ISSUE NO. 271

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

NIJ

National Institute of Justice **JOURNAL**

Predicting Recidivism Risk: New Tool in Philadelphia Shows Great Promise

- Ballistic Body Armor: A Closer Look at the Follow-Up Inspection and Testing Program
- Sex Offenders Monitored by GPS Found to Commit Fewer Crimes
- Ending Modern-Day Slavery: Using Research to Inform U.S. Anti-Human Trafficking Efforts
- The Pitfalls of Prediction

Also in this issue

The Grant Progress Assessment Program: Looking Back on Success and Moving Forward

Kristen M. Zgoba Wins the Peter P. Lejins Research Award

U.S. Department of Justice
Office of Justice Programs

810 Seventh St. N.W.
Washington, DC 20531

Eric H. Holder, Jr.
Attorney General

Mary Lou Leary
Acting Assistant Attorney General

Greg Ridgeway
Acting Director, National Institute of Justice

This and other publications and products of the National Institute of Justice can be found at:

National Institute of Justice
http://www.nij.gov

Office of Justice Programs
Innovation • Partnerships • Safer Neighborhoods
http://www.ojp.usdoj.gov

NCJ 240695

NIJ

BUILDING
KNOWLEDGE TO
MEET THE CHALLENGE OF
CRIME AND JUSTICE

National Institute of Justice

Greg Ridgeway
Acting Director, National Institute of Justice

The *NIJ Journal* is published by the National Institute of Justice to announce the Institute's policy-relevant research results and initiatives. The Attorney General has determined that publication of this periodical is necessary in transacting the public business of the U.S. Department of Justice as required by law.

Subscription Information

Online	https://puborder.ncjrs.gov/Listservs/nij/reg.asp
Phone	301-519-5500
	800-851-3420
Mail	NCJRS
	P.O. Box 6000
	Rockville, MD 20849-6000

World Wide Web Address
http://www.nij.gov/journals/welcome.htm

Contact NIJ
National Institute of Justice
810 Seventh St. N.W., Washington, DC 20531, USA
http://www.nij.gov/about/contact.htm

Director's Message

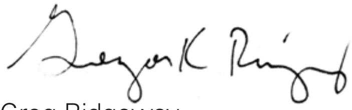
Before I arrived at NIJ in July 2012, I was already impressed with NIJ's research, especially with the innovation and intellectual curiosity that has led to major changes in the way law enforcement works — historic examples include the development of body armor and crime mapping. Now that I have been on board at NIJ for eight months and acting director since January 7, I am even more impressed with NIJ's potential for propelling innovation forward. We have remarkable people at NIJ and operate from a prominent national platform. I am committed to making decisions that will continue to ensure rigorous science and leading-edge activities.

Several articles in this issue of the *NIJ Journal* illustrate NIJ's commitment to innovation. The cover story, for example, shows how Geoffrey Barnes and Jordan Hyatt used sophisticated statistical techniques to create a computerized system that goes a long way toward predicting which probationers are most likely to violently reoffend within two years of returning to the community. Not only does their work illustrate innovation, it exemplifies two other primary NIJ goals: researcher-practitioner partnerships and translational criminology. Barnes and Hyatt formed a partnership with Philadelphia's Adult Probation and Parole Department and translated their work as they went along. They continually tailored the models to be what practitioners in the Department needed most. They custom-fit their research for the end-user.

The article about the pitfalls of prediction is a piece I wrote while I was an NIJ grantee at RAND; I presented it to a group of law enforcement agencies that were developing predictive policing programs. It is gratifying to see that the timing allowed the *NIJ Journal* to publish it in this issue with other articles about the ways researchers are using data to keep communities safer while also saving public safety dollars and practitioner time.

People who study innovation tell us that great ideas happen when networks of people connect. With NIJ's new Office of Research Partnerships, we are making deeper and stronger connections with researcher and practitioner networks — such as the National Science Foundation, the Academy of Criminal Justice Sciences, the International Association of Chiefs of Police (IACP) and the International Association of Crime Analysts (IACA). NIJ plans to actively participate in IACP and IACA's annual conferences this year.

As acting director, I intend to continue to foster the interchange of ideas between researchers and practitioners and to learn from each other so we can better understand how to use data to respond to the nation's most pressing criminal justice issues.



