoc.com

NIJ Guidance

The National Institute of Justice (NIJ) has identified some key outcome variables and other parameters of interest for this technology, and has provided some guidance on possible evaluation designs. Applicants may depart from this guidance by providing appropriate rationale.

The advantages of a regional automated database relative to a fragmented, paper-based system of recordkeeping are patently obvious:

- Investigators can search for stolen property throughout a region, no longer being confined to their local jurisdictions.
- Investigators can track suspects in terms of their fencing patterns and the kinds of items they steal.
- Investigators can identify pawnbrokers who continually receive large quantities of stolen goods.
- Regional victims have a greater likelihood that their stolen property will be returned.

Acquiring the data that documents these gains is a formidable challenge, however, because pre-database records are scattered over local jurisdictions in paper storage. These challenges apply equally to all standard designs. Other challenges to experimental and quasi-experimental designs are also noted in this evaluability assessment.

Technology Summary: Pawnbroker databases are used to provide property crime investigators with electronic access to records of all pawnbroker transactions in their region. These databases allow investigators to search for property, individuals, or suspicious behaviors across multiple jurisdictions. Previously this investigation was paper based; the new system frees investigators' time for deeper investigations.

Scope of Evaluation: Two evaluation types are possible: a randomized design and a quasi-experimental design. In a randomized design, outcomes from a team or teams of investigators using the system would be compared against outcomes from a team not using the system. However, police agencies with access to pawnbroker data would likely be reluctant to restrict access for a control group, and isolation of the experimental and control groups may prove difficult. An alternate design would be to analyze time panel data from Minnesota's Automated Pawn System (APS) to observe changes in rates of property recovery among counties as they were added to the system.

Summary of Evaluability Assessment Activity: The evaluability assessment was composed of a review of available documents and printed materials on each of the systems in use today as well as indepth interviews with users and directors of the three regional pawnbroker databases; the Regional Pawn Data Sharing System (RPDSS) (Washington, D.C. area), the Automated Pawn System (Minnesota, Wisconsin), and the Florida Integrated Network for Data Exchange and Retrieval (FINDER). A site visit was conducted at the Montgomery County (Maryland) Police Department, users of the RPDSS.

Finding: Although much anecdotal evidence exists as to the benefits of pawn databases, no evaluation has yet directly shown benefits in terms of labor saved, property value recovered, or increased arrest rates of property criminals. A quasi-experimental evaluation in a data-rich system such as Minnesota's APS could provide valuable information on the benefits of this technology.

1. Initial Screening

Background

Brief technology description

Within the past decade, electronic data-sharing systems have emerged as a tool to help law enforcement agencies track and retrieve stolen goods that criminals have attempted to launder through secondhand markets. Pawnbroker venues are of special interest to law enforcement for a variety of reasons. Foremost, pawnshops can provide a relatively quick and easy way to dispose of stolen goods, which, according to some researchers, may even create a market incentive for theft (Fass and Francis, 2004).

Pawning occurs when an individual offers an item of value as collateral for a loan (Moseley, May 2005). Pawning may not differ substantially from other secondhand markets, according to some researchers, except perhaps in the prevalence of pawnshops in certain areas and their ability to allow customers to unload a variety of goods quickly (Fass and Francis, 2004). Most States already regulate pawnshops and require pawnbrokers to collect some basic data on any merchandise that passes through their stores. However, the details of the regulations themselves may vary substantially from State to State, especially concerning the type of information or data that pawnshop owners are required to collect on the customers who deliver goods to their stores (Moseley, May 2005). In addition, it is not clear how successful States have been in tracking pawnshop transactions and enforcing regulations. Older tracking systems for pawnshops frequently use handwritten slips that either pawnshop owners or law enforcement officers themselves must manually enter into databases. The process, by many accounts, can be difficult and time-consuming, especially given the high volume of goods that can pass through a pawnshop and the frequent understaffing of those shops. All in all, older tracking systems that rely on hand-written data frequently do not accommodate the quick turnaround time that law enforcement needs to solve theft cases (Fass and Francis, 2004; Business Watch International, 2006).

Electronic data-sharing systems, by contrast, can facilitate the tracking of stolen merchandise by being substantially quicker and easier to use. Pawnbroker databases can operate through a variety of interfaces—either through specialized software that pawnbrokers use to upload basic data that law enforcement can access immediately, or through a similar Internet-based data entry system. Law enforcement can then search pawnbroker databases through a variety of search functions that identify items by serial numbers or identifying markings (when available) or by description (Hurley, 2000). When States require pawnshops to collect information on their customers, law enforcement can use databases to search by seller profiles. Seller profiles can be especially useful in targeting frequent pawners, who may be more likely to traffic in stolen goods (Fass and Francis, 2004; Dougherty and Liao, 2005). The difficulties that hinder identifying stolen goods through pawnbroker databases are the same difficulties that confront theft cases in general: many stolen goods do not have identifying markings or serial numbers that owners have recorded. To increase the chances of recovering stolen goods, some pawnbroker databases have widened their net to include other secondhand goods outlets, like flea markets, precious metal or antique dealers, or even eBay.

What is the background/history of this technology?

Maturity (i.e., lab prototype? Field rollout? Multiple generations/manufacturers?)
The pawnbroker databases available today are actively in use in many departments throughout the country.

Time in the field?
The oldest pawnbroker database system is the Automated Pawn System of Minnesota (and now western Wisconsin). APS has been in place since 1997, though the number of districts using the system has grown over the years. Two of the most frequently discussed databases, LeadsOnline.com (which serves a variety of secondhand goods markets) and RPDSS, have been in use since 2000 and 2004, respectively.

Prevalence in the field? (Is site a first/early adopter?)
Although the actual number of municipalities or law enforcement agencies that use record management systems is unknown, a review of recent literature on pawnbroker databases suggests that such systems have become increasingly popular and visible within the past decade. In addition to RPDSS, used in the Washington, D.C., metropolitan area, other databases include APS in Minnesota and western Wisconsin, the FINDER system in Florida (a comprehensive system including more than just pawn transactions), and LeadsOnline. LeadsOnline, the pawn database supported by the National Pawnbroker Association, has been aggressively marketed to law enforcement departments across the Nation and is now used in at least 12 States. The company's Web

site boasts a clientele of nearly 600 law enforcement agencies that subscribe to its services throughout the U.S.

What do we already know about technologies like these?

The preliminary evaluations of some pawnbroker databases have been markedly positive. According to a press release by the Metropolitan Washington Council of Governments, RPDSS resulted in more than 1,300 arrests and the recovery of more than $4 million in stolen property (MWCG, 2005). Other evaluations of RPDSS offer examples of particular cases in which the database was instrumental (according to the accounts of law enforcement officers) in apprehending suspects and recovering stolen goods (Hawdon and Ryan, 2006). The database, for example, can show when particular individuals are pawning items frequently—sometimes several times a day—and can tally pawning transactions and identify trends that may point to criminal activity (Moseley, June 2005). Promising data like these may serve to further popularize such systems within the next few years.

However, several evaluations of pawnbroker databases have focused primarily on the perceptions of law enforcement officers who use such databases. Furthermore, the response rates of some studies have been low, and those studies relied heavily on anecdotal data and respondent recall (Moseley, June 2005). The success of pawnbroker databases depends on a number of factors, including an agency's ability to purchase and maintain such a database; the ease of use of the database; the agency's ability to train detectives, investigators, and pawnbrokers on the uses and functions of the database; and the degree of buy-in from pawnbrokers themselves. This last issue may be especially significant, given the documented objections on the part of some pawnbroker representatives to collecting personal, identifying information on their customers and turning it over to the police. (Although several States already require that pawnshops collect customer data, compliance and enforcement may have been inconsistent in the past.)[10]

What could an evaluation of this technology add to current knowledge?

There is ample anecdotal evidence of the benefit of this technology but very few evaluations have measured real benefits in terms of labor saved, property recovered, and reduced crime. The few evaluations that have been conducted have extrapolated value saved from surveys with relatively low response rates. A well-done evaluation would be able to quantify benefits much more accurately.

Which audience(s) would benefit from this evaluation?

- Law enforcement agencies
- Pawnbroker industry (potentially)
- Database companies and developers who sell such technology

[10] In 2001, after law enforcement in Fort Worth, Texas began using LeadsOnline.com, representatives of the city's pawnbroker industry objected to the increased scrutiny of their operations, and cited privacy rights within the Texas Constitution to question whether police agencies had probable cause to force businesses to disclose personal information (Office of Consumer Credit Commissioner, Texas, 2001).

What could they do with the findings?
Showing the benefit to law enforcement and thus to the community of these systems may allow law enforcement to better defend the use of these systems against political attacks.

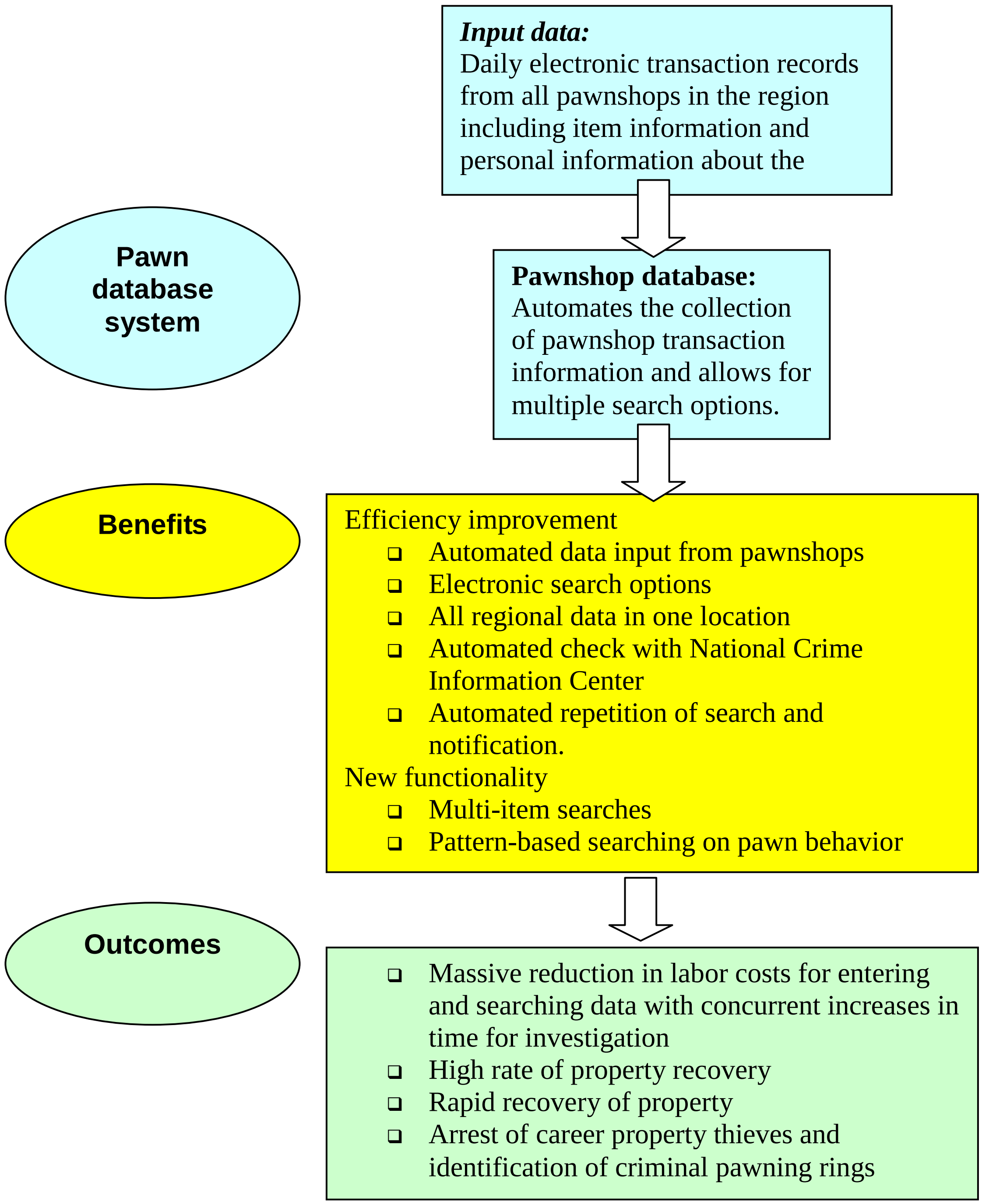
At what stage of adoption/implementation is the technology?

For the initial site visit we focused on the RPDSS system as implemented in the Montgomery County (Maryland) Police Department. The technology has been in place there for a year and a half. It has been actively used by property crime detectives during that time.

What efficiencies or primary/secondary outcomes are expected?

Sketch the logic by which technology use should affect goals (see exhibit 1).

Exhibit 1. Pawnbroker Database Logic Model



Are there operational alternatives that could be used for comparison?

The operational alternative to these systems is the paper-based system that requires investigators to sift through tickets by hand or enter tickets by hand into a department-specific spreadsheet. A comparison of police districts that use the system and that do not use the system is unlikely to be productive, because an investigator would be challenged to establish that the comparison district is a suitable counterfactual.

A more useful approach would rest on a pre-post design where the same police department would compare outcomes before and after the system had been implemented. An even more useful comparison would use panel data to compare the outcomes across districts that varied in when they introduced the system. Minnesota is especially attractive given the active use of a pawnbroker system across the State and in bordering Wisconsin.

Other comparisons might be useful. Pawnbroker systems provide data; but analyses are the means that convert the data into information. A useful comparison might examine how different agencies use their data: Can they search for patterns according to seller identity? Once they find a pattern, what do they do with the information?

Is the site interested in being evaluated?

Montgomery County, Maryland, police are willing to participate in an evaluation. They have already taken part in a user survey of the system conducted by researchers at Virginia Tech, entitled "An Evaluation of the Regional Pawn Data Sharing System" (Hawdon and Ryan, 2006). This evaluation did not directly measure any outcomes, but estimated the benefit of the system based on respondent answers to an open-ended survey question.

Is the site planning an evaluation?

No evaluation of the system is currently planned.

Data Sources

What data systems exist that would facilitate evaluation?

The only statistics readily available from the system are statistics on usage, such as the number of searches conducted. No outcome statistics are reported within RPDSS. Each county is responsible for its own outcome reports. These may be annual, semi-annual, or at best monthly reports documenting outcomes such as number of items recovered, number of arrests, and (pawn) value of items, laboriously gathered from paper-based case files. Systemwide outcomes have been difficult to gather. A report on RPDSS success from June 2005 attempted *to gather information on outcomes in an e-mail survey of investigators (Moseley, June 2005). Those who responded reported significant benefits but represented only 16 of the 474 investigators using the system (3 percent).*

What key data elements are contained in these systems?

None of the key outcome events are tracked by RPDSS.

Are there data to estimate unit costs of labor and capital?

Few data are available to estimate costs of labor and capital. RPDSS was developed on a $1-million grant partnering NIJ with the Council of Governments in the National Capital Region. Ongoing system costs involve the storage of data on servers, the maintenance of the system, and the service contract with the vendor, BWI. Although data on these costs could be identified, no data on the cost of training or reduced labor costs for data-entry are currently available for RPDSS.

Are there data for possible comparison technologies or other solutions?

Quantitative data on comparison technologies is not currently available. The alternative to RPDSS is a paper-based system requiring hand entry of data into a database. The benefit of RPDSS is a shift of labor from data-entry to investigation, which may not lead directly to reduced labor costs. Furthermore, the outcome data for the paper-based system will be as difficult to obtain as that for the RPDSS system. It may be possible to involve a county in the National Capital Region that has not yet joined RPDSS and compare it with a county already in the system. However, RPDSS has been very successful in recruiting all of the populous counties near the capital. Counties not currently in the RPDSS system are likely to have lower populations and fewer resources than counties within the system.

In general, how useful are the data systems to an impact evaluation?

The data systems currently available within RPDSS are not very useful to an impact evaluation. Data on outcomes would have to be generated before any evaluation could be conducted. Counties vary in the statistics they use to document outcomes, so there is no standard. An evaluation could demand significant time resources of participating police departments.

2. Site Visit Screening

The Intervention

Has the organization implemented a policy or training for the technology's use?