Greg Ridgeway
Acting Director, National Institute of Justice

Editor's Note: *Read an interview with Greg Ridgeway in* Amstat News, *the magazine of the American Statistical Association, at http://magazine.amstat.org/blog/2012/10/01/nij-ridgeway/.*

NIJ

February 2013
NIJ JOURNAL / ISSUE NO. 271

Contents

NIJ Bulletin

Newest Research Findings

Most Police Officers Wear Their Body Armor

NIJ funded the Police Executive Research Forum (PERF) to conduct a survey that asked law enforcement officers whether and why they use body armor. PERF sent surveys to 1,370 randomly selected police officers across the country; 1,080 (78 percent) of the surveys were completed and returned.

Results are encouraging — they suggest that body armor policies are effective, officers understand the importance of wearing body armor, and most officers are knowledgeable about body armor care and maintenance.

The survey found that officers wear their armor because they know it protects them and not because they have had firsthand experience needing its protection: 73 percent indicated that they had never been involved in a situation in which their armor protected them from possible injuries, but 90 percent said that they wore it because they thought it was "critical for safety."

Nearly all of the officers (99 percent) reported wearing body armor either all or most of the time when required to do so. Approximately half (49 percent) identified "agency policy" as a reason they wear body armor. Most officers (93 percent) reported that their agency required them to wear body armor. A smaller number (78 percent) said that their agency had a written policy. Less than 1 percent reported being disciplined for not wearing their armor.

▶ Read "A Practitioner's Guide to the 2011 National Body Armor Survey of Law Enforcement Officers" on NCJRS.gov. Keyword: NCJ 240225.

▶ Visit NIJ's topic page on body armor. Keyword: body armor.

Labor Trafficking in San Diego County

As many as 31 percent of unauthorized Spanish-speaking workers in San Diego County have experienced an incident that meets the official definition of human trafficking, and about 55 percent have experienced abuses or exploitative practices at the hands of their "coyotes" (those who guide them across the border) or their employers, according to a recent NIJ-funded study.

Researchers used advanced sampling methods and gained unique access to 826 unauthorized immigrant workers to gather reliable data about labor trafficking in San Diego County. The study looked at trafficking violations and abusive practices that occurred during migration or at the workplace. Trafficking violations included any direct infringement of freedom of movement.

Of six labor sectors singled out in the study, construction had the highest rates of trafficking violations (35 percent) and abusive labor practices (63 percent). Agriculture had the lowest rate of both trafficking violations (16 percent) and abusive labor practices (27 percent). The researchers speculated that the insulated and close-knit network of migrant farmworkers in northern San Diego County protects workers against victimization.

▶ Read the full report on NCJRS.gov. Keyword: NCJ 240223.

▶ Read more about the study on NIJ's human trafficking page. Keyword: labor trafficking.

▶ Learn more about human trafficking in this issue of the *Journal*: "Ending Modern-Day Slavery: Using Research to Inform U.S. Anti-Human Trafficking Efforts," page 26.

Publications
IN BRIEF

Geography and Public Safety, Volume 3, Issue 2: Using Geographic Information Systems (GIS) in Public Health and Public Safety

Researchers in criminal justice and public health are more regularly working with practitioners because they share an interest in poverty, chronic health problems and violence issues. The Chicago CeaseFire project, for example, used a public health framework to drive its epidemiological approach to serious gun violence prevention.

The latest issue of *Geography and Public Safety* draws on the parallels and intersections between criminal justice and public health approaches to violent crime intervention and prevention. Articles include:

- Integrating Emergency Department and Police Data to Locate and Prevent Violence: The Cardiff Model
- Using Public Health Strategies to Reduce Violence in "Hot Spots" in East Palo Alto, California
- Comprehensive Community-Based Information System to Reduce Youth and Gang Violence in Los Angeles County and Beyond
- What Can We Learn from Public Health? — An Example of Sharing Law Enforcement Spatial Data with Community Partners

▶ Read the issue at http://cops.usdoj.gov/Publications/e05122469c.pdf.

▶ *Geography and Public Safety* is published through a partnership between NIJ and the Office of Community Oriented Policing Services. Learn more about crime mapping at http://www.nij.gov/nij/topics/technology/maps/welcome.htm.

News & Notes

NIJ Partners with National Science Foundation

During his tenure at NIJ, former Director John Laub emphasized the importance of building partnerships that advance research and understanding in areas of common interest. To that end, NIJ established the Office of Research Partnerships and has formed or enhanced several partnerships over the last two years. NIJ also signed several memorandums of understanding (MOUs) that promote fiscal efficiencies through interagency activities.