Montgomery County provides training for all officers and pawn investigators before they gain access to the system. The training is minimal because the system has been designed to be user friendly. Additional questions are answered by the chief pawn investigator.

Who are the users?

The users are primarily property crime investigators, typically one per county. These systems greatly increase efficiency in their work and have received wide praise. Sworn officers who choose to do so may also have read-only access to the database to conduct their own investigations into property crimes.

Who/what are the targets?

The target is any use of pawnshops for the purposes of property crime. The pawn database identifies stolen property among pawned items and property criminals who use pawnshops to sell stolen items.

Who/what gets excluded as a user or target?

Property criminals who sell stolen property directly or personally use the stolen items will not be identified. Only those who come in contact with pawnshops can be identified. Furthermore, criminals who use pawnshops outside the larger National Capital Region will not have their transactions noted in the database. Because pawnshop tickets include personal descriptions of each pawner, criminals who use false names may *still potentially be identified with this system.*

Have the characteristics of the user or target population changed over time?

Both the user population and the target population have grown as RPDSS has become more widely adopted by districts in the National Capital Region. Additional districts may join in partnership in coming months, increasing both the number of departments with search access and the base of pawn data available for search.

What values/outcomes do users see/envision in the technology?

The most significant value observed by users of the pawn database is the increase in efficiency. Prior methods included entering hundreds of paper tickets into a district-specific database. This would require at least one full-time employee. Oftentimes entry of paper tickets was delayed and incomplete due to limited manpower. Furthermore, searching was extremely time-consuming, including the check of both the database and the unfinished paper tickets. Because searching was very time-consuming, searches were not conducted *multiple times even though a stolen item might take days or weeks to appear among pawned items.*

The pawn database system fully automates data entry and greatly enhances searching. Searches can be automated to repeat several times and automatically notify the detective of a hit. All items with a serial number are automatically checked against NCIC records of stolen property. Data is accurate and timely, allowing stolen items to be recovered much more quickly than in the past. The dramatic reduction in man-hours required for data entry and search has allowed property crime detectives to spend a greater amount of time on investigations. This also results in an increase in recovery of smaller items such as cameras, which, in the past, were too small to warrant the man-hours required for investigation.

Furthermore there are additional functionalities that were not possible in the past. For the first time data is available from other districts, allowing property crime detectives to recover property that was taken outside their district or even their state before it was pawned. This is a common tactic and this functionality has greatly increased pawn detectives abilities. Multi-item searches can be conducted to identify multiple items stolen during one robbery even if these items were pawned at different locations. Searches can be conducted based on suspicious behavior, such as the pawning of more than $20,000 worth of merchandise within one month, even if the individual pawned items at several different pawnshops. Other behaviors such as multi-district pawning or pawning of many new-in-box items in a short period can also be searched. Searches can also be automated to notify detectives if former property criminals break parole by pawning items. The pawnbroker database essentially enhances property crime detectives' efficiency in property recovery and ability to identify career property criminals.

The outcome of these increases in efficiency and new functionality is a much higher rate of recovery of property and identification and arrest of property thieves.

What are the limitations/obstacles to using the technology?

Property crime detectives can only identify stolen property within the districts participating in the database system. Criminals may relocate their transactions to areas that are not yet in a pawn database. Also, not all pawn systems allow searches based on behavior patterns. All systems allow searches based on specific property items, but some of the most productive searches for investigation may not be available.

What outcomes could be assessed? Using what measures?

The operative question is "What outcomes should be assessed?" One might think of a pawnbroker database as a tool for reducing property crime or, at least, *for increasing the rate at which property crimes are cleared and thieves are prosecuted. If this were the most important question, then an evaluator might examine reported thefts in a district, or clearance rates for reported thefts, or both.*

One might think of a pawnbroker database as a tool for recovering lost property. If this were the most important question, then the investigator might examine the rate at which property is recovered. The investigator might monetize the recovery using the value of the property.

Still a third question might presume that neither of the above two questions are important, and that the real key to a pawnbroker database is that it provides the police with a less expensive method of performing the same job that they performed using paper records.

If the first question were important, than an investigator might use public records of reported thefts and of crimes cleared by arrest. Any investigator should be aware of the limitations of arrest/clearance data.

If the second question were important, than an investigator would want a list of stolen property and a list of recovered property, ideally with enough description to monetize the dollar gain to the victim (or the insurance company as his or her proxy) from recovering the stolen property.

If the third question were important, the answer would seem to depend on a qualitative evaluation. How has investigator time been increased by his or her being relieved from coding data? What productive use of his or her time has resulted?

Of course, an investigator might be interested in all the above questions. Each might be incorporated into an evaluation.

Designing a Study

Are there other operational environments for which the technology is well suited?

What are the constraints in such environments?
A pawnbroker database would seem to be most effective when the pawnbroker is honest and the pawner is dishonest. Other than avoiding time required to report to the system, the pawnbroker has no incentive to hide or falsify transactions. If the pawnbroker is dishonest, serving as a launderer for stolen goods, then both the pawnbroker and the pawner have an incentive to obscure transactions.

Similar systems exist in other settings. With respect to firearms, the National Rifle Association summarizes Federal law:

> Sale of a firearm by a federally licensed dealer must be documented by a Federal Form 4473, which identifies and includes other information about the purchaser, and records the make, model, and serial number of the firearm. Sales to an individual of multiple handguns within a 5-day period require dealer notification to the Federal Bureau of Alcohol, Tobacco and Firearms. Violations of dealer recordkeeping requirements are punishable by a penalty of up to $1,000 and 1 year's imprisonment (National Rifle Association, 2006).

This system might be used to track the origin of guns used in crimes, but a 2004 report by the Inspector General of the U.S. Department of Justice indicates that the system does not function well (U.S. Department of Justice, Office of the Inspector General, 2004)

Do the technology "events" permit randomly generated applications of the technology?

If not, can comparison samples be formed? With what difficulties?
A random design experiment would be difficult to establish in Montgomery County or anywhere else where theft investigators already have access to a pawnbroker database. Conceivably, an evaluator could randomly assign a pool of police theft investigators to a group that has access to the pawnbroker database and another group that lacks access to the database. Still, it seems unlikely that an agency that had access to pawnbroker data would agree to restrict some of its investigators from having access.

It is possible, however, that a Federal grant could pay for access to a pawnbroker database for a large agency that did not already have access to pawnbroker data. A condition would be that the agency would randomly assign its investigators to two units. One would pursue theft investigations using traditional means, e.g., without access to the pawnbroker data. The other would have access to the pawnbroker data. At the end of the experiment, all investigators would have access to the pawnbroker data.

Of the three questions posed earlier, this random assignment design could answer the second (Does the property recovery rate increase?) and the third (How is investigator time enhanced?). This design could partially answer the first question (Does the clearance rate increase?) but it could not answer the question of whether theft rates are sensitive to use of the pawnbroker data. It is, however, difficult to see how the treatment group could be isolated from the control group, so the estimated treatment effect would be biased.

Moreover, this randomized design does not deal with an additional important issue: How does the success of a pawnbroker database increase as the coverage increases? It would be difficult to answer this question experimentally.

Montgomery County does not seem to offer an ideal location for a quasi-experimental design. An evaluator might consider the State of Minnesota as an alternative. Minnesota is an attractive site for a quasi-experiment for several *reasons. Minnesota provides panel data from across the state and overlapping into Wisconsin. (Panel data comprise a time-series of cross-sectional units.) Counties introduced pawnbroker data at different times, so the time-series aspects of these data would be informative. We would expect favorable outcomes in county X as county X introduces a system and favorable outcomes as county Y that lies next to county X introduces its own system. Because the Minnesota data date back to 1997, the time series would be longer than a decade (by the time an evaluation is funded), and a time series of this length provides the prospect of a strong quasi-experimental design based on instrumental variables.*

We are uncertain of data availability from Minnesota. County-specific theft data and clearance rates are almost certainly available, so the first question would be answerable. We do not know if the counties could provide summary data on theft recovery rates, although we presume that such statistics could be assembled from public records. Alternatively, insurance companies might be willing to provide recovery rates for items reported stolen across Minnesota counties over time. Answering the third question might best be done using qualitative techniques—such as questionnaires. Although others have found survey response rates to be low, the survey need not be burdensome because it is only intended to ask about time savings, and techniques to promote followup (multiple mailings and telephone followup) could be implemented across Minnesota police agencies without high costs.

The strategy proposed for Minnesota could also be applied to RPDSS in the Washington, D.C. area. The constraint is that the system has been in place only since 2004, so it does not provide the same advantages as the lengthier time series in Minnesota.

How many times would the technology be applied in 1 year?

The technology is used continuously.

Will modest but statistically significant effect sizes be detectable given sample sizes?

Given the prevalence of property crime, we would expect that sample sizes will be adequate to detect a moderate or even small treatment effect. This assumes that pawnbrokers play an appreciable role in laundering stolen property.

How many units—if any—would have to be procured for an evaluation?

No procurement would be required unless an evaluator proposes a random assignment design.

What does a control/comparison group receive?

The comparison group would receive theft investigations based on a paper record system and whatever reporting requirements were in place at the time that the pawnbroker database was adopted.

What kinds of data elements are available from existing data sources?

The pawnbroker database only reports pawnbroker loans. Outcome data would have to come from other sources.

What specific input, process, and outcome measures would they support?

The existing sources only collect process-type measures such as the number of times a system has been accessed and the number of items that have been placed on hold by an investigator. The data in the system contain specific information on the value of items pawned and descriptions of both the pawned items and the individual who pawns them. However, nothing included in the data distinguishes legitimate pawn transactions from illegal pawn transactions. The pawn database is merely a tool to help investigators identify suspicious transactions.

How complete are data records? (Attain samples if possible.)

This is unknown.

Can user and/or target populations be followed over time?

Yes, this is a feature of a pawnbroker database. Loans to the same individual are indicative of theft for profit.

Can the dosage of technology used be identified?

Yes, depending on the design. At one level, the dosage depends on how contiguous agencies have implemented the system. At another level, dosage depends on the number of investigators that a police agency allocates to investigating theft relative to the amount of theft reports. Regardless of the utility of the pawnbroker database, theft investigations take time, and we would expect that outcomes will be sensitive to the inputs of investigator time.

Can data systems help diagnose implementation problems?

No, this seems unlikely.

What threats to a sound evaluation are most likely to occur?

A random design experiment suffers from the problem that the control subjects (investigators who lack access to the pawnbroker data) would be contaminated *by treatment subjects (investigator who would have access to the data). This problem might be reduced by randomly assigning sites to receive pawnbroker databases, but the expense of such a study is prohibitive, and anyway, treated counties would still contaminate untreated counties.*

The problem with a quasi-experiment is that one can never be sure that the counterfactual is valid. A lengthy time-series is an ameliorative, however, for two reasons. The first is that a panel design allows the evaluator to based inferences on a difference of difference approach. The second is that a lengthy panel may lend itself to the use of instrumental variables.

What changes is the site director willing to make to support the evaluation?

We are uncertain that any changes are required.

3. Overall

Would you recommend that the technology be evaluated? Why or why not?

Yes, either a random design experiment or a quasi-experiment would yield useful information about how pawnbroker databases affect theft and theft recovery. The utility of a pawnbroker database has face validity. Although pawnbrokers are not the only outlet for laundering stolen property, they are both convenient and visible to thieves. Alternative sources, such as third parties who purchase for resale, would be much harder to locate (because their operations are covert) and likely are less willing to deal in small commodity items. One would reason that removing pawnbrokers as a convenient source for laundering stolen property could greatly reduce the incentive for theft for profit.

Furthermore, if either Minnesota or some other location could provide suitable and readily available outcome data, an evaluation should be able to provide a compelling estimate of treatment effectiveness.

What type of evaluation design would you recommend?

As noted above, we recommend either a random design experiment or a quasi-experimental design in a place that provides suitable panel data.

References

Business Watch International (2006). "Regional Pawn Data Sharing System." Retrieved October 18, 2006, from http://www.bwipolice.com/ProductsandServices/RegionalPawnDataSharingSystemRPDSS/Introduction/tabid/105/Default.aspx

Dougherty, M., and Liao, H. (2005). "Examining Pawnshop Activity in the St. Louis Area Using GIS." Greensboro, NC: The IACP 29th LEIM Conference, May 2005.

Fass, S., and Francis, J. (2004). “Where Have All the Hot Goods Gone? The Role of Pawnshops.” *Journal of Research in Crime and Delinquency* 41(2): 156–179.

Hawdon, J., and Ryan, J. (2006). “An Evaluation of the Regional Pawn Data Sharing System.” Blacksburg, VA: Department of Sociology, Virginia Polytechnic Institute and State University.

Hurley, J.T. (2000). “Property Crime and Pawnshops: Correlation or Coincidence?” *Police Magazine* (May). Retrieved September 15, 2006, from http://ci.ftlaud.fl.us/police/hurley2.html

LeadsOnline.com (date unknown). At a Glance: LeadsOnline (Law Enforcement Automated Database Search). Retrieved September 13, 2006, from http://www.leadsonline.com

Metropolitan Washington Council of Governments (2005). “Regional Pawn Data Sharing System Among First in Nation”. Retrieved October 18, 2006, from http://www.mwcog.org/news/press/detail.asp?NEWS_ID=138

Moseley, B. (May 2005). “NIJ’s In Progress Review: Regional Pawn Data Sharing System (RPDSS).” Washington, D.C.: National Institute of Justice.

Moseley, B. (June 2005). “Regional Pawn Data Sharing System—Success Reports.” Washington, D.C.: National Institute of Justice, CommTech Program.

National Rifle Association. (Date unknown). “A Citizen’s Guide to Federal Firearms Laws.” Retrieved November 8, 2006, from http://www.nraila.org/GunLaws/FederalGunLaws.aspx?ID=60

Office of Consumer Credit Commissioner (2001). Pawnshop Electronic Data Transfer (EDT) Project. Fort Worth, TX: Public Meeting, November 7, 2001. Retrieved October 17, 2006, from http://www.occc.state.tx.us/pages/industry/Pawn/edt/FtWorth.html

U. S. Department of Justice, Office of the Inspector General. (2004). “Inspection of Firearms Dealers by the Bureau of Alcohol, Tobacco, Firearms and Explosives: Executive Digest,” Report Number I–2004–005 (July) Retrieved November 8, 2006, from http://www.usdoj.gov/oig/reports/ATF/e0405/exec.htm

Evaluability Assessment of Radio Frequency Identification Device (RFID) Use in Correctional Settings

Staff Contact: Steven T. Wieland
Telecommunications Manager
Ohio Department of Rehabilitation and Correction
614–387–0863

NIJ Guidance

The National Institute of Justice (NIJ) recommends an evaluation of Radio Frequency Identification Device (RFID) technology in the site assessed below (or other appropriate correctional settings). In particular, NIJ is interested in a combined quantitative and qualitative study of inmate behavior and safety at the Northeast Pre-Release Center (NEPRC) facility. It appears likely that an interrupted time series design could identify RFID's impact on fights and other infractions. Further, NIJ is interested in RFID's impact on officer efficiency (e.g. the monitoring and tracking of inmates) and investigations (e.g. substantiating allegations of misconduct).

Applicants may depart from this guidance by providing appropriate rationale.

Technology Summary: Radio Frequency Identification Device technology has been in existence for more than 30 years, but its application in correctional settings is relatively new, dating back only to 1997. The use of RFID in correctional facilities is designed to improve prison management, offering a more efficient means of locating inmates, confirming counts, and alerting officials to escapes. Overall, it holds promise for improving inmate behavior (i.e., reducing infractions and assaults) and for providing a safe and secure environment for staff and inmates. Two Ohio adult correctional facilities, Ross Camp and Northeast Pre-Release Center, were selected as the focus of this evaluability assessment. Both facilities have fully operational RFID, and because each used a different vendor to install RFID, an evaluability assessment of both provides an opportunity to learn about differences in implementation and potential outcome measures.

Scope of Evaluation: The overall conclusion from this assessment is that an evaluation of the use of RFID technology at NEPRC that employs an interrupted time series design (both impact and process evaluation) is currently feasible. An impact evaluation at Ross Camp is not feasible, due primarily to lack of outcome data. Absent an impact evaluation, a process evaluation at both facilities is still recommended.

Summary of Evaluability Assessment Activity: The assessment of the feasibility of evaluating RFID technology began with a review of the literature and a Web-based search to identify RFID vendors and agencies that are currently using RFID. In addition, technology experts at the National Law Enforcement and Corrections Technology

Centers (NLECTC) and NIJ staff were interviewed. Our research revealed that two vendors, ElmoTech and TSI Prism/Alanco (TSI), are the primary providers of RFID technology for correctional institutions. ElmoTech and TSI provided the Urban Institute (UI) with a list of agencies that are using, or are in the process of implementing, RFID to monitor inmates and/or staff. Currently more than 4,200 inmates and staff in 7 States are tracked using RFID technology.

Additional screening, including input from vendors, revealed nine mature and four planned applications of RFID technology in correctional settings. On the basis of the background information compiled and discussions with NIJ, it was mutually decided that the Ohio Department of Rehabilitation and Correction's (ODRC's) Ross Camp and NEPRC would be the locations for the RFID site visits.

1. Initial Screening

Background

Describe the technology. What is the background/history of this technology?

RFID technology has been in use for more than three decades, mostly in the context of inventory tracking. In recent years, both the use and number of applications of RFID have grown exponentially. The use of RFID technology by Wal-Mart and the Department of Defense for inventory and supply-chain management has fueled the growth and use of this technology throughout other industries (Justice Technology Information Center 2005). Moreover, significant developments in the technology and reductions in cost have led to an open standard system that can be used for any application and applied to any object (Beck, 2006). RFID technology has been implemented in various retail and commercial industries to prevent theft (Justice Technology Information Center, 2005). In addition, in 2001 the United Kingdom implemented the Chipping Goods Initiative in an effort to reduce the cost of property crime, relieve pressure on police resources, and trace the ownership of stolen goods (Adams, 2004; Home Office, 2006). RFID technology has also been recognized for use within corrections, law enforcement, and even homeland security (Justice Technology Information Center, 2005).

The first application of RFID in a correctional setting was in 1997 at California State Prison, Corcoran, where it was used to track staff for safety purposes. In 2002, RFID technology was piloted at a Michigan juvenile facility, its first known application for use with inmates (Reza, 2004). To date, RFID has been implemented (or is in the process of being implemented) for use with inmates in 13 facilities in 7 States across the Nation (see attachment A for complete list of sites, including facility name, location, type of facility, number of RFID units, year of implementation, implementation status, targets, and vendor).

As used in correctional settings, RFID technology consists of three components: (1) an RFID chip, which is embedded in a bracelet or anklet that also has the ability to detect body mass index (issuing an alert if the bracelet is removed or is not within one finger's width of the skin); (2) a series of Data Extension Units (DEUs), which operate like antennas to read and transmit information stored on the RFID chip; and (3) computer software that enables correctional officials to document—in almost real time—the whereabouts of inmates. With a sufficient number of DEUs in a facility, RFID technology has the ability to track the locations of inmates every 30 seconds, with software mapping the locations and movements over time in a fashion similar to Global Positioning System (GPS) technology, but at a fraction of the cost.

On its most basic level, RFID use in corrections can help confirm counts of inmates and serves as an additional perimeter control device. The software can also enable more sophisticated applications. Correctional officials can enter information on inmates' schedules and where they are supposed to be at certain times of day (e.g., classes, cafeteria, cells), issuing an "out of place" alert if inmates deviate from those schedules. The software can also be programmed to issue alerts when certain inmates, such as rival gang members, are in close proximity to one another. And, because the system maintains historical data of inmates' locations, RFID can also be a useful tool for investigating assaults, pinpointing which inmates were at a location where an assault took place and aiding in the substantiation of allegations of sexual and other assaults.

Maturity (Time in field)

RFID has been used in correctional facilities since 1997.

Prevalence in the field

Two vendors, TSI and Elmotech, currently offer RFID implementation in correctional settings. Among nine correctional agencies that have already implemented RFID, ODRC is one of the earliest adopters of RFID technology, with RFID fully operational in two separate correctional institutions, each of which implemented RFID using a different vendor. The ability to examine applications of two different vendors' technologies in one site visit prompted us to select Ohio as the focus of the evaluability assessment.

What do we already know about technologies like these?

RFID use in correctional settings is relatively new and has not been subject to rigorous evaluation. The only study identified through an extensive literature review was an assessment of implementation of RFID at a Michigan juvenile facility, which found that, during a 3-year test period, no escapes occurred and violent incidents were reduced by 65 percent compared to pre-RFID incidents (Reza, 2004). Beyond this one assessment, anecdotal evidence supplied by vendors, and personal impressions of the few correctional agencies that have already invested in RFID technology, very little is known about this technology's potential impact on efficiencies and outcomes.

What could an evaluation of this technology add to current knowledge?

An evaluation of RFID will enhance knowledge of how to apply this technology to improve prison and jail operations and manage correctional populations.

Which audience(s) would benefit from this evaluation?

An evaluation of RFID would benefit directors of departments of corrections, wardens, and line-level officers. The application of RFID in correctional settings is relatively new, and Ohio represents one of the earliest adopters. Word of mouth has attracted representatives from departments of corrections across the Nation to visit Ohio and learn more about the technology, but much of the information they collect is based on perceptions rather than any hard numbers on impact and costs/benefits.

What could they do with the findings?

There is much to be learned and documented about the process of implementing and using RFID in correctional settings that would be of use to the corrections community and help guide DOCs in making an educated investment in RFID rather than relying solely on the information provided by vendors. Agencies that have already invested in RFID would naturally be interested in knowing whether it has an impact on prison management as well as the various uses of RFID in a correctional setting. Agencies contemplating investing in RFID would also be interested in these findings. For example, if an RFID evaluation demonstrates that it is effective in both detecting inmate misbehavior as well as possibly discouraging it, more correctional agencies might consider investing in it.

At what stage of adoption/implementation is the technology in the targeted site?

ODRC has enjoyed a long history of being at the forefront of innovative correctional practices. ODRC Director Terry Collins first became interested in the possibilities of RFID during his tenure as Director of Prisons. He was particularly interested in installing RFID for perimeter control around Ross agricultural camp, a correctional institution of approximately 350 inmates who run a full farm operation. Because Ross operates as an honor camp, Director Collins was interested in testing the technology for tracking, scheduling, and alerting correctional staff to out-of-place inmates and perimeter violations. Collins was also interested in RFID's capacity to support investigations of allegations of staff assaults on inmates, as well as inmate-on-inmate assaults. Shortly after releasing a request for proposals (RFP) for Ross, Director Collins secured money through the Prison Rape Elimination Act (PREA) to implement RFID at the Northeast Pre-Release Center, a 570-inmate women's institution in Cleveland. Although the focus of the RFID technology at NEPRC was similar to that at Ross, a greater emphasis was placed on preventing and supporting investigations of sexual assaults. In both correctional facilities, only inmates are currently equipped with RFIDs, but ODRC is contemplating using RFID-equipped identification cards for staff sometime in the future.

Ross Camp is a 350-inmate mixed-security institution that neighbors the 1,600-inmate medium-security Ross Correctional Institution in rural Chillicothe, Ohio (approximately 30 miles south of Columbus). The camp is part of a 1,800-acre working farm where inmates raise and slaughter cattle used to feed inmates. The camp borders on a Veterans Administration hospital and a high school, with a river 400 yards to the east and a major highway nearby. As an “honor camp,” inmates are free to move about the facility and surrounding campus, and frequently check in and out of the facility to report to and return from their farm work. In 2004, ODRC issued an RFP for RFID implementation at Ross Camp and selected TSI as the vendor. RFID is used at Ross Camp primarily as a means of enhancing perimeter control based on concerns about escapes, with a secondary use in determining whether inmates had reported to school and other programs and, if not, where they are located. RFID has also been used to identify who ate (for diabetics) and to prevent “doublebacks”—inmates getting back in the cafeteria line for a second meal. Ross has little in the way of disciplinary issues, so correctional officials did not believe that RFID would reduce inmate violence.

The Northeast Pre-Release Center is a 570-inmate minimum-/medium-security women’s prison located in Cleveland, Ohio, with an average inmate stay of 24 months. It is a dormitory-style facility with two, four, or six cots to a room. In 2005, ORDC issued an RFP for RFID at NEPRC and selected Elmotech as the vendor. Because funds for installation at NEPRC came from PREA, the primary purpose of RFID at NEPRC is to reduce inmate-on-inmate sexual assaults and to aid in the investigation of actual and alleged assaults. RFID at NEPRC is also used to confirm if an inmate is where she is supposed to be and to document the date, time, and location of fights. In addition, as with Ross Camp, NEPRC uses RFID to enhance perimeter control and complement body counts by providing an electronic “running count” of inmates.

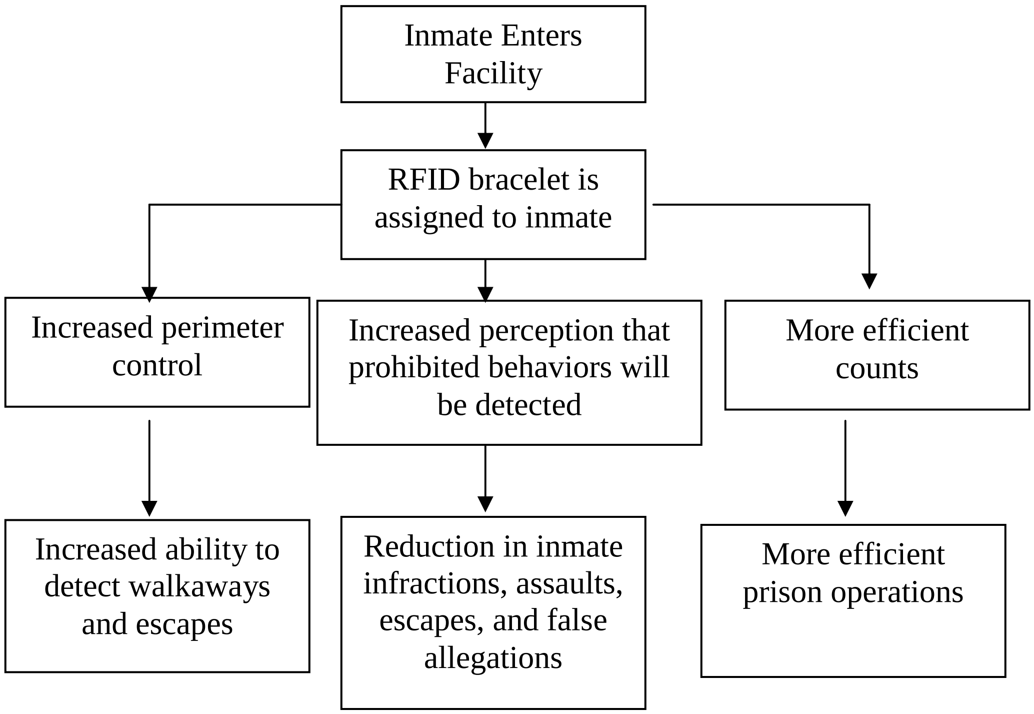
RFID became fully operational at both Ross Camp and NEPRC in August 2006.

What efficiencies or primary/secondary outcomes are expected?

RFID use in correctional facilities is designed to improve prison management, offering a more efficient means of locating inmates, confirming counts, and alerting officials to escapes. It also has the ability to aid in investigations. Overall, it holds promise for improving inmate behavior (i.e., reduced infractions and assaults).

Sketch the logic by which technology use should affect goals (see exhibit 1).

Exhibit 1. RFID Logic Model



Is the technology well suited and appropriately specified given these goals?
It is, but an evaluation would need to occur to explore exactly how the technology is implemented and used by correctional staff.

Are there operational alternatives that could be used for comparisons?

The operational alternative would be no RFID use, which in this context would mean the identification of a comparison institution that is not currently using RFID. Given the variation in ODRC's facilities in terms of size, design and, population, it would be difficult if not impossible to select such a comparison institution.

Is the site interested in being evaluated?

Both Ross Camp and NEPRC are keenly interested in being evaluated.

Is the site planning an evaluation?

Currently, ODRC has no plans for formal evaluation of RFID.

Data Sources

What data systems exist that would facilitate evaluation?

ODRC maintains an Institutional Climate Database for each facility, which documents numbers of inmate escapes; walkaways; drug finds; weapons finds; disruptive incidents; use-of-force incidents; cell extractions; Rules Infraction Board (RIB) hearings; drug test results; homicides; suicides and suicide attempts' and inmate-on-inmate physical assaults, sexual assaults, and fights. Although base rates are low for most of these measures, significant numbers of inmate fights occur at NEPRC (an average of 36 per year) and NEPRC also has a relatively high number of RIB hearings, averaging 197 each year. Both alleged and confirmed incidents of inmate sexual assaults are extremely low, averaging five and three per year, respectively.

Although ODRC maintains similar incident data for Ross Camp, incidents for Ross Camp are combined with data for Ross Correctional Institution, precluding the use of incident data to assess the impact of RFID on inmate behavior at Ross Camp.

At the facility level, data are also maintained on inmate locations, movements, and out-of-place alerts. Currently, those data are only maintained for 30 days and are then purged from the system. However, given that this information would support a process evaluation, it is likely that ODRC would agree to maintain these data longer in support of an evaluation.

What key data elements are contained in these systems?

See data systems discussion above.

Are there data to estimate unit costs of labor and capital?

Currently the cost data are embodied in the original RFPs released for Ross Camp and NEPRC, the contracts of which totaled $425,000 and $390,000 respectively. There are no maintenance costs to date, as both installations are still under warranty. Labor costs and benefits are also difficult to estimate, as RFID does not replace staff, it simply aids them in locating inmates, identifying infractions, and conducting investigations.

Are there data for possible comparison technologies or other solutions?

No. The only comparison would be business as usual before implementation of RFID.

In general, how useful are the data systems to an impact evaluation?

If the goal is to look at the global impact of RFID in prison (i.e., aggregate inmate behavior) rather than the local impact (e.g., tracking individual inmates on RFID), then the data should be suitable for impact evaluation purposes.

Is this site worthwhile?

Yes.

2. Site Visit Screening

The Intervention

Has the organization implemented a policy and/or training for the technology's use?

Training has thus far been vendor supplied, with more training offered at NEPRC than at Ross Camp. Mostly the training has been "on the job," and procedures have been adapted over time to suit the way the technology has been used as well as the problems that have been encountered with the technology. For example, officers have developed a system for calling into the central command area to clear false alarms.

Who are the users?

The users of the technology are correctional staff at all levels who play a role in managing and accounting for the whereabouts of the inmate population. In addition, the investigators at each institution use the technology to research assaults and other inmate incidents.

Who/what are the targets?

The targets are the inmates at each of the two institutions. At Ross Camp, there are 350 mixed-security male inmates. NEPRC houses 570 minimum-security female inmates.

Who/what gets excluded as a user or target?

No one is excluded—all inmates have RFID bracelets.

Have the characteristics of the user or target population changed over time?

Ross Camp as had an increase in short-term inmates over the last several years. Although the female inmate population across Ohio has increased significantly in recent years, the composition of women at NEPRC has been relatively stable.

What values/outcomes do users see/envision in the technology?

According to the correctional staff we interviewed, RFID serves as a useful management tool. Although it does not serve as a substitute for headcounts, it reinforces counts and aids tremendously in identifying where out-of-place inmates are located. This saves staff time and effort in tracking down inmates, which can be very time consuming, particularly at Ross. This could become particularly critical during inmate escapes, none of which have occurred at NEPRC or Ross Camp since the time of RFID implementation. Although escapes are rare, with RFID, correctional officers would know that an escape occurred—and which inmate escaped—within minutes.

ODRC also believes that RFID is saving time and money in investigations, and that it is particularly useful when used in tandem with closed-circuit televisions (CCTVs) at NEPRC.

Because of inmates' perceptions that they are closely monitored and their whereabouts are known at all times—perceptions that are reinforced when they are caught out of place —ODRC believes that RFID may actually prevent rules infractions, assaults, and thefts from taking place.

What are the limitations/obstacles in using the technology?