In September 2012, Laub signed an MOU with the National Science Foundation (NSF). Although NSF's mission to promote scientific progress is far broader than NIJ's, both share an interest in the social, behavioral and forensic sciences, particularly as they relate to the juvenile and criminal justice systems. The MOU will give NIJ and NSF greater flexibility to coordinate workshops, identify opportunities for collaboration, sponsor research and leverage resources.

▶ Read more at http://www.nij.gov/nij/about/director/nsf-mou.htm.

.gov

Look for multimedia links throughout this issue of the *NIJ Journal.* On the NIJ website, look for the following new content:

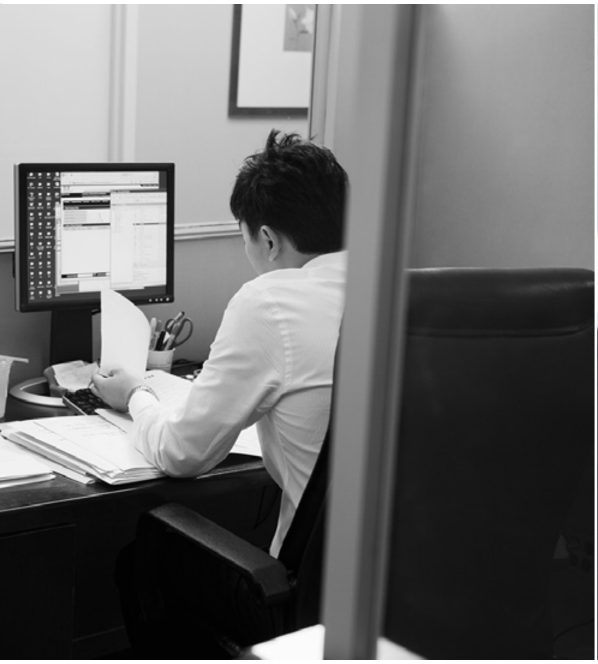
- Michael Jacobson on reforming the New Orleans criminal justice system
- New database of NIJ-funded projects, including descriptions
- New Web topic page on collecting arrestee DNA
- Updated Web topic pages on pursuit management and police smart radios
- Current and forthcoming solicitations
- Expert Web chat about the nexus between performance and health in officers
- Interviews with researchers and practitioners from the 2012 NIJ Conference
- New Computer Forensic Tool Testing program reports
- Past issues of the *NIJ Journal*

Follow @OJPNIJ on Twitter

Like NIJ on Facebook https://www.facebook.com/OJPNIJ

NIJ

http://www.nij.gov



Predicting Recidivism Risk: New Tool in Philadelphia Shows Great Promise

by Nancy Ritter

Tool uses random forest modeling to identify probationers likely to reoffend within two years of returning to the community.

For some time to come, our cities, counties and states will face the tremendous challenge of trying to do their work with fewer resources. That challenge is perhaps no more pressing than in the nation's corrections system, where fiscal realities demand the downsizing of prison populations.

In 2009, the Pew Center on the States estimated that 1 in 45 adults in the U.S. was under some form of community correctional supervision.[1] As ever-increasing numbers of offenders are supervised in the community — witness the massive "realignment" of prisoners in California — parole and probation departments must find the balance between dwindling dollars and the lowest possible risk to public safety. The good news is that researchers and officials in Philadelphia, Pa., believe they have developed a tool that helps find that balance.

Seven years ago, criminologists from the University of Pennsylvania and officials with Philadelphia's Adult Probation and Parole Department (APPD) teamed up to create a computerized system that predicts — with a high degree of accuracy — which probationers are likely to violently reoffend within two years of returning to the community.

"We were asked to develop a new risk-forecasting tool to help the financially strapped probation department tailor their officers' caseloads to the risk level of probationers," said Geoffrey Barnes, who, with fellow researcher Jordan Hyatt from Penn's Jerry Lee Center of Criminology, created and evaluated the tool. "The goal was to ensure that officers who were supervising probationers with a high risk of recidivating would have a smaller caseload than officers who were supervising folks with a lower risk."

The tool — which has been successfully used in Philadelphia for four years — assesses each new probation case at its outset and assigns the probationer to a high-, moderate- or low-risk category. Although this is not a new concept, what is unique is that the tool uses "random forest modeling," a sophisticated statistical approach that considers the nonlinear effects of a large number of variables with complex interactions (see sidebar, "What Is Random Forest Modeling?" on this page). Historically, corrections officials — in Philadelphia and elsewhere around the country — have used simpler statistical methods, such as linear regression models, to try to get a handle on the risk that a probationer may pose to the community.

Random forest modeling, as applied to criminal justice, was pioneered by criminologist Richard Berk, also at Penn, who acted as a consultant on the NIJ-funded project.

Pre-Random Forest Times

Before the creation and implementation of the new risk-forecasting tool, Philadelphia — like many of the nation's parole and probation departments — used a one-size-fits-all supervising strategy. Every offender saw his or her probation officer about once a month for 20-30 minutes.

"Most of APPD's probationers were supervised under a strategy that mandated only two and a half hours of interaction per year," said Barnes.

"When they contacted us, the department's leaders expressed a strong desire to reform this policy — to focus more supervision on those with the largest risk of future violence and devote far fewer resources on those who presented little or no risk of reoffending."

Parole and probation departments must find the balance between dwindling dollars and the lowest possible risk to public safety. Researchers and officials in Philadelphia believe they have developed a tool that helps find that balance.

What Is Random Forest Modeling?

Random forest modeling is the technique used by Richard Berk — working with NIJ-funded researchers Geoffrey Barnes and Jordan Hyatt — to build the risk prediction tool for Philadelphia's Adult Probation and Parole Department. Random forest modeling could best be described as hundreds of individual decision trees.

In the simplest statistical terms, here is how it works: Data are organized using a technique called "classification and regression trees." The computer then runs an algorithm that selects predictors at random and repeats and repeats this process to build several hundred trees — which then allow the randomly selected predictors to average themselves into a single outcome. In the case of the Philadelphia tool, this outcome was assignment to one of three risk categories (high, moderate or low) for probation-supervision purposes.

The final NIJ report describes random forest modeling — and the fine-tuning that the research partnership went through as they built three iterations of the risk prediction tool — in much more detail (http://www.ncjrs.gov/pdffiles1/nij/grants/238082.pdf).

Is random forest modeling an improvement over more traditional actuarial prediction analyses? Barnes and Hyatt say yes.

"It allows for the inclusion of a large number of predictors, the use of a variety of data sources, the expansion of assessments beyond binary outcomes, and taking the costs of different types of forecasting errors into account," Barnes said.

To develop the tool they had in mind, Barnes and Hyatt were actually "embedded" in APPD. Over the next few years, they built three iterations of a model that makes virtually instantaneous forecasts regarding offenders who are due to be released to the community.

Since APPD began using on-demand risk forecasting, the agency has handled well over 120,000 new "case starts," referring to the time when an offender begins probation. (Note that about one-third of the offenders have had more than one probation "case start," so this number actually reflects about 72,000 individual offenders.)

In 10-15 seconds, the tool assigns a new probationer to one of three categories. The lowest level of risk is assigned to those who are predicted to not commit any new offense in the next two years. The moderate-risk level identifies those who are likely to commit a crime, but not a serious one. The high-risk level is for those who are most likely to commit a serious crime, which APPD defines as murder, attempted murder, aggravated assault, rape and arson.

Community supervision is based on the determined risk level. Probation officers who are supervising high-risk individuals are given the smallest caseloads.

Getting Started

Although the random forest model developed in Philadelphia can be adapted by other jurisdictions, it is not an off-the-shelf tool. Obviously, the data are unique to the probationers who are under APPD supervision. And the "outcomes," or risk-level assignments, are also unique to Philadelphia because APPD officials set their own parameters based on resources and every manner of policy, operational and political reality that the tool is asked to consider.

> "The real achievement of the final model in Philadelphia is not that it is right two-thirds of the time but that it produces this accuracy by balancing the relative costs of the different kinds of errors."

Hyatt offers this analogy: You could take the engine out of a custom sports car, but it probably wouldn't work the same way in another car — and it might not even work at all. Therefore, another jurisdiction using random forest modeling to build a risk-prediction tool would need its own statisticians, computer whizzes and agency officials working in concert. Although many of the same questions would be asked, the "answers" — specifically, the outcomes around which the tool would be designed — would be different.

The first thing a jurisdiction interested in creating a random forest risk-forecasting tool must do is determine what data already exist in electronic form. It is very possible, say Barnes and Hyatt, that a jurisdiction will discover it has far more data than it realizes — criminal histories in the court system, local prison records and separate police records.

"Every jurisdiction probably has access to data that they haven't even thought about," said Barnes. "We capture so many types of information as a matter of course, as part of the day-to-day routine. I suspect few people realize how enormously powerful it could be to — with just a few manipulations — convert it into numbers that could forecast future behavior."