The most common problem with the RFID technology as experienced by both Ross and NEPRC staff is false alerts. False alerts can occur when an inmate is sitting on the floor and the bracelet's RFID signal is picked up by the DEU on the floor below where the inmate actually is. Signal blockage can also cause false alerts, as was the case with inmates under the metal-roofed pavilion at NEPRC (the vendor added additional DEUs to correct the problem). Signals may also be blocked if a male inmate is sleeping with his hand against the wall or if a female inmate has her ankle on the ground.[11]

The sizing of the RFID bracelets may also cause problems. The bracelets used at Ross are difficult to resize, and in both facilities if inmates gain or lose weight the bracelets will not fit. Overly tight bracelets prompt complaints from inmates and naturally require adjustment, but overly loose bracelets are perhaps more problematic, in that they issue an

[11] Due to the standard width of the RFID units, they are used on wrists for male inmates but are more suitable for use on female inmates' ankles.

"inmate missing" alert. When RFID was first implemented at both facilities, these types of alerts were frequent, but as correctional officers have become more accustomed to the technology they have decreased significantly.

Other issues with bracelets include their battery life (when batteries die, the unit reports "inmate missing"), and the cleaning and maintenance of the bracelets, which can be time consuming for correctional officers.

One final limitation to use of the technology is unique to Ross, in that correctional officials at that facility do not find the software to be user friendly, which may limit their use of the technology to its fullest capacity.

What outcomes could be assessed? Using what measures?

Improved prison management. With regard to prison management issues, it would be useful to know the extent to which RFID has increased correctional officers' efficiency and perhaps saved officers' time. For example, the time it might take to track down the location of an inmate might be better spent patrolling the grounds or conducting counts. However, given that any evaluation would likely be retrospective, it is not feasible to collect hard data on how officers spent their time before and after RFID implementation. This outcome would have to be addressed qualitatively through interviews or focus groups with correctional officers.

Improved inmate behavior. Theoretically, one would anticipate that RFID increases inmates' perceptions of the risk of being detected while committing an offense or infraction. One would also expect that any effect that RFID had on improved management overall would have a secondary effect on inmate behavior. The best means of measuring inmate behavior is through an analysis of inmate infraction data before and after RFID implementation.

Better investigations. More specific to sexual assault, RFID may have an impact on inmate reports of victimizations. Fewer false allegations may be reported if inmates learn that RFID helps refute false claims. Likewise, RFID may increase the number of inmates who report actual sexual assaults because they have more confidence in the system based on evidence supplied by RFID. Theoretically, this outcome could be measured by analyzing the number of sexual assault complaints filed before and after RFID implementation, as well as the share of those complaints that are substantiated at time one versus time two. However, given the low base rate of sexual assault allegations (5 per year on average), this would be difficult to confirm quantitatively.

Designing a Study

Are there other operational environments for which the technology is well suited?

Any type of correctional facility should be well suited for this technology.

What are the constraints in such environments?

The constraints are mostly in cost and implementation time. Large facilities and those that have extensive grounds will require the installation of many more DEUs, and it takes time to calibrate the reception area around each DEU.

Do the technology "events" permit randomly generated applications of the technology?

This technology is not conducive to random assignment, as any efficiencies gained in prison management would be lost under such a scenario.

Can comparison samples be formed? With what difficulties?
The only possible comparison sample would be a comparison institution. Due to the variations in institution size, design, and location, however, this is not an appropriate evaluation approach.

How many times would the technology be applied in one year?

The technology, once applied, remains within the institution indefinitely.

Will modest but statistically significant effect sizes be detectable given sample sizes?

The only likely evaluation method would be an interrupted time series design. The base rate of inmate infractions, however, is relatively low. However, if one examines changes in rates of both inmate fights and RIB hearings, these data should be sufficient to detect a difference between pre- and post-implementation if one exists.

How many units—if any—would have to be procured for an evaluation?

The units have already been procured. However, this technology would be even more powerful if correctional officers also were equipped with RFID devices, particularly with regard to investigating allegations of sexual assault or other charges of inmate abuse. Such a study would require the procurement of additional units (one for each correctional staff person).

What does a control/comparison group receive?

The "comparison group" would be the institution prior to RFID implementation, so it would receive nothing.

What kinds of data elements are available from existing data sources?

See data elements question above.

What specific input, process, and outcome measures would they support?

See above.

How complete are data records?

The data are in the process of being converted to a new system. However, the data that are maintained on inmate behavior are extremely rich and detailed.

Can user and/or target populations be followed over time?

Target populations may be followed over time at the institutional level; once an inmate left an RFID-equipped facility, he or she would drop out of the sample. However, RFID use should not be measured at the individual level, but rather at the institutional level whereby aggregate changes in infractions can be assessed over time.

Can the dosage of technology used be identified?

No.

Can data systems help diagnose implementation problems?

To some extent, the data system associated with the RFID software enables the generation of reports for different types of alerts and the responses of correctional officers. Those data would therefore aid in learning whether correctional officers are responding to alerts promptly and appropriately.

What threats to a sound evaluation are most likely to occur?

The greatest threat to an evaluation is a Type II error (failing to reject the null hypothesis when it is false). The relatively low base rate of inmate infractions and fights may not provide sufficient statistical power to detect a significant reduction in events from pre-RFID implementation to post-RFID implementation.

Another potential evaluation threat is that, because RFID may actually increase *detection* of infractions, records of official infractions may increase and therefore may not reflect any improvement in inmate behavior.

What changes is the site director willing to make to support the evaluation?

ODRC might be willing to consider using RFID with correctional staff, but union issues may make that difficult to sell.

3. Overall

Would you recommend that the technology be evaluated?

An impact evaluation is possible at NEPRC but not at Ross Camp. Ross' base rate of inmate infractions is already quite low, so any impact on infractions is unlikely to be detectable. Moreover, the infractions data for Ross cannot be disaggregated from that of its larger neighboring correctional institution.

What type of evaluation designs would you recommend?

The most appropriate evaluation design for identifying changes in inmate behavior as a result of RFID implementation would be a retrospective interrupted time series design employing autoregressive integrated moving average (ARIMA) modeling. Employing weekly RIB and inmate fight data, ample pre- and post-intervention data points should be available to support this approach. This quantitative evaluation should be complemented with qualitative data collected through semi-structured interviews with correctional staff and the facility investigator to learn their perceptions of the impact of RFID, as well as focus groups with inmates to learn whether RFID use promotes a safer prison environment. Given the data restrictions at Ross Camp, an impact evaluation is recommended for NEPRC only.

Even without an impact evaluation, this technology still merits a full process evaluation at both facilities. Such an evaluation would support prospective new adopters in making informed decisions about whether to invest in the technology and ways in which it can be effectively applied in a correctional setting.

Plans for Future Expansion

Director Collins has expressed an interest in implementing RFID in a second women's prison, as well as in Ross Correctional Institution, the 2,600-inmate medium-level facility located 500 yards away from the Ross Camp. Implementation at Ross Correctional Institution, although expensive, would enable ODRC to test out RFID's capabilities at identifying gang members and triggering alerts when rival gang members are in close proximity to one another. Director Collins is also contemplating equipping correctional officers with RFID, which would further support investigation efforts and has the potential to protect correctional officers from false allegations of misconduct. In a perfect world, Director Collins would like to see RFID implemented during the construction of a new prison, placing DEUs and CCTVs strategically throughout the facility to enhance surveillance and monitoring of inmates. Although no plans are underway for new prison construction in Ohio, Director Collins believes this would be the most cost-effective approach to RFID implementation and operation.

References

Adams, Chris. (2004). “Property Crime Reduction Through the Use of Electronic Tagging Systems.” *The Chipping of Goods Initiative: A Final Review*. Home Office Police Scientific Development Branch.

Beck, Adrian. (2006.) “Shrinkage and Radio Frequency Identification (RFID): Prospects, Problems and Practicalities.” In *The Handbook of Security* Martin Gill, ed. (pp. 462–482). New York: Palgrave Macmillan.

Gluckman, George. (2005.) “Profiling the Future: New Technologies are Changing the World of Profiling.” *Law Enforcement Technology* 32 (8): 64–71.

HighJump Software. (2004). “True Cost of Radio Frequency Identification (RFID).” 3M. Retrieved February 28, 2006, from http://www.highjumpsoftware.com/promos/rfid-cost-report.asp

Home Office. (2006). “Chipping of Goods.” Retrieved February 28, 2006, from http://www.chippingofgoods.org.uk/

IDTechEx. (2006). “The RFID Knowledgebase.” Retrieved October 24, 2006, from http://rfid.idtechex.com/knowledgebase/en/

Justice Technology Information Center. (2005). “Technology Primer: Radio Frequency Identification.” *TechBeat*. National Law Enforcement and Corrections Technology Center. National Institute of Justice.

Reza, J.D. (2004). "Do You Know Where Your Offenders Are?" Law Enforcement Technology 31(6): 118–123.

Swedberg, Clare. (2006). “Police Tested RFID-Enabled Badges.” *RFID Journal*, March 10. Retrieved August 17, 2006, from http://www.rfidjournal.com/article/articleprint/2191/-1/1.

Swedberg, Clare. (2005). “LA County Jail to Track Inmates.” *RFID Journal*, May 16. Retrieved August 17, 2006, from http://www.rfidjournal.com/article/articleprint/1601/-1/1,

Walker, J. Douglas. (2005). “Radio Frequency Identification: Why the Justice System Should Tune In.” *Trends*. National Center for State Courts.

Attachment A: Correctional Agencies Using RFID

Facility Name	Location	Type of Facility	Number of RFID Units	Implementation Year	Implementation Status	Targets	Vendor
ODRC Ross Correctional Center	Chillicothe, OH	Minimum/ medium/closed facility	**350**	2004	Fully implemented	Male and female inmates	TSI Prism
Logan Correctional Center	Lincoln, IL	Medium security facility	**2,000**	2003	Fully implemented	Male inmates	TSI Prism
W.J. Maxey Training School for Boys	Whitmore Lake, MI	Youth detention/ rehabilitation center	**250**	2002	Fully implemented	Male juvenile inmates	TSI Prism
Marion Treatment Center	Marion, VA	Mental health facility	**N/A**	2006 (expected by yearend)	Not fully implemented	Male inmates	TSI Prism
Minnesota Correctional Facility-Faribault	Faribault, MN	Medium-security, level-three facility	**150**	2002	Fully implemented	Male inmates	ElmoTech
Pitchess Detention Center North	Castaic, CA	Low and medium security facility	**300**	2004	Fully implemented	Male inmates	ElmoTech
St. Peter Regional Treatment Center	St. Peter, MN	Mental health facility	**100**	2005	Fully implemented	Male and female inmates	ElmoTech
ODRC Northeast Pre-Release Center	Cleveland, OH	Minimum/ medium security compound	**704**	2006	Fully implemented	Female inmates	ElmoTech
Southern Nevada Correctional Center	Jean, NV	Medium level facility	**200**	2006	Fully implemented	Juvenile male and female inmates and staff	ElmoTech

Facility Name	Location	Type of Facility	Number of RFID Units	Implementation Year	Implementation Status	Targets	Vendor
California State Prison - Corcoran State Hospital	Corcoran, CA	Minimum, medium, closed, and high security facility including protective housing unit.	**200**	1997	Fully implemented	Staff	TSI Prism
Minnesota Correctional Facility-Stillwater	Stillwater, MN	Minimum security and closed facility.	**100**	2006 (expected by yearend)	Not Implemented	Male inmates	TSI Prism
St. Joseph Community Supervision Center	St. Joseph, MO	Minimum security halfway house	**50**	2006 (expected by yearend)	Not Implemented	Male and female inmates	TSI Prism
Farmington Supervisory Center	Farmington, MO	Minimum security halfway house	**50**	2006 (expected by yearend)	Not Implemented	Male and female inmates	TSI Prism

Secure Continuous Remote Alcohol Monitoring (SCRAM) Technology Evaluability Assessment

Staff Contact: Brian Barton
Director
Marion County Community Corrections
708–341–9361

NIJ Guidance

The National Institute of Justice (NIJ) recommends, with qualifications, an evaluation of Secure Continuous Remote Alcohol Monitoring (SCRAM) in the site assessed below (or other appropriate community corrections settings). NIJ is not convinced that an appropriate control group could be constructed because of the obstacles to random assignment and data access necessary for propensity scoring. NIJ would consider an application that overcame these obstacles.

Applicants who propose to evaluate this technology (or other SCRAM implementations) are encouraged to consider the outcome variables (including detection and deterrence of violations, compliance with the conditions of community release, and cost savings from jail diversion) and obstacles (including small numbers and unavailable or incomparable control groups) identified below. NIJ encourages applicants to identify sites where randomization is possible or where matched comparison groups can be readily constructed.

Applicants may depart from this guidance by providing an appropriate rationale.

Project Summary: Secure Continuous Remote Alcohol Monitoring is a relatively new technology designed to continuously monitor pretrial clients and offenders under community supervision for alcohol consumption and issue alerts to community corrections officers when alcohol has been consumed. We selected Marion County, Indiana, as the focal point of our evaluability assessment of SCRAM. Marion County Community Corrections (MCCC) is the agency with the largest number of clients using SCRAM, with approximately 280 SCRAM users at any given time. Marion County has been using this technology since 2003, with judges employing SCRAM as a sanction or condition of pretrial release for those who have been charged with or sentenced for driving under the influence (DUI) or domestic violence offenses. Marion County officials invested in SCRAM in an effort to relieve jail overcrowding and because SCRAM enables clients to remain in the community, drive a motor vehicle, and maintain employment during the course of their sentence or pretrial release period.

Scope of Evaluation: A rigorous outcome evaluation of SCRAM would be possible if Marion County agreed to random assignment to SCRAM or an alternative sanction. To date, one judge has expressed an interest in learning more about what participation in an evaluation involving random assignment would entail. Another possible evaluation design would be a retrospective evaluation employing propensity scores to identify a comparison group.

Summary of Evaluability Assessment Activity: To understand the prevalence of SCRAM and to assess the feasibility of evaluating SCRAM technology, Urban Institute (UI) staff began with a review of the literature and a Web-based search to identify agencies currently using the technology. In addition, UI had several phone and e-mail communications with Alcohol Monitoring Services (AMS), the manufacturer and sole proprietor of SCRAM technology, to identify potential agencies. Informal interviews with technology experts at the National Law Enforcement and Corrections Technology Centers (NLECTC) were also conducted. The results of the literature review, telephone interviews, and conference calls led to the conclusion that SCRAM monitoring of offenders in the community is a relatively new application in the criminal justice arena, but is quickly being adopted by community corrections agencies across the country.

UI's initial screening identified five mature applications of SCRAM technology. These were found at Marion County Community Corrections (Indiana), Michigan Department of Corrections, the City and County of Denver (Colorado), Maricopa County Adult Probation (Arizona), and Eastern Missouri Alternative Sentencing Services. Michigan Department of Corrections served as the beta testing site for SCRAM in 2002. However, MCCC, with approximately 280 persons being monitored using SCRAM, has one of the largest caseloads of any agency using SCRAM, and therefore was selected for this evaluability assessment.

1. Background

Describe the technology. What is the background/history of this technology?

Secure Continuous Remote Alcohol Monitoring is an automated alcohol-monitoring device that uses transdermal testing to measure the amount of alcohol in person's body, known as transdermal alcohol content (TAC). When alcohol is consumed, ethanol migrates through the skin and is excreted through perspiration. SCRAM measures TAC levels by taking a sample of one's perspiration. Traditional methods of measuring alcohol consumption commonly employ a portable or stationary device, such as a Breathalyzer, which measures blood alcohol content (BAC). BAC relies upon fuel cell technology and provides a one-time view of a person's alcohol consumption. SCRAM, on the other hand, allows for continuous testing regardless of the location of the person under supervision, which increases the sampling detection. Moreover, whereas the BAC burnoff rate is relatively high, dissipating within a few short hours after a last drink, TAC levels remain high for a much longer duration, increasing the possibility of detection of alcohol consumption. The SCRAM device also measures body temperature as a means of determining whether the bracelet has been removed or tampered with so as to block perspiration from being read by the device.