As they developed the risk prediction tool in Philadelphia, Barnes and Hyatt mined raw data from six different databases. The team then tested hundreds of different predictors using many different approaches, all the while fine-tuning the delicate balance between APPD's resources and the forecasting accuracy that was achievable. Eventually, three models went live. The third, Model C, has been in operation since November 2011 and uses 12 of the strongest predictors of risk of reoffending, including prior jail stays, the probationer's ZIP code and the number of years since the last serious offense.

Every jurisdiction would be looking at its own very unique data set that reflects decisions, made by people who have long since retired, about what should and should not be rolled over into their next system. Therefore, it would be especially helpful for one of the team members to understand what data were taken from an older system — be they paper records from jails, courts or police, or old computer records — and used in newer systems.

"You definitely need a computer professional on the team from the beginning," said Barnes. "Ideally, this would be someone familiar with the way the jurisdiction has kept its records."

Finally, it is important to be mindful of simple geography. It will come as no surprise that data-sharing among jurisdictions in the U.S. is quite limited, particularly in terms of the kind of instantaneous forecasting that this tool is designed to perform. For example, the APPD tool uses criminal history data only from Philadelphia; data from other states, and even other parts of Pennsylvania, are not used, which means that the forecasts do not necessarily indicate each probationer's universal level of risk.

"Offenders who represent a serious danger outside the city of Philadelphia could very easily be forecasted as low risk within these boundaries, particularly if they usually live, work and offend elsewhere," Hyatt explained.

The bottom line is that, as in every scientifically based endeavor, data are paramount.

"The key," added Barnes, "is to ensure that all of the data sources are immediately available through the agency's data network, although it is important to note that the data do not need to be up-to-the-minute accurate to be useful."

Forecast Begin- and End-Points

After dealing with the availability of data, the next step is to determine when the forecasting begins (called the "unit of prediction") and when it ends (the "time horizon"). The beginning point can be any moment in the lifespan of an offender's case — when bail is set, when charges are filed, at sentencing, when the offender enters the correctional system or when the offender first reports for probation.

In Philadelphia, officials chose the start of probation and a time horizon of two years. The APPD tool therefore predicts the likelihood of a probationer committing a violent crime within two years of returning to the community. Although any time period can be used, it is important to understand that the accuracy of forecasting a longer period depends on the depth of data available.

"If, for example, you want to forecast what is going to happen over the next five years, you have to use data from at least five years ago and before," Hyatt said.

Once the unit of prediction and time horizon are determined, the next step is to decide what "forecasting outcomes" the tool should be set up to predict. Researchers such as Barnes and Hyatt can guide practitioners through this process, but the practitioners themselves must ultimately make the decisions because resources, personnel, operational and even political realities must be considered. In Philadelphia — after weeks of examining caseloads and staffing levels — officials decided that approximately 15 percent of their probation population should be classified as high risk, 25-30 percent as moderate risk, and 55-60 percent as low risk.

Barnes and Hyatt acknowledge that someone picking up the final report they submitted to NIJ at the end of the grant could be a bit overwhelmed by random forest modeling. The forecasting tool now being used in Philadelphia, for example, looks at 500 decision "trees" (hence random "forest") as it runs a risk assessment of a new probationer. But, they insist, there is no reason that criminal justice practitioners should shy away from the technology.

"If you think about it," said Barnes, "private companies do this every day — they crunch data to decide who's likely to buy peanut butter, for example, and they send coupons to those folks."

Of course, both researchers are quick to point out that forecasting criminal behavior is not coupon clipping, but the principles of data analysis, they say, are the same.

Determining an Acceptable Error Rate

No prediction tool is perfect. Anyone who has watched a weather forecaster predict 8 inches of snow — then dealt with crying children who have to go to school when only a dusting falls — knows that predictions are occasionally wrong.

The key in building a random forest prediction tool for any aspect of the criminal justice system is balancing the risk of getting it wrong. This process involves determining, in advance, an acceptable error rate. And this demands intensive collaboration between researchers and practitioners, one in which agency officials — not statisticians — must make crucial policy decisions. In particular, this means determining prespecified levels of "false positives" to "false negatives."

A false negative is an actual high-risk person who was mistakenly identified as moderate or low risk. A false positive is an actual low- or moderate-risk person who was identified, and therefore supervised, as high risk.

As the practitioners work side-by-side with the researchers to set these parameters, they will inevitably encounter the need to make tradeoffs they can live with. This is referred to as the "cost ratio." Before the risk prediction tool can be built, the numerical ratio of these costs must be approximated. It is not enough to simply say that false negatives are generally more costly than false positives. Rather, an actual *value* must be provided.

Here is how Hyatt explained the process in Philadelphia: "Basically, we had to determine precisely how much more costly it would be to mistakenly classify a probationer in a lower-risk category who then went on to commit a serious crime than it would be to intensely supervise someone who is actually a low-risk probationer because the tool had assessed him as high risk."

Most jurisdictions that contemplate building a random forest risk-prediction tool would likely do what they did in Philadelphia: set a higher relative cost for false negatives than for false positives. Philadelphia's APPD decided on a cost ratio where false negatives were 2.6 times more costly than false positives. But any jurisdiction that wishes to design and implement a similar tool would have to determine its own cost ratio or error rate.

As Barnes and Hyatt noted, there is no single 'right answer' in choosing the unit of prediction, the time horizon, the definition of outcomes or the cost ratio. Every jurisdiction that wants to build a random forest model prediction tool must commit to this very delicate balancing act — one in which researchers can assist, but that, in the end, requires practitioners to do the heavy lifting.

"Random forest modeling allows you to add different variables without sacrificing your ability to make accurate predictions."

"I cannot emphasize this enough," Barnes added. "Balancing these different types of errors with the model's overall accuracy rate is not the job of the team's statisticians. Because an agency's leadership has to live with the consequences of any error that occurs once the forecasting tool goes live, they must decide what level of accuracy they can live with and the balance of potential errors they prefer."

Accuracy

The model that has been used in Philadelphia for just over a year (Model C) has an accuracy rate of 66 percent when considering all three (high-, moderate- and low-risk) categories. In their final report to NIJ, Barnes and Hyatt offer a detailed account of the development and accuracy of the three generations of risk-prediction models, including much more detail about the separate accuracy rates for the three risk categories; for example, probationers who were categorized as high risk are 13 times more likely to commit a new serious offense within the two-year forecast period than either low- or moderate-risk probationers. The NIJ report is available at http://www.ncjrs.gov/pdffiles1/nij/grants/238082.pdf.

All three iterations of the Philadelphia model were validated using a sample of probation cases from 2001, which gave the researchers a 10-year period in which to assess the long-term offending of the probationers. That said, of course, any forecasting tool, including this one built using random forest modeling, will make mistakes.

But, said the researchers, when it comes to figuring out how to be more effective in using corrections system dollars, everyone should understand that choices will always have to be made — and the goal is to make the most accurate choices in as cost-effective a manner as possible.

As Barnes put it, "The real achievement of the final model in Philadelphia is not that it is right two-thirds of the time but that it produces this accuracy by balancing the relative costs of the different kinds of errors."

"The point," Hyatt added, "is that random forest modeling allows you to add different variables without sacrificing your ability to make accurate predictions. By working hand-in-hand with their practitioner and policymaker partners, researchers can come up with the right ratio of variables that work in their own unique jurisdiction, both from a practical standpoint in terms of the data that are available and from a standpoint of political and policy exigencies which decision-makers are comfortable putting into a forecast tool."

It is important to understand that the NIJ-funded project discussed in this article looked only at the creation and effectiveness of the prediction tool itself — not at the effectiveness of the subsequent supervision or treatment of APPD probationers. In other words, the project did not, for example, consider whether (and to what extent) intense supervision and exposure to more aggressive interventions may have caused a high-risk probationer to not commit another serious crime.

The Benefits of Random Forest Modeling

One of the most compelling attributes of random forest modeling is that — unlike linear regression analyses — it is not necessary to know in advance what data will be useful in predicting behavior or which variables will affect the predictive power. In more traditional statistical procedures, only a limited number of predictors are used to try to forecast future behavior. But random forest modeling does not require users to be so choosy.

The tool can be programmed to simply not consider a factor based on other variables. In other words, data can be "over-included," and the tool will simply filter them out. For example, the tool may say, "I don't see much of a juvenile record for this individual, but I do see, from an earlier branch in the tree, that this person is 60 years old, so I wouldn't expect to find much of a juvenile record; but, regardless, now that he is 60, this is probably not a very important factor now."

This is not the case with regression equations, where every time another variable or predictor is added, something is lost. With random forest modeling, variables can be added

Proponents say one of the most compelling reasons for building a risk analysis prediction tool using random forest modeling is the simple matter of fiscal resources.

without losing predictive capacity. Indeed, it is this feature that can help bring researchers, practitioners and even politicians to the same table while the tool is being developed. It helps garner buy-in, as it were, from skeptics.