The SCRAM system has three components: the SCRAM bracelet, the SCRAM modem, and SCRAMnet. The SCRAM bracelet is an 8-ounce device that is attached to a client's ankle and is worn around the clock. It is made up of two parts: (1) a sensor pack, which tests vapors through the skin; and (2) a data-storage component, which collects, stores, and transfers data regarding alcohol consumption as well as tamper detection and systems control. The modem is connected to a landline and at a prescheduled time each day, the bracelet will transmit data through the modem using secure radio frequency. The modem stores alcohol readings, tamper alerts, body temperature, and diagnostic data from the bracelet; it then transmits data from the SCRAM bracelet, via the Internet, to SCRAMnet. The modem also downloads monitoring and reporting schedules from SCRAMnet to the supervising agency. SCRAMnet is a Web-based application in which offender data is collected, analyzed, and stored. Agencies employing SCRAM technology can use SCRAMnet to control testing, synchronization, and reporting schedules of monitored subjects.

Maturity

SCRAM is manufactured by Alcohol Monitoring Services. AMS has trademarked SCRAM and is the sole proprietor of this technology. SCRAM is a relatively new product: the first patent for SCRAM was filed in 1991, and in 1993 the first operational SCRAM prototype was completed and a patent was granted. In 2002, the first 100 preproduction SCRAM units were introduced and beta testing of SCRAM began. In 2003 the first commercially available SCRAM units were introduced to the field.

Prevalence in the field

According to AMS, SCRAM is currently available in 35 States and is used by more than 600 courts and agencies throughout the Nation (see attachments A and B). Use by individual agencies varies greatly: some have few as 1 or 2 clients; others monitor more than 200 persons with SCRAM.

What do we already know about technologies like these?

SCRAM is the first and only commercially available secure continuous remote alcohol-monitoring device. Other remote noncontinuous technologies are available, but as agencies become aware of SCRAM, they are more apt to choose it over competitors because it is more tamperproof and provides more accurate measures of alcohol use at roughly the same cost as other alcohol-monitoring devices.

What could an evaluation of this technology add to current knowledge?

The only formal evaluation of SCRAM our preliminary literature review identified is one based on 2.5years of data in Alaska. The study found that the system, which was implemented in a rural area via Alaska’s satellite telecommunications network, operated reliably and was successfully used on supervised offenders in areas with extreme weather conditions.[12] The evaluation, however, was restricted to an assessment of the

[12]McKelvie, Alan R. 2006. “An Implementation of Remote Alcohol Monitoring In Alaska.” Justice Center, University of Alaska at Anchorage.

technology's performance and did not examine its impact on correctional supervision or offender behavior. The majority of knowledge regarding SCRAM is limited to reports by AMS, beta testing of SCRAM at the Michigan Department of Corrections, and various media reports. However, there is no empirical literature available on the impact of SCRAM, and its recent and widespread use beckons an evaluation in order to inform agencies and the larger criminal justice arena of its potential benefits.

Which audience(s) would benefit from this evaluation?

Judges, corrections officials, probation, parole, and community supervision staff would all greatly benefit from an evaluation.

What could they do with the findings?

Agencies that have already invested in SCRAM would naturally be interested in knowing whether it has an impact on detection of alcohol consumption among their clients, as well as the inclination of SCRAM clients to engage in alcohol use. Communities contemplating investing in SCRAM would also be interested in these findings. For example, if a SCRAM evaluation demonstrates that it is effective in both detecting alcohol consumption as well as possibly discouraging it, more community correction agencies would invest in it. This would equip judges with a new intermediate sanction appropriate for DUI and domestic violence offenders, which could free up jail space and save money. In addition, corrections, probation, parole, and community supervision officers could increase their ability to monitor offenders and do so more effectively.

At what stage of adoption/implementation is the technology in the targeted site?

SCRAM is fully implemented in the five sites we identified and has been operational in MCCC since 2003—around the time SCRAM was first introduced.

What efficiencies or primary/secondary outcomes are expected?

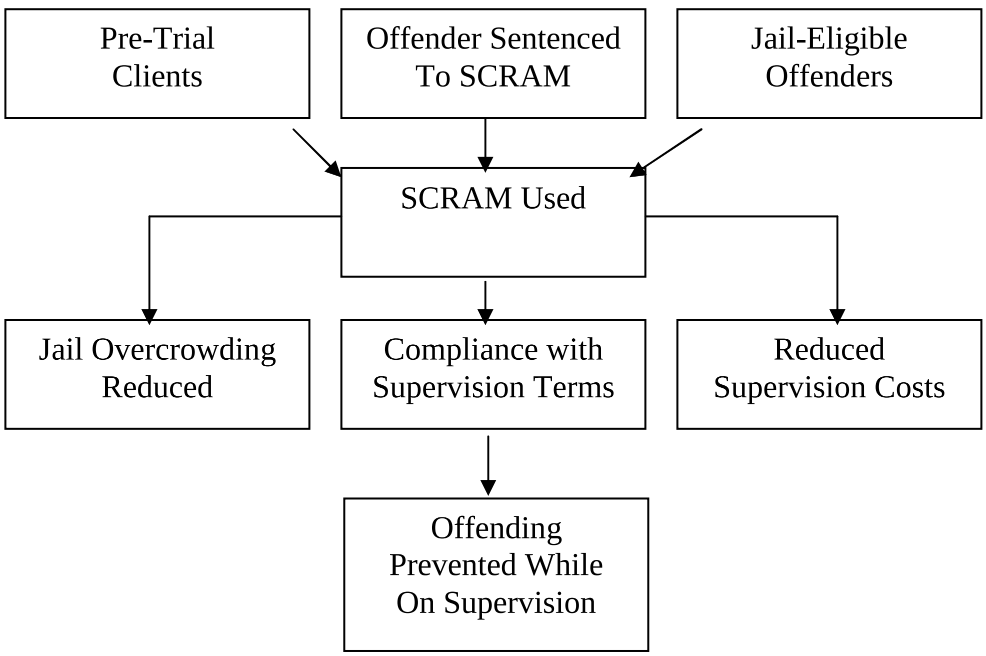
The primary outcome of SCRAM is its potential to increase the detection of prohibited alcohol use among SCRAM clients. Secondary outcomes include reduced alcohol consumption as well as increased compliance with other conditions of supervision. Depending on how it is used in sentencing decisions, SCRAM also has the potential to reduce jail overcrowding by diverting would-be inmates to a sanction of SCRAM in the community.

The basic outcome logic of this technology is that offenders with histories of alcohol abuse can be supervised under sentence or pretrial release in the community, where they can maintain their jobs and day-to-day activities, including driving, through the continuous monitoring of their alcohol use. The primary outcome suggested is that jail overcrowding can be reduced, or at least minimized. In addition, supervision costs using this technology are much lower than those of incarceration. Theoretically, this technology may also reduce technical and criminal offenses during the period of supervision and reduce longer-term recidivism.

The goals of the use of this technology are to provide a safe and secure alternative to incarceration. The objectives are to: 1) reduce jail overcrowding; 2) decrease supervision costs; 3) increase detection of alcohol use while under supervision; and 4) reduce reoffending by deterring alcohol consumption, which serves as a precipitator to DUI and domestic violence offenses.

Sketch the logic by which technology use should affect goals (see exhibit 1)

Exhibit 1. SCRAM Logic Model



Is the technology well suited and appropriately specified given these goals?

It is logical to purport that SCRAM has the potential to increase the detection of alcohol consumption and to reduce actual alcohol consumption among SCRAM clients. The extent to which SCRAM successfully reduces the jail population depends in large part on how clients assigned to SCRAM would have been supervised were SCRAM not available. It could be, for example, that SCRAM use simply provides an extra measure of supervision for those who would have received a community supervision sentence anyway (thereby widening the net of community supervision rather than decreasing the jail population).

Are there operational alternatives that could be used for comparisons?

The operational alternative to SCRAM would be other forms of supervision that are typically used on clients who are charged or sentenced with similar offenses. These alternatives include home detention with electronic monitoring through the use of radio frequency technology, global positioning systems (GPS), and various forms of conditional supervision.

Is the site interested in being evaluated?

All of the agencies UI contacted are interested in being evaluated. MCCC would greatly welcome an evaluation.

Is the site planning an evaluation?

None of the sites contacted indicated that they have planned an evaluation.

Data Sources

What data systems exist that would facilitate evaluation?

The possible data sources for evaluation purposes are threefold: (1) case-level data on clients on SCRAM and other forms of home detention supervision (i.e., electronic monitoring and GPS monitoring); (2) systemwide court data on all persons who are sentenced to jail, community supervision or pretrial release; and (3) AMS data on elevated TAC levels and tampering incidents.

What key data elements are contained in these systems?

Case-Level Data on SCRAM and Other Home Detention Clients

Marion County maintains extensive electronic data on clients on SCRAM, as well as those on GPS and electronic monitoring, including demographic information, current offense, criminal history, risk level, drug testing dates and results (if applicable), violations of terms of supervision, and employment status. This database, however, is case based and does not allow for the creation of reports that aggregate across the entire client base. Nonetheless, the data exist and could be extracted manually to track outcomes for treatment and control groups.

Systemwide Court Data

Electronic data on all persons charged with criminal offenses are maintained by the Marion County Circuit Court Clerk from 1998 to the present. These data include name, age, sex, race, initial charge, case summary and chronology, disposition, and sentence. Pretrial and sentenced persons can be tracked through the system using a unique ID number associated with the individual.

AMS Data

AMS collects data that are downloaded daily from the bracelets regarding TAC and temperature readings, elevated TAC alerts, and signs of tampering. An AMS representative indicated that AMS has maintained all of the downloaded data since 2003.

Are there data to estimate unit costs of labor and capital?

AMS charges community corrections agencies $1,500 for purchase of one SCRAM bracelet and modem set. However, Marion County opted to lease the units at a daily rate of $1.70 per unit over a 3-year period. Additional fees of $5 per day are charged to cover AMS' monitoring costs. Marion County in turn charges its SCRAM clients $12 per day in supervision fees, which, given an average 50-percent collection rate, roughly covers the costs of SCRAM

Are there data for possible comparison technologies or other solutions?

Marion County is not employing any other alcohol detection system at this time. However, the county maintains data on those under home detention with electronic monitoring and GPS supervision. Either of these community sanctions could serve as a comparison technology.

In general, how useful are the data systems to an impact evaluation?

While the data systems do not allow for easy extraction of information, the data are available and would support a rigorous impact evaluation.

2. Checkpoint

Is a site visit worthwhile?

Of the five sites identified, MCCC is the most viable option for a site visit.

3. Site Visit Screening

The Intervention

Has the organization implemented a policy and/or training for the technology's use?

Yes. AMS provides training for staff who use SCRAM technology and there is a certain amount of on-the-job training from MCCC staff who are familiar with the SCRAM system.

Who are the users?

The primary SCRAM users are judges, who use SCRAM as a community supervision sanction, and corrections officers, who receive daily reports from AMS and respond to alerts about members of their caseloads who test positive for alcohol use or have tampered with the SCRAM unit.

Who/what are the targets?

Currently SCRAM is used primarily for DUI and domestic violence cases, along with a handful of drug cases.

Who/what gets excluded as a user or target?

The technology is aimed at offenders for whom alcohol use influences or precipitates their criminal behavior or puts others at risk. Persons who do not have histories of alcohol abuse or misuse are excluded.

Have the characteristics of the user or target population changed over time?

MCCC initially used SCRAM on DUI cases. As use of the technology became known, judges began to use SCRAM for any offender for whom alcohol served as a gateway to criminal behavior or violence.

What values/outcomes do users see/envision in the technology?

Ideally MCCC would like persons on SCRAM to attain permanent abstinence from alcohol use. However, more realistic outcomes envisioned by MCCC include reduced alcohol consumption; increased compliance with treatment and other forms of supervision; and decreased recidivism. On a macro level, MCCC envisions that SCRAM use will result in decreased jail overcrowding.

What are the limitations/obstacles in using the technology?

Originally a major limitation to using SCRAM was its cost. AMS had initially only given agencies the option to purchase the units. Now that AMS is leasing the units, MCCC has the ability to offset the leasing costs through the collection of supervision fees from SCRAM clients. Another limitation noted by MCCC was that clients must download the information from the bracelet using a landline, which many clients do not have. Therefore some clients must make special arrangements to access a landline so that data from the bracelet can be downloaded. Equipment failure was also noted as a limitation. MCCC notes that the current equipment is much better than the equipment they first used. Monitoring individuals with equipment failures, such as batteries running down and other malfunctions, can also be labor intensive.

What outcomes could be assessed? Using what measures?

Although it is the primary stated objective of MCCC's investment in SCRAM, a reduction in jail overcrowding is not a feasible outcome measure for evaluation purposes. The implementation of SCRAM has been incremental from 2003 to the present, making an interrupted time-series design inappropriate for evaluation purposes because it would be too difficult to identify intervention points.

Alcohol detection rates of those on SCRAM compared to those on other forms of supervision may be difficult to assess as well. Since MCCC employs no alternative alcohol detection system, SCRAM by definition would be more likely to detect alcohol use than any nontechnological means (e.g., self-reported alcohol use by clients). However, alcohol-related offenses, other offending behavior, compliance with other conditions of supervision, and jail admission can all be assessed.

Designing a Study

Are there other operational environments for which the technology is well suited?

The most suitable environment for this technology is a community setting.

Do the technology "events" permit randomly generated applications of the technology?

Yes, provided judges agree to participate in a study involving random assignment.

How many times would the technology be applied in 1 year?

The number of new SCRAM clients each year is approximately 186. Pretrial clients are on SCRAM an average of 120 days. Sentenced offenders are on SCRAM for an average of 180 days.

Will modest but statistically significant effect sizes be detectable given sample sizes?

The statistical power will depend on the sample size (which depends on the number of participating judges and their SCRAM-eligible caseloads) as well as the expected effect size of the intervention (which is likely to be small to moderate). Without more specific information on the number of SCRAM-eligible clients who could be assigned to treatment or control groups, statistical power cannot be fully assessed at this time.

How many units, if any, would have to be procured for an evaluation?

MCCC currently has 350 units in-house. As of October 11, 2006, MCCC monitors 287 offenders using SCRAM. Because we are unable to assess SCRAM's prospects for expansion at this time, it is difficult to know whether additional units would need to be procured for evaluation purposes.

What does a control/comparison group receive?

A control group would have similar characteristics to SCRAM clients (i.e., histories of DWI or alcohol-precipitated violence) but would receive some other form of community supervision or conditions of pretrial release, such as home detention with electronic monitoring, GPS, or conditional release (e.g., curfews, license suspension). Any evaluation design would require a researcher to determine the exact composition of the control group (e.g., a mix of home detention, GPS, and conditional release) or whether it would be more appropriate to compare the SCRAM treatment to multiple comparison groups (e.g., one for home detention, one for GPS, and a third for conditional release). These decisions will rest to a large extent on sample sizes.

What kinds of data elements are available from existing data sources?

See data source discussion above.

What specific input, process, and outcome measures would they support?

Input measures include number and type of clients put on SCRAM, AMS data on alcohol use and tampering by SCRAM clients, and duration of SCRAM monitoring.

Process measures are currently not well-documented by MCCC, but could be collected through the use of a data collection instrument requiring supervision officers to document the ways in which they respond to tampering and alcohol use alerts.

Outcome measures include AMS data on alcohol use, and MCCC and County Circuit Court data on violations of conditions of release, new arrests, new convictions, jail admissions, and potentially employment information.

How complete are data records?

See data source discussion above.

Can user and/or target populations be followed over time?

Persons on SCRAM can be followed over time during the duration that they are required to be on SCRAM. After they are released from SCRAM supervision the only way to follow their involvement with the criminal justice system would be check their names against court, police, and corrections records.

Can the dosage of technology used be identified?

The only feasible dosage measure would be duration of time on SCRAM monitoring. As referenced above, the average time a client is on SCRAM ranges from 120 days (for pretrial clients) to 180 days (for sentenced offenders).

Can data systems help diagnose implementation problems?

AMS collects data on equipment failures and triggers. Although MCCC does not currently collect data on individual corrections officers' responses to SCRAM alerts, data collection systems could be developed for such a purpose.

What threats to a sound evaluation are most likely to occur?

The greatest threats to evaluation are: (1) nonrandom assignment of participants to treatment and control groups due to judges deviating from the random assignment protocol; and (2) lack of statistical power to detect an impact if one exists, due to a small effect size and/or a small sample size. With regard to sample size, much will depend on the number of judges who agree to participate in random assignment and the size of their SCRAM-eligible caseloads.