"Adding variables that individual stakeholders cared about — even if we, as criminologists, didn't think they would have much predictive power — helped our APPD partners feel that we were hearing them and responding to their concerns," said Barnes. "This feature helped them get behind what we were trying to do as we built the forecasting tool, and, importantly, it helped everyone understand the risks that the policy-makers, in particular, faced."

The bottom line is that any data can be used in a random forest tool, depending on the wishes of officials and other key players. Data that may be statistically unimportant — but politically important — can be built into the tool. For example, a jurisdiction might want to consider the number of a probationer's violent co-offenders; although APPD ended up not using that data in its tool, another jurisdiction may find such data predictive.

Another advantage of random forest modeling is its ability to identify highly nonlinear effects for each individual predictor. Consider, for example, the bivariate relationship between a soon-to-be-probationer's age and the likelihood that the tool would forecast him to be high risk. It is not surprising that the youngest probationers in Philadelphia were forecast to present the greatest danger of a serious-crime reoffending. However, the random forest analysis also showed something else.

"A bit more surprising is how quickly the probability of a high-risk forecast dropped as the offender got just a few years older," said Hyatt. "By the time the incoming probationer turns 27, the likelihood of receiving a high-risk forecast is not appreciably different from that of a 40-year-old — and, after the age of 40, the amount of risk seems to drop once again until it reaches a level that is effectively zero at age 50 and beyond."

Resources, Equitability and Fairness

Why should officials in the criminal justice system think about building a risk analysis prediction tool using random forest modeling? Proponents say one of the most compelling reasons is the simple matter of fiscal resources.

"We just do not have the ability to pay for the most intensive level of supervision for every probationer," said Barnes. "We don't have the ability to sentence every prisoner to life. We have to be very careful about how we allocate precious resources and, for public-sector workers — be they probation officers, police officers or corrections officers — the most precious resource is time."

The random forest model prediction tool, he said, allows agencies to base their personnel and policy decisions on a scientifically proven method.

Another reason to consider constructing such a sophisticated prediction tool is that, quite frankly, "prediction," in some form or another, is already occurring. Everyone involved in the criminal justice system — from judges to probation officers, from police leaders to politicians who write the laws and determine budgets — is making judgments, essentially predictions, about the relative risk of an offender.

Researchers Hyatt and Barnes believe that by using random forest modeling to build the actuarial risk-assessment tool for Philadelphia's APPD, they have ensured that those predictions are being made in the fairest, most equitable way possible.

"Using random forest modeling gave us the assurance that we made use of the best science available to identify the most dangerous offenders," said Barnes. "It has ensured that we're preserving resources and that the people who are subject to the policy decisions based on those risk assessments are being treated in a fair and consistent way."

"You may not like being on high-risk probation," he added, "but from a procedural justice standpoint, you at least know that the decision was made the same way for everybody."

Under one-size-fits-all procedures being used in many jurisdictions around the country, probation officers are given an enormous amount of discretion. This means that probationers who actually have a similar risk of reoffending could be — and therefore likely are — treated in disparate ways based on who their probation officer is and any number of other factors.

The random forest model prediction tool allows agencies to base their personnel and policy decisions on a scientifically proven method.

However, in addition to ensuring that offenders are assessed consistently in terms of their risk level, the tool being used in Philadelphia — and the policy decisions that APPD has put into place to operationalize the results — ensures that offenders who are identified as being at a certain risk level are all treated the same. Every probationer whom the tool scores as high risk is treated under the same high-risk protocol; this standardizes both their reporting requirements and the rules that they have to follow — including, of course, any likelihood that they will be sanctioned for a technical violation.

This equitability is something that researchers Barnes and Hyatt — and the probation professionals who have been successfully using the tool — believe in.

"Because every probationer is put into the same model, the same decision points will be hit as the model produces the risk-category analysis," Barnes said. "Two offenders with the same data values — even if they come from different parts of the city, even if they are different kinds of people — will go through the same scoring process in the same way."

"And that," Barnes argued, "is a far sight more equitable than a probation officer perhaps taking a dislike to you and deciding that you need to come in more frequently because you remind him of somebody who victimized a close relative a few months ago."

This is not to say, however, that human judgments don't play a role.

"Human judgments are important," Hyatt added. "But one thing that has been consistently found every time that this sort of technology has been used to forecast human behavior is that these actuarial decision-making models do a better job — and produce more accuracy in a more consistent fashion — than human gut reactions ever could."

As with any kind of new technology-based tool, however, there is an inevitable intersection of science and human nature — including ethics — that must be grappled with.