A secondary threat to evaluation concerns the time it may take to recruit study participants and track them over time to assess outcomes. If too few judges are willing to participate in a randomized controlled trial (RCT), the flow of eligible candidates for assignment to treatment and control groups may be slow. If it takes more than a year to recruit a sufficient N of study participants, and outcomes are tracked for the sample for at least 6 months (the average time clients are on SCRAM), this could amount to an evaluation that spans 3 years or more, which could be costly. This is a legitimate threat to an RCT design, as MCCC assigns only 186 clients to SCRAM each year: Almost half of all judges would need to participate in an RCT in order to obtain treatment and control groups of 50 persons each within a year's time (and that assumes that all eligible study candidates will agree to participate).

What changes is the site director willing to make to support the evaluation?

The major issue impacting an evaluation is the ability to identify a control group. MCCC is willing to approach judges to help identify a way to do so. It is difficult to discern at this time whether enough judges could be recruited to support such an approach.

4. Overall

Would you recommend that the technology be evaluated? Why or why not?

Provided that an RCT could be employed, this technology should be evaluated. Another approach entailing a weaker evaluation design would be to retrospectively compare SCRAM users to a control group identified through the use of propensity scores. This would require a researcher to gain access to MCCC's client database as well as the County Circuit Court Clerk's database to extract and analyze data. MCCC's database is rich, but is not designed in way that supports easy data extractions. The County Circuit Court Clerk's database is searchable online at the case level, but we do not know at this time whether aggregate data can be exported from that system.

Even without random assignment, this technology still merits a full process evaluation so that prospective new adopters can make informed decisions about whether to invest in the technology.

What type of evaluation designs would you recommend?

The most rigorous design would involve random assignment of persons at the pretrial or sentencing stage to either treatment (SCRAM) or control (home detention) groups. Following both groups over time will enable the collection of data on whether the groups differ in terms of violations of conditions of supervision and measures of recidivism (arrests, convictions, and returns to prison). An alternative design would be a retrospective evaluation comparing outcomes of those monitored by SCRAM versus those assigned to other forms of community supervision.

Attachment A: Interviewed Agencies Currently Using SCRAM

Location	Implementation Year	Number of Units	Criminal Justice Application	Targets	Interest in Evaluation	Outcomes*	Data Systems
Marion County Community Corrections (IN)	2003 started with 20 units	**287** on (350 in-house)	Community supervision.	287 offenders with driving under the influence (DUI) and domestic violence (DV) cases; some drug cases.	High level of commitment and interest.	Reduced alcohol consumption. Attain permanent abstinence (although not likely). Decrease in substance abuse. Increase in compliance with substance abuse treatment.	Have access to data on violations and sentencing information; AMS provides reports regarding violation/triggers.
Michigan Department of Corrections	2003 started with 30 units	**100** on (260 in-house)	Probation and parole supervision.	100 parolees and probationers convicted of a felony; primarily Operating Under the Influence (OUI) offenses.	Very interested in an evaluation; Would like to be able to show that it is more effective than other methods (Sobrieter).	Increased reporting of violations.	They have a case management system that compiles general offender data; AMS provides reports regarding violation/triggers.
City and County of Denver (CO)	2003	**90**	Pretrial supervision	90 offenders with DUI/DV or any alcohol-related offense.	Very interested in advancing the knowledge and education of such technology.	Increase in victim safety. Decrease in substance abuse. Increase in compliance.	AMS provides reports regarding violation/triggers. Should be able to get access to other data.
Maricopa County Adult Probation (AZ)	2003 started with 10 units	**65**	Probation; Some lower courts are using it.	65 probationers mostly from DUI courts as needed from DV or drug court.	Very interested in evaluation and strong commitment in technology from department.	Increased compliance with orders. Decreased alcohol consumption. Increase in sobriety. Increase in successful periods of being monitored.	Automated database. AMS provides reports regarding violation/triggers.
Eastern Missouri Alternative Sentencing Services	2004	**111**	Probation; Condition of bond; Attorney referral for pretrial alcohol-related offenses.	111 offenders with alcohol-related offenses.	Possibly	Increased abstinence. Increase in compliance.	AMS provides all the data they use.

* As defined by site, may not be quantifiable.

Attachment B. Agencies and Counties Currently Using SCRAM by State

Alabama
Mobile County Community Corrections

Alaska
Aleutians East Borough
Anchorage Borough
Bristol Bay Borough
City and Borough of Juneau
City and Borough of Sitka
City and Borough of Yakutat
Denali Borough
Fairbanks North Star Borough
Haines Borough
Kenai Peninsula Borough
Ketchikan Gateway Borough
Kodiak Island Borough
Lake and Peninsula Borough
Matanuska-Susitna Borough
North Slope Borough
Northwest Arctic Borough

Arizona
Gila County
Maricopa County Community Corrections
Maricopa County DUI Court
Maricopa County DV Probation
Pinal County
Yavapai County

Arkansas
Sebastian County

California
Contra Costa County
Kern County
Los Angeles County
Orange County
Sacramento County
San Francisco City and County
Santa Barbara County
Santa Clara County
Solano County
Yuba County

Colorado
Adams County
Arapahoe County
Baca County
Bent County
Boulder County
Broomfield City and County
Chaffee County
Cheyenne County
Crowley County
Custer County
Denver City and County
Douglas County
El Paso County
Elbert County
Fremont County
Garfield County
Gilpin County
Jackson County
Jefferson County
Kiowa County
Kit Carson County
Larimer County
Las Animas County
Lincoln County
Logan County
Mesa County
Morgan County
Otero County
Park County
Phillips County
Pitkin County
Prowers County
Pueblo County
Rio Blanco County
Sedgwick County
Teller County
Washington County
Weld County
Yuma County

Delaware
Kent County
New Castle County
Sussex County

Florida
Alachua County
Baker County
Bradford County
Broward County
Charlotte County
Collier County
De Soto County
Escambia County Community Corrections
Gilchrist County
Glades County
Hardee County
Hendry County
Indian River County
Jackson County
Lee County
Leon County
Levy County
Manatee County
Martin County
Miami–Dade County
Okaloosa County
Okeechobee County
Orange County
Osceola County
Palm Beach County
Pinellas County
Santa Rosa County
Sarasota County
St Lucie County
Union County
Volusia County Drug Court
Walton County

Georgia
Chatham County DUI Court
Clarke County DUI Court
Cobb County Drug Court
Hall County DUI Court

Idaho
Ada County
Benewah County
Bonner County
Boundary County
Kootenai County

Shoshone County

Illinois
DuPage County

Indiana
Hancock County
Hendricks County Probation
Marion County Community Correction
Boone County
Delaware County
Fayette
Hamilton County
Hendricks County Superior Court Probation
Henry County
Johnson County
Madison County
Morgan County
Putnam County
Shelby County
Tippecanoe County
Vigo County
Bartholomew County
Blackford County
Brown County
Clay County
Dearborn County
Decatur County
Elkhart County
Franklin County
Grant County
Huntington County
Jackson County
Kosciusko County
La Porte County
Lagrange County
Lake County
Monroe County
Porter County
Ripley County
St Joseph County
Steuben County
Wells County

Iowa
Dallas County
Jasper County
Marion County
Polk County
Story County
Warren County

Louisiana
Acadia Parish
Calcasieu Parish
East Baton Rouge Parish
Iberia Parish
Iberville Parish
Jefferson Davis Parish
Jefferson Parish
Lafayette Consolidated Government
Livingston Parish
St. Martin Parish
Terrebonne Parish
West Baton Rouge Parish

Maryland
Anne Arundel County
Baltimore City County
Howard County
Prince Georges County
Wicomico County

Michigan
3rd Circuit Court
4A District Court
5th District Court
6th Circuit Court
16th Circuit Court
17th District Court
18th Circuit Court
18th District Court
19th District Court
21st Circuit Court
21st District Court
23rd District Court
27th District Court
28th District Court
31st District Court
32A District Court

34th District Court
35th District Court
37th Circuit Court
37th District Court
38th District Court
39th District Court
40th District Court
41A District Court
41B District Court
42nd District Court
43rd District Court
44th Circuit Court
44th District Court
46th Circuit Trial Court
46th District Court
47th District Court
48th District Court
52nd District Court
55th District Court
56A District Court
58th District Court
59th District Court
61st District Court
64A District Court
70th District Court
72nd District Court
74th District Court
76th District Court
88th District Court
89th District Court
Benzie County Probation and Parole
Berrien County Probation and Parole
Clare County Sheriff
Eaton County Probation and Parole
Grosse Pointe Municipal Court
Kalamazoo County Probation and Parole
Kent County Probation and Parole
Lake County Probation and Parole
Livingston County Probation and Parole
Macomb County Probation and Parole
Manistee County Probation and Parole
Mason County Probation and Parole
Michigan Department of Corrections
Muskegon County Probation and Parole
Oakland County Probation and Parole
Oceana County Probation and Parole

Ottawa County Probation and Parole
Van Buren County Probation and Parole
Washtenaw County Probation

Minnesota

Aitkin County
Anoka County
Beltrami County
Blue Earth Community Corrections
Brown County
Carver County
Chippewa County
Chisago County
Crow Wing County
Dakota County
Dodge County
Douglas County
Fillmore County
Hennepin County Community Corrections
Isanti County Community Corrections
Jackson County
Le Sueur County
Martin County
McLeod County
Meeker County
Morrison County
Murray County
Nicollet County
Olmsted County
Ramsey County Community Corrections
Renville County
Roseau County
Scott County
Sherburne County
Sibley County
Stearns County Community Corrections
Steele County
Washington County
Watonwan County
Wright County

Mississippi

Alcorn County
Attala County
Benton County
Bolivar County

Calhoun County
Carroll County
Chickasaw County
Choctaw County
Clay County
Coahoma County
De Soto County
Grenada County
Hinds County
Holmes County
Humphreys County
Issaquena County
Itawamba County
Kemper County
Lafayette County
Lauderdale County
Leake County
Lee County
Leflore County
Lowndes County
Madison County
Marshall County
Monroe County
Montgomery County
Neshoba County
Newton County
Noxubee County
Oktibbeha County
Panola County
Pontotoc County
Prentiss County
Quitman County
Rankin County
Scott County
Sharkey County
Sunflower County
Tallahatchie County
Tate County
Tippah County
Tishomingo County
Tunica County
Union County
Warren County
Washington County
Webster County
Winston County

Yalobusha County
Yazoo County

Missouri
Barton County
Bates County
Benton County
Boone County
Buchanan County
Butler County
Caldwell County
Camden County
Camden County
Cape Girardeau County
Carroll County
Cass County
Cedar County
Chariton County
Clay County
Clinton County
Cole County
Cooper County
Crawford County
Dade County
Dallas County
Dunklin County
Franklin County
Greene County
Henry County
Hickory County
Howard County
Jackson County
Jasper County
Jefferson County Courts
Johnson County
Laclede County
Laclede County
Lafayette County
Lawrence County
Lincoln County
Macon County
Miller County
Mississippi County
Missouri Probation and Parole
Moniteau County
Montgomery County

Morgan County
New Madrid County
Newton County
Perry County
Pettis County
Phelps County
Platte County
Polk County
Pulaski County
Randolph County
Ray County
Saline County
Scott County
St Charles Associates and Circuit Court
St. Charles Drug Court
St Clair
St Francois
St Louis County
St Louis City
St. Louis County Circuit Court
St. Louis County Justice Services
Texas County
Vernon County
Warren County

Montana
Carbon County
Musselshell County
Stillwater County
Yellowstone County

Nebraska
Arthur County
Chase County
Dawson County
Douglas County
Dundy County
Frontier County
Furnas County
Gosper County
Hayes County
Hitchcock County
Hooker County
Keith County
Lancaster County
Logan County

McPherson County
Perkins County
Platte County
Red Willow County
Sarpy County
Thomas County

Nevada
Clark County
Washoe County

New Mexico
San Juan County

New York
Orange County
Rockland County
Suffolk County

Ohio
Akron Municipal Court
Ashland County
Carroll County
Chardon Municipal Court
Columbiana County
Crawford County
Cuyahoga County Municipal Court
Delaware County
Fairfield County
Franklin County Municipal Court
Fulton County
Guernsey County
Harrison County
Henry County
Hocking County
Holmes County
Jefferson County
Knox County
Licking County
Lucas County
Mahoning County
Marion County
Medina County
Miami County
Morgan County
Morrow County

Muskingum County
Oregon Municipal Court
Perry County
Pickaway County
Portage County
Richland County
Rocky River Municipal Court
Ross County
Seneca County Common Pleas Court
Stark County
Summit County Common Pleas Court
Summit County Juvenile Court
Tiffin Municipal Court
Tuscarawas County
Vinton County
Wood County

Oklahoma
Cleveland County
Creek County
Delaware County
Garvin County
Grant County
Kay County
Logan County
McClain County
Oklahoma County
Osage County
Ottawa County
Pawnee County
Payne County
Rogers County
Tulsa County

Oregon
Malheur County

Pennsylvania
Allegheny County
Blair County
Butler County
Cambria County
Centre County
Chester County
Franklin County
Lackawanna County Drug Court

Lycoming County
Mercer County
Sullivan County
Susquehanna County
Venango County
Washington County
Wayne County
Wyoming County

South Carolina
Pending SCRAM program—discussion underway

South Dakota
Entire State covered by service providers or State program

Texas
Andrews County
Angelina County
Bexar County
Bowie County
Brazoria County
Brazos County
Burnet County
Cameron County
Cass County
Collin County District Court
Dallas County District Court
Denton County District Court
El Paso County
Ellis County District Court
Fort Bend County
Galveston County
Harris County
Henderson County
Hidalgo County
Houston County
Jim Wells County
Johnson County
Kaufman County District Court
Kleberg County
Midland County
Nacogdoches County
Nolan County
Palo Pinto County District Court
Parker County
Rockwall County District Court

San Patricio County
Tarrant County District Court
Taylor County
Travis County
Willacy County
Williamson County

Utah

Department of County District Court
Murray Justice Court
Salt Lake City County
Taylorville Justice Court
Uintah County District Court

Vermont

Addison County
Bennington County
Caledonia County
Chittenden County
Essex County
Franklin County
Grand Isle County
Lamoille County
Orange County
Orleans County
Rutland County
Washington County
Windham County
Windsor County

Washington

Adams County
Benton County
Columbia County
Douglas County
Ferry County
Franklin County
Garfield County
Grant County
Lincoln County
Okanogan County
Pend Oreille County
Pierce County
Skagit County
Spokane County
Walla Walla County

Whitman County
Yakima County

Wisconsin

Dane County
Dodge County
Fond du Lac County
Grant County
Jefferson County
Kenosha County
La Crosse County
Milwaukee County
Racine County
Rock County
Sheboygan County
St Croix County
Walworth County
Washington County
Waukesha County
Winnebago County

Wyoming

Albany County
Big Horn County
Campbell County
Carbon County
Converse County
Crook County
Fremont County
Goshen County
Hot Springs County
Johnson County
Laramie County
Lincoln County
Natrona County
Niobrara County
Park County
Platte County
Sheridan County
Sublette County
Sweetwater County
Teton County
Uinta County
Washakie County
Weston County
Laramie County

Lincoln County
Natrona County
Niobrara County
Park County
Platte County
Sheridan County
Sublette County
Sweetwater County
Teton County
Uinta County
Washakie County
Weston County

Evaluability Assessment of Trace Detection Technology

Staff Contact: Lt. Sean Stewart
Pima County Sheriff's Department
Corrections Bureau
520-547-8384
sean.stewart@sheriff.pima.gov

NIJ Guidance

The National Institute of Justice (NIJ) recommends that applicants consider evaluating this technology along cost effectiveness lines, whether the application is sited in Pima County or some other system. While Pima County uses their technology in a mail room, other systems might use the technology to screen visitor possessions or suspicious parcels found in facilities. Applicants are encouraged to consider such outcome variables as numbers of items detected, persons prosecuted, and inmates sanctioned. On the cost side, evaluations should consider all possible costs including startup, training, and maintenance. Applicants may also want to consider that an effective detection technology may drive traffickers of contraband to explore other avenues of penetration, so that measurement of outcomes needs to account for possible displacement effects.

Technology Summary: In recent years, several systems have become available for the detection of trace amounts of drugs and explosives. This report discusses the range of technologies and focuses specifically on the use of ion mobility spectrometry for detecting illegal drugs. In this process, a swab is taken from the area in question. Inside a scanner, the sample is heated until particles of the substance are vaporized. These vaporized ions are then analyzed and compared against a library of potential narcotics and other substances, and any positive matches are identified.

Scope of Evaluation: Illegal drugs can be transported into jails and prisons by visitors and incoming inmates, by corrupt guards and other officials, and through the mail. Pima County officials felt that corruption was absent among Pima County guards and that visitors could not transport drugs because the jail does not permit personal contact. Jail officials used trace detection technology to periodically check inmates returning from community release and to routinely examine incoming mail. An evaluation would examine (1) the effectiveness of trace detection technology at intercepting illegal drugs coming in through the mail, (2) the consequence that successful detection has for drug use in the jail, and (3) the effect that reducing drug use has on day-to-day jail operations. These three questions are progressively more difficult to answer.

Summary of Evaluability Assessment Activity: During the evaluability assessment, Abt Associates, Inc., conducted a thorough search of manufacturer reports and field uses of trace detection technologies in the United States. We also interviewed individuals from Sandia Laboratories, the Specialized Crimes and Narcotics Task Force of Kingsville, TX, and staff from the Pima County Adult Detention Center. We completed a site visit at the Pima County Detention Center in order to view the technology in practice.

Finding: We find trace detection technologies to be evaluable. Three designs are suggested which attempt to quantify the end benefit of this technology rather than simply the process outcomes.

1. Initial Screening

Background

Drug and explosive detection technology is frequently divided into two categories: bulk detection and trace detection systems. While bulk detection is used to identify large masses of visible substances, trace detection is used to identify the minuscule, sometimes invisible, residue of substances. Traditionally, law enforcement officers have used specially trained canines to aid in the detection of contraband (in its bulk and trace forms). However, the increased versatility of electronic detection technology has allowed law enforcement to test large amounts of potentially contaminated materials (for example, the thousands of letters that pass through a prison mail room each day) with an efficiency that may not be feasible with the use of a canine. (NLETC, 2003)

Trace detection technology is capable of identifying chemicals—be they illegal drugs or explosives—in the form of gas vapors or particulate matter.

- *Gas vapors*: The detection of vapors involves the use of vacuum technology with a sensitivity to gas-phase molecules. All solids and liquids emit varying amounts of gas vapors, which are dependent on a variety of environmental conditions. At particular temperatures, the amount of vapor that a substance emits can allow trace detection technology to identify the substance at hand. (Thiesan, 2004) One of the advantages of gas vapor technology is that such equipment need not touch a potentially contaminated service to detect the presence of an illicit substance.

- *Particulate matter*: The detection of particulate matter involves the use of sampling-pad technology that swipes contaminated surfaces to collect microscopic solids. A few micrograms in weight is sufficient for identifying an illegal drug. Because trace detection of particles involves touching a potentially contaminated surface, technicians who use this equipment must engage in special procedures to avoid cross-contaminating surfaces. (Thiesan, 2004) It is important to note that the amount of particulate matter present on a given item can be related very loosely to the amount of a contraband substance present in a given environment. Cross-contamination is very common, and can occur through second-hand contact, even if no contraband material is present. (Parmeter, 2000)

Trace detection technology comes in seven principal forms, three of which are commonly used to detect narcotics: ion mobility spectrometry, chemical reagents, and mass spectrometry.[1] *Ion mobility spectrometry* can operate in vapor mode (by collecting air samples) or in particle mode (by swiping a surface), and can detect the presence of narcotics through the collection of as few as 100 picograms. Ion mobility spectrometry identifies chemical substances by measuring the drift speed of ions. *Chemical reagents* detect particle matter through the use of aerosol sprays and swipes, which are calibrated to change color in the presence of different substances. Chemical reagents, for example, have been developed to identify the presence of cocaine or marijuana on surfaces, and can identify material with a concentration of as few as 5 picograms. (B. Butler, 2002 and Thiesan, 2004) *Mass spectrometry* identifies trace substances by swiping contaminated surfaces to identify a material's molecular weight and fragmentation patterns. Thiesan, et al. (2004) describe it primarily as a "mass

[1] The other four technologies—chemiluminescence, thermo-redox, surface acoustic wave, and ultraviolet fluorescence technology—are used principally in the detection of explosives (Thiesan, 2004, pp. 33-40).

filtering technique," in which a substance is collected, ionized, and passed through a filter that identifies its charge-to-mass ratio. Like ion mobility spectrometry, mass spectrometry has a minimum detection level of 100 picograms for any given substance. (B. Butler, 2002)

What is the background/history of this technology?

Maturity (i.e., lab prototype? Field roll-out? Multiple generations/manufacturers?)
Trace detection technology is widely available for law enforcement purposes, with several private manufacturers supplying well-tested, increasingly advanced equipment.

Time in the field?
No report or article that we obtained gave an estimated date at which trace detection technology first became available to correctional officers in the field. Within the past decade, however, NIJ and the Federal Bureau of Prisons began to focus on providing trace detection equipment to help prisons curtail the amount of drugs that are smuggled into jails through visitors or mail services. (Wright, 2001)

Prevalence in the field? (Is site a first/early adopter?)
Trace detection technology itself has been widely available for many years. In recent years, its sensitivity and accuracy has increased substantially, and the price of many models has become more affordable, making it a more attractive option for correctional facilities. We have no definitive estimate of the number of prison facilities that employ trace detection technology, but with more manufacturers marketing to the correctional "market," we believe that the number is increasing. NIJ's evaluation of contraband detection technology in its Mailroom Scenario Evaluation (2000-2002), as well as its three-day experiment in Pima County's jail mailroom (2002), may have increased the visibility of such technology, while showcasing the versatility that it can offer to correctional facilities.

What do we already know about technologies like these?
At this time, most of the literature on the use of trace detection technology in correctional facilities discusses the results of studies that were funded by NIJ and its National Law Enforcement and Corrections Technology Center. The Mailroom Scenario Evaluation tested the success rates of different trace detection technologies at identifying the presence of multiple narcotics, and a three-day "experiment" at Pima County compared two different ion mobility spectrometers. Butler, R.F. (2002) described the mailroom evaluation as a "limited" scenario evaluation, which simulated a prison mailroom environment by modeling an evaluation center after the mailroom center at the U.S. Penitentiary in Leavenworth, Kansas. This evaluation tested various ion mobility spectrometers and a chemical reagent spray to measure their ability to identify a number of drugs, including marijuana, cocaine, heroin, methamphetamine, ecstasy, and LSD. (R.F. Butler, 2002) One of the principal findings of this evaluation was that the minimum detection levels taken from the literature of vendors is not necessarily accurate, and that minimum detection levels are dependent on a variety of factors, including drug type and environmental conditions. (Wright, 2001) The three-day Pima County experiment was, to a large degree, designed to better understand the needs of mailroom staff with regard to the ease-of-use of particular equipment. Several officers at Pima County, for example, appreciated the portability of one of the ion mobility spectrometers, which allowed staff to test letters after they had been opened by inmates, or swab in particular areas of the facility after visitors have left. (Falcon, 2005)

The results of these two evaluations highlighted a number of key factors that correction facilities need to consider when deciding which type of trace detection technology to purchase. These factors include, among other things: purchasing costs, maintenance cost, screening speed (which involves amount of time it takes a system to identify trace matter), the sensitivity of the system to different types of drugs, system portability, ease of use (which includes training requirements), and safety issues. (Parmeter, 2000) Other issues mentioned in both evaluations involved the varying rates of false-positives given by almost all detection equipment, much of which was attributed to cross-contamination. Trace detection technology, for example, cannot distinguish between mail that contains smuggled drugs, and mail that was simply touched by hands that had previously handled a certain drug. (Falcon, 2005) Mail contamination can also occur when tainted mail touches clean mail and thereby contaminates it. (USDOJ, 2004) While this contamination will continue to remain an unavoidable problem for correctional facilities that employ trace detection technology, there is nevertheless training that staff can undertake to limit the extent to which they themselves become the unwitting agents of cross-contamination through poor handling procedures.

What could an evaluation of this technology add to current knowledge?

Illegal drugs enter jails and prisons through various vectors, including the mail. An evaluation would not test the reliability and specificity of the trace detection technology at detecting illegal drugs, because the performance of the technology is not in question. Rather, the question is how the technology works in an operational setting: How much illegal drugs are removed from entering a jail through the mail and at what cost? What proportion of illegal drugs used in jails is transported through the mail, and hence, what is the benefit of using trace detection technology?

Interdicting a drug shipment is not necessarily the final step when applying the technology. Identifying a drug shipment is sometimes a trigger for investigating an offender’s conduct by using telephone taps and by investigating the letter’s source. This can lead to prosecutions of the sender, the recipient, or both.

Which audience(s) would benefit from this evaluation?

- Law enforcement agencies and correctional facilities
- Database companies and developers who sell such technology

What could they do with the findings?

The use of trace detection technology to inspect incoming mail is only one step toward reducing drug use in jails and prisons. Presuming that a jail is able to control other vectors into the jail (visitors, incoming inmates, corrupt guards), then the jail could use the results from the research to judge whether the additional reduction in imported drugs would be worth the staffing and equipment costs of implementing trace detection tools in the mailroom.

At what stage of adoption/implementation is the technology?

Pima County has been using the IonScan system (from Smiths Detection) since July of 2005. It has been in continuous use in the mailroom since that time with little change in protocol. The system is used at least every other day to test suspicious mail items.

What efficiencies or primary/secondary outcomes are expected?

Sketch the logic by which technology use should affect goals (see Exhibit 1)

Exhibit 1: Logic Model

Method: Trace Detection
Uses ion mobility scanner to identify trace amounts of drugs in articles of mail sent to the facility.

Proximal Outcomes
- Removes contaminated items from circulation and prevents drugs from entering the facility by the most common method.
- Reduces or eliminates the number of inmates who are high.
- Provides prosecutable evidence for the crime of introducing contraband into a jail.

Distal Outcomes
- Reduces frequency and magnitude of inmate-inmate violence.
- Reduces inmate-guard violence.
- Facilitates operations by increasing inmate's trust in the ability of staff to maintain order.

Are there operational alternatives that could be used for comparison?

No system in place at Pima County Adult Detention Center provides the same accuracy or speed as the IonScan system. However, the Center has used several alternatives for the detection of incoming contraband materials. These include:

- Visual inspection: Mailroom staff inspect each item of personal mail by hand. In some cases, an ultra-violet light may be used to highlight stains or non-ink lettering. This method usually cannot identify the specific drug or concentration.
- X-ray: All mail entering the facility is X-rayed (the only "search" permitted in legal mail). X-rays will identify bulk materials in the mail and have some ability to distinguish between organic materials (less dense) and metallic materials (more dense). However, the X-ray machine is not very specific and cannot identify trace amounts.

- Drug dogs: The Pima County Detention Center has on occasion used drug dogs to confirm the presence of drugs in suspect mail. However, the mail must be separated widely so that the dog can specifically identify the item in question. This can be time-consuming and does not eliminate the possibility of cross-contamination from other items of mail.
- Lab results: Each positive result on the IonScan is confirmed at the county laboratory through chemical assays of the material. This process is as accurate as the IonScan machine, but requires additional training and is more time consuming. It also cannot be accomplished without damaging the letter in question.

However, the trace detection technology augments rather than replaces extant technology. The identification of drugs through trace detection technology is necessarily "value added" because it can be applied to a large number of parcels, which is impractical for laboratory analysis.

Is the site interested in being evaluated?

Pima County Adult Detention Center would be willing to participate in an evaluation. They have previously cooperated in NIJ studies such as the Adult Drug Abuse Monitoring program.

Is the site planning an evaluation?

No evaluation is currently planned.

Data Sources

What data systems exist that would facilitate evaluation?

The IonScan system can produce a printed receipt with the time, the name of the drug, and the amount detected. Currently Pima County staff prints a receipt for every positive test above the usable threshold (approx. 3-4 on a scale of 20). Copies of these receipts are kept after the letter itself is sent to the county lab. In addition to paper receipts, the Pima County staff keeps a logbook of all of the positives identified. In the first 14 months of use, 81 items were found to have usable quantities of drugs. This is approximately one to two per week. Currently the logbook is in handwritten form and no electronic version is available. In addition to the Pima County Detention Center materials, each sample is sent to the county lab for confirmation. The lab results are then reported with detailed information on the type and amount of drug identified.

What key data elements are contained in these systems?

The logbook contains the same information as the printed receipt from the IonScan machine. Mailroom staff writes the date of discovery, the item in question, and the type and amount of drug found.

Are there data to estimate unit costs of labor and capital?

The data on labor and capital costs have not yet been specifically collected, but could be identified with little additional effort. There are several costs associated with the use of the IonScan system:

- IonScan unit: ~$60,000

- Training: two hour training for mailroom staff (unknown cost)
- Maintenance: There are several “consumables” associated with the IonScan system. Swabs can be used up to 25 times (if only negative results) before they must be discarded. The carbon filter and tubes must be replaced periodically. If contaminated by a very high dose of drugs, the system may need service from the vendor. Pima County staff suggested annual maintenance costs were $500-700.

Currently the system is used approximately daily to check mail that is *suspicious* based on a visual inspection of the opened letter. Drugs are smuggled by dissolving them into a fluid that is then used to write the letter. Gel pens are especially useful. The bulk amount of fluid may cause damage to the paper surface, and the bulk amount of ink or crayon or other material may raise suspicion.

The amount of mail deemed suspicious and hence subject to trace the number of mailroom personnel limits detection analysis. Adding staff would add costs.

Furthermore, the identification of contraband may trigger an extended investigation of the sender and the recipient. Although this might be limited to interviews, it sometimes involves wiretaps of phone calls and criminal investigation and prosecution. This adds costs to the process.

Offsetting some of these costs, the system is portable and has been used on occasion to check inmates as they return from visits (including employment) outside the jail. Because of the amount of ambient drugs in areas where prisoners travel when released temporarily, testing positive is not necessarily indicative of recent drug use, but it can precipitate questioning and investigation.

Are there data for possible comparison technologies or other solutions?

Currently each positive “hit” on Pima County’s IonScan device is confirmed by lab analysis of the suspect item. Lab analysis is operationally more difficult because it is offsite, requires significant training, and can only test a small portion of the item in question. (Usually a small piece of a letter is removed and prepared for lab analysis.) The two systems are not independent because Pima County mail staff tells county lab technicians where on the letter the drug is to be found. However, lab reports on the items can be compared to the IonScan paper receipts to confirm the amount and type of drug tested. Visual inspection of mail (the method used prior to introduction of the IonScan device) did identify some instances of incoming drugs, but no records are available from that period of time.

In general, how useful are the data systems to an impact evaluation?

The printed receipt from the IonScan device is a useful record for an impact evaluation, however it does not show the number of items that were tested but shown to be negative (mailroom staff estimate that about 20% of items that were tested were positive), nor the proportion of mail that passes through the mailroom untested. Protocols could be put in place to record both positive and negative readings over the course of an evaluation.

Nevertheless, the log is a definitive count of those that tested positive, and we presume that few such suspicious items were detected prior to the use of IonScan. Thus the log provides a definitive measure of drugs removed from prison circulation. Furthermore, the log provides a time series, so an evaluator could derive a measure of the deterrent effectiveness of using the IonScan. Given a longer time series, an evaluator might correlate the receipt of drug-infused mail with disciplinary actions or prosecutions of mail senders or recipients, providing an assessment of whether those collateral actions are a further deterrent to mailing drugs to prisoners.

2. Site Visit Screening

The Intervention

Has the organization implemented a policy or training for the technology's use?

Two types of knowledge are required to use this system effectively. In Pima County, only suspect mail is tested, so the first skill is the ability to recognize an item of mail that is likely to be contaminated. Mailroom staff lacks explicit training in this, but with experience they acquire an ability to recognize paper that may have been tampered with. The second is training in the use of the detection system itself. This training is conducted by Smiths Detection and takes approximately two hours. During that time, the trainee learns not only how to run the IonScan device, but how to service it when maintenance is required. In Pima County, outside help from the vendor was required only once since the beginning of the trial in July of 2005.

Who are the users?

Users are correctional officers and mailroom staff who need to identify trace amounts of drugs in mail or on an inmate's person.

Who/what are the targets?

In Pima County, trace detection is primarily used in the mailroom. It is used to detect drugs in incoming personal mail before it reaches the recipient. Currently the system is only used to test items that are identified by the staff as suspicious. The system is portable and can be transported to other locations for testing the hands of visitors or inmates returning from a work-release program.

Who/what gets excluded as a user or target?

The law protects legal mail (mail from an attorney to his client) from search prior to delivery to the recipient. In the event that legal mail looks suspicious, jail authorities will contact the mailing attorney to assure that the attorney actually sent the letter, because a prisoner's associates sometime counterfeit the attorney's letterhead. Attorneys may give permission to open a suspicious letter. Otherwise, legal mail cannot be opened and tested with the IonScan device.

Mail that does not catch the attention of mailroom staff may pass through to an inmate without testing. Although checking every item of mail would be too time consuming for the available staff, a random check of some proportion (~20%) could identify other instances of contraband entering the jail. An evaluation could show the frequency at which mail not deemed suspicious actually contained contraband.

Have the characteristics of the user or target population changed over time?

The characteristics of the user and target population have not changed significantly over time. Mailroom staff have become more adept at identifying the area of a suspicious letter most likely to contain drug compounds. Also, the number of positive "hits" in proportion to total mail has decreased since the system was put in place. This is attributed to increased awareness among inmates and inmates' contacts of the likelihood of detection and a decrease in attempts to send drugs through the mailroom.

What values/outcomes do users see/envision in the technology?

The primary goal of the use of this technology in Pima County is to prevent the introduction of drugs into the facility. By detecting trace amounts of drugs, mailroom staff can remove the contaminated item from circulation and prevent inmates from using the drugs. The use of drugs in a correctional facility causes problems on many levels. Inmates who are high are more likely to be violent towards one another and to correctional officers. In addition, a trade in drugs promotes debts and antagonisms between inmates that can lead to gang-related violence. If drugs are entering a jail or prison, inmates lose trust in the ability of correctional officers to maintain order, and may resort to making homemade weapons for self-defense from other inmates. Use of these weapons greatly increases the number of severe injuries inflicted. Pima County staff estimate that 90% of contraband materials enter the prison through the mailroom. Trace detection technology can eliminate or greatly reduce this source of contraband materials and thus promote a less violent population.

Trace detection can also be used to test individuals, and has been used at the nearby Mission prison to test inmates returning from a work-release program. In this case, trace detection can be used as an enforcement technology, allowing prison staff to identify inmates who have broken work-release rules by using drugs. In one testing, 10 inmates admitted using drugs when confronted by the evidence of trace amounts of drugs on their hands. The evidence can be used in some cases to prosecute individuals, either who break parole by consuming drugs or who anticipate receipt of drugs through the mailroom. Pima County has successfully prosecuted five cases of attempting to bring contraband into a facility over the course of the 18 months since the use of trace detection technology began.

What are the limitations/obstacles to using the technology?

The small-scale trace detection machines now available are all based on swabbing an item and testing the residue on the swab. (Large non-swab technologies such as the phone-booth sized "sniffers" used in airports are available, but are significantly more expensive.) On a contaminated item there may be only a single area that contains contraband material. This means that a user must know where to test in order to obtain a positive result. This becomes an important issue when trace detection is used in other arenas. For example, border control officers must know where to take samples in a car that is being inspected. Each swab covers only a few inches of material, so there is a possibility that significant quantities of drugs may pass undetected if they are in an area that is not tested. Drug dogs, in contrast, can find drugs at a distance. However, the benefits of specificity and ease of use make trace detection machines very useful despite the limitation of swabs.

Although not a limitation, another concern in the use of trace detection technology is that it is extremely sensitive. Pima County officials worked with the county lab staff to agree upon a threshold of concern. Trace amounts less than that threshold were considered not high enough to represent a usable amount and may be the result of cross-contamination. Because mail can easily be cross-contaminated, staff must be trained to test the inside of mail rather than just the envelope in order to avoid positive results that may not indicate deliberate attempts to send contraband. Any study should note not just the detection of a drug, but the amount detected to avoid the inclusion of trace amounts so small that they represent only cross-contamination.

What outcomes could be assessed? Using what measures?

The primary outcome that can be assessed is the number and type of drugs prevented from entering a facility through the mailroom. Specific process measures include:

- Proportion of mail tested.
- Proportion of tested mail that tests positive (above threshold).
- Type and amount of drug detected for each positive item.
- Type and amount of drug confirmed through county lab results.
- Reduction in attempts to send contraband through the mailroom as evidenced by a reduction in the number of items that test positive since the introduction of trace detection technology in 2005.
- Disciplinary actions and prosecutions initiated because of mailroom detection.

These are process measures. Outcome measures involve the amount of illegal drug use in the jail. Collecting such outcome measures would require separate data collection. Potential outcome measures are:

- Drug use in the jail.
- Sequela of drug use in the jail.

Absent random urine testing (which does not appear to happen in the Pima County jail), there seems to be no way of monitoring drug use over time. An evaluator might acquire a snapshot of drug use within the jail using a survey. It would require clever survey technology to acquire valid responses. Furthermore, the utility of the survey would be limited to identifying the prevalence of illegal substance use and the method of acquisition.

Designing a Study

Are there other operational environments for which the technology is well-suited?

What are the constraints in such environments?
The technology is applicable to long-term and medium-term detention. Pima County officials noted that mail inspection is unnecessary for short-term detention, because short-term detainees do not receive mail. Trace detection has also been used by border-control officers to test automobiles entering the country. However, particulate matter trace detectors are limited by the officers' ability to recognize where to take a sample. This technology is only useful in environments where the area to be searched is small.

Do the technology "events" permit randomly generated applications of the technology?

If not, can comparison samples be formed? With what difficulties?

A true randomized experiment would be difficult to design. An evaluator could select a random sample of jails and then assign some to the treatment condition and some to the control condition, but this design seems impractical. For one, given differences across jails, the sample of jails would have to be large to assure adequate statistical power. For another, trace detection technology (and staff to administer the technology) is expensive. Thus, an evaluator might ask: How has trace detection technology affected drug use in the Pima County jail or in some other jail?

There appears to be no way to induce randomly generated events. If perpetrators (e.g., those who mail drugs into jails) were deterred by mail inspection, then a random pulse of inspection would cause a deterrent effect that would contaminate any random period of abatement.

Moreover, the first research question is whether or not the introduction of trace detection technology results in a higher discovery rate of contraband than would have been true in the absence of trace detection technology but in the presence of routine mail inspection (opening the mail, performing visual inspection, and using x-ray). Prior to the introduction of trace detection technology, mailroom inspectors sometimes sent mail to a laboratory for testing. We presume (but we did not ask) that the earlier laboratory results are available, so that an evaluator could determine the frequency at which contraband was detected by laboratory analysis prior to the introduction of trace detection methodology. This frequency could be compared with the frequency at which contraband was confirmed through the combination of trace detection technology and laboratory confirmation after the introduction of trace detection technology. An inference about the effectiveness of trace detection technology at identifying contraband could be determined from inspecting the time series.

A remaining question is the rate at which the trace detection technology misses mailed contraband because of the relatively high threshold for "suspicion" necessitated by the number of mailroom staff. A stratified random sample of non-suspicious mail could provide an estimate of what is missed by the current application of the technology. This is a useful evaluation question because the jail could increase mailroom staff if that would be worthwhile, and research findings could provide evidence useful for justifying a larger budget for the mailroom.

How many times would the technology be applied in one year?

The technology is applied throughout the year on an almost daily basis. Given greater staffing resources, an evaluation could involve using the technology significantly more frequently with minimal additional costs in terms of "consumable" resources.

Will modest but statistically significant effect sizes be detectable given sample sizes?

Sample size should not be an issue. The power of detecting an effect from the time series should be high unless there is a great deal of variance from period to period in the frequency of positive laboratory tests during the period predating the use of trace detection technology. Before the introduction of trace detection techniques, we suspect that the number of positive tests approaches zero on a monthly basis. Thereafter, the number approaches about 10-15 on a monthly basis. This should provide ample statistical power.

When sampling the non-suspicious mail and testing it for embedded drugs, the sample need not be large. An evaluator is not testing a hypothesis about the frequency of embedded drugs in suspicious and non-suspicious mail. The evaluator is simply attempting to estimate the frequency of embedded drugs in the non-suspicious mail. If this were a simple random sample, the standard error for the estimate would be $\sqrt{P(1-P)/N}$, where P is the proportion of drug-positive letters in the non-suspicious mail and N is the sample size. An evaluator who tests 25 mailed items per day for 48 weeks would have a sample of 6000, which seems more than adequate to estimate the proportion of non-suspicious letters that contain contraband.

How many units—if any—would have to be procured for an evaluation?

None unless this evaluation was to be done in a jail that does not currently have the technology. Manpower, not the machinery, appears to be the constraint. We presume that Pima County would allow an evaluation team to use the Pima County equipment to test non-suspicious mail provided that the evaluator provides the manpower and pays for the incremental costs.

What does a control/comparison group receive?

Based on the time series, the comparison group would receive visual inspection, x-ray scan, and occasional laboratory testing. The treated group would receive visual inspection, x-ray scan, use of IonScan, and subsequent laboratory testing.

Based on the random selection of non-suspicious mail, the treatment group (e.g., the non-suspicious mail) would receive visual inspection, x-ray scan, and IonScan; the comparison group would not receive the IonScan.

What kinds of data elements are available from existing data sources?

The Pima County Jail mailroom maintains a log of contaminated mail that was intercepted since the program's inception. In addition to this log, they keep copies of a "receipt" printed for every positive result from the IonScan machine

What specific input, process, and outcome measures would they support?

Extant data, when incorporated into the time-series design, can help identify patterns in the detection of contraband embedded in inmate mail, but these data are insufficient to support the entire evaluation. Currently, the data include only information such as the time and date of the test, the amount and type of drug found, and lab confirmation on the drug found. Additional information on legal outcomes such as prosecution is only available for a very few cases. Final outcomes, such as reduction in drug use, are not monitored.

To elaborate, the detection of embedded contraband is one step in a process. Once the contraband is detected, jail authorities may talk with the inmate or they may take investigative steps that can culminate in a prosecution. An evaluator should carefully trace the steps taken when contraband is detected in the mail and the results from internal and external investigations.

Additionally, while removing contraband from the mail is important, that act itself does not provide that the jail has made a major impact on within-jail drug use. We understand that the jail does not perform routine random urine testing, so there is not extant data regarding drug use. It seems unlikely that an investigator could perform random urine testing in a jail setting, and furthermore urine testing would not explain all that would be of interest to an evaluation. An inmate survey that asks inmates to report the prevalence of drug use in the jail and to assess how drugs move into the prison could prove more valuable. This would be a difficult survey to design. It might be part of a larger health survey that would also ask about illicit drug use, and it might ask respondents to report on general drug use in the jail population. Despite the difficulty of asking about drug use, a useful survey might only need to provide an estimate of how drugs move into the jail. If jail officials are correct that 80% of drugs move into the jail through the mail, then this fact alone can tell an evaluator the importance of intercepting drugs in the mailroom.

An evaluator might propose to conduct a time series of the rate at which contraband is discovered in the mail. The time series would run for a period before the introduction of trace detection technology and for a period after the introduction of trace detection technology. If there were a deterrent effect, we would expect the frequency of detecting drugs to peak with the introduction of the technology (perhaps following a break-in period) and then to moderate to a fairly steady state. We understand that the jail has experienced a few significant events relating to discipline/prosecution of inmate traffickers. An evaluator should identify those events and see if they have an effect on reducing the importation of illegal drugs.

How complete are data records? (Attain samples if possible.)

We requested but did not receive data records.

Can user and/or target populations be followed over time?

We did not discuss this during the site visit. It may be worthwhile to track patterns in the use of mail to import drugs into the jail. Presumably the same inmate has little or no chance to repeat his or her behavior, so there would be little reason to track a target population over time. But there may be patterns such that inmates from particular places, or inmates that are associated with particular gangs, have higher smuggling rates than do other inmates.

Can the dosage of technology used be identified?

Yes, the threshold dosage is that currently used in Pima County. The effect of any other dosage could be inferred from the sample of non-suspicious mail.

Can data systems help diagnose implementation problems?

Yes, the random sample of non-suspicious mail can help diagnose implementation problems in the form of non-optimal search rules or in the form of understaffing.

What threats to a sound evaluation are most likely to occur?

Because the analysis would be based on a time series, there is the prospect that the time series would not adequately control for external factors. This would be mitigated by the sharp interruption that the use of trace detection technology would introduce into the time series.

We are uncertain about inmates' willingness to report on drug use in custody. However, questions about the general vector of drugs into the jail would seem to be less threatening than would questions about substance use per se.

What changes is the site director willing to make to support the evaluation?

We did not discuss the proposed structure of this evaluation with jail administrators. However, the proposed design does not impose heavy costs of jail administrators.

3. Overall

Would you recommend that the technology be evaluated? Why or why not?

We recommend that the technology be evaluated, but we do not see value in simply answering the question about whether or not the use of trace detection technology helps identify drug-embedded mail. The utility of preventing such mail from entering the jail depends on whether interdiction appreciably reduces the volume of drug use in the jail. This cannot be inferred from interception rates without confirmation that the mail is the principal vector of drugs into the jail. A survey seems like the best way to get at this information.

Furthermore, the question remains: What is the advantage of reducing drug use in the jail? Jail administrators have an explanation, which we sketched above in the logic model. Can an evaluator confirm that the presence of drugs has this deleterious effect? We doubt that this can be demonstrated empirically, but testimonial evidence would be valuable.

What type of evaluation design would you recommend?

See above. We recommend an interrupted time-series analysis of the number of drug-embedded letters intercepted in the jail mail. We recommend experimental testing of a stratified random sample of non-suspicious mail to estimate the percentage of contraband intercepted by the current mail inspection program and how that percentage could be increased by expanding the use of trace detection technology.

We recommend a process analysis to understand how the trace detection works in practice and especially to understand how jail administrators use the results of mailroom inspections beyond the obvious step of interdicting the drug-embedded item of mail. We recommend that special events of public discipline and prosecution be introduced into the time-series analysis to test for deterrence.

We recommend a survey of jail inmates to understand how drugs enter into the jail. If an appreciable proportion of drugs enter the jail through other vectors, then the effectiveness of mail intercepts is necessarily diminished.

References

Butler, B. 2002. “Keeping Drugs Out of Prisons and Jails: How Can Technology Help Address the Problem?,” Sabre Systems, Inc., p. 23.

Butler, R.F. 2002. Mailroom Scenario Evaluation, final report prepared for the National Institute of Justice, pp. 10-12. Available at http://www.ncjrs.gov/pdffiles1/nij/grants/199048.pdf.

Falcon, W. 2005. “Special Technologies for Law Enforcement and Corrections,” *NIJ Journal, 252,* pp. 23-4.

National Law Enforcement and Corrections Technology Center (NLETC). 2003. “You Don’t Have Mail,” *TechBeat*, Summer 2003.

Parmeter, J.E., D.W. Murray, and D.W. Hannum. 2000. “Guide for the Selection of Drug Detectors for Law Enforcement Applications,” Sandia National Laboratories, prepared for NIJ, Office of Science and Technology (NIJ Guide 601-00).

Thiesan, L., D. Hannum, D.W. Murray, and J.E. Parmeter. 2004. “Survey of Commercially Available Explosives Detection Technologies and Equipment 2004, “ Sandia National Laboratories, report for the National Law Enforcement and Correction Technology Center, a program of NIJ.

U.S. Department of Justice. 2004. “Drug Detection in Prison Mailrooms,” *In Short: Toward Criminal Justice Solutions,* November 2004, Office of Justice Programs, Partnerships for Safer Communities.

Wright, S. and R.F. Butler. 2001. “Technology Takes on Drug Smugglers: Can Drug Detection Technology Stop Drugs from Entering Prisons?,” *Corrections Today, 63 (4).*