For example, some have argued that using some variables, such as an offender's ZIP code — particularly in a city as highly segregated as Philadelphia — can be a proxy for race. Others note that individuals who are categorized as high risk and therefore more intensely supervised are probably going to incur more technical violations of the terms of their parole. Certainly, just as any policy decision that has moral and ethical ramifications (and most do),

The Role of Ethics in Statistical Forecasting

The ethical considerations inherent in trying to predict future events — such as criminal offending — are not new. Indeed, as the NIJ-funded researchers who worked on the Philadelphia risk assessment tool point out, one of the reasons some offenders are sentenced to longer prison terms is to prevent crimes that they might commit if they were not incarcerated.

Geoffrey Barnes and Jordan Hyatt, from the University of Pennsylvania, believe that random forest modeling offers a different — and potentially more accurate — approach for building a prediction tool. Nonetheless, they recognize the ethical crux that lies at the heart of building such a tool: deciding *which* "predictors," or fact variables, are acceptable to use.

In their final report, for example, they ask, "Would it ever be permissible … to include an offender's racial background as a predictor variable in one of these models? If not, what about the use of predictors such as residential location or familial circumstances, which could indirectly communicate the offender's racial identity into the forecasting model?"

Would it be permissible to use controversial predictors in "lower-stakes" forecasting models — to control admission into a treatment program or govern supervision decisions, for example — but prohibit their use in "higher-stakes" decision-making such as sentencing?

Furthermore, some note, aren't the age of criminal-behavior onset, possession of a juvenile record or the neighborhood a person resides in (factors that *could* be used as prediction variables) all "extrajudicial" factors? As such, should they be considered in an individual criminal justice decision?

Considering potential "collateral consequences" of decision-making based on a forecasting tool is also an important part of the process. As mentioned in the main article, for example, Philadelphia's Adult Probation and Parole Department used the random forest prediction tool to identify offenders who were at a high risk of committing a serious crime in the two years following return to the community — and these people were supervised more closely, under more stringent parole terms and conditions. This could increase the likelihood that technical violations of their parole would be more likely to be detected and punished, including imposing additional custodial sanctions.

There are no easy answers to these questions, but they will have to be addressed head-on as increasingly technologically advanced forecasting methods become available for use in our nation's criminal justice system.

"One thing that has been consistently found every time that this sort of technology has been used to forecast human behavior is that these actuarial decision-making models do a better job — and produce more accuracy in a more consistent fashion — than human gut reactions ever could."

it is important that these issues are clearly understood and squarely addressed (see sidebar, "The Role of Ethics in Statistical Forecasting," on this page).

The Key: A Strong Partnership

Barnes and Hyatt emphasize that building the random forest prediction tool in Philadelphia was a tremendously iterative process — and one that required day-to-day collaboration with APPD.

"You don't put all the data into the computer the first time and hit the button and say, 'OK, we're done,'" Barnes said. "The model comes out and you look at it. Everyone sits down around a table and discusses it. The statisticians describe the problems they faced. The database guys look at it and say, 'Well, yes, but you are using this variable in the wrong way,' and the practitioners look at it and say, 'We really can't have 35 percent of our caseload on high-risk supervision. It's not going to work. That number has got to come down.'"

This gets at the constantly evolving nature of the random forest tool.

"You constantly are building new things to try to deal with changes in the environment, changes in the data, changes in what people think are predictive, changes in chronological theory over time," Hyatt noted.

Recommendations from the Research

Given the need to balance fiscal realities with an overarching mission to protect public safety, criminal justice professionals are beginning to look — with the same creativity and vigor as private-sector professionals — at sophisticated statistical tools to solve problems. Therefore, it is likely that risk-prediction tools using random forest modeling may play an important role in the future of our criminal justice system.

A tool like the one developed in Philadelphia provides an opportunity to advance the capabilities of the criminal justice system to protect communities, particularly for jurisdictions with large probation populations that must be managed with fewer dollars. Indeed, as the Philadelphia project demonstrated, the random forest-based risk-prediction tool has helped probation officials manage cases more efficiently. For nearly four years now, they have been

A tool like the one developed in Philadelphia provides an opportunity to advance the capabilities of the criminal justice system to protect communities, particularly for jurisdictions with large probation populations that must be managed with fewer dollars.

able to concentrate resources on a small number of probationers who require more active supervision, rather than on those who are unlikely to reoffend regardless of how they are supervised.

In their final report, Barnes and Hyatt recommend 12 steps that could serve as a blueprint for a jurisdiction that is considering building a random forest model risk prediction tool:

1. Obtain access to reliable data that are consistently and electronically available.
2. Define the unit of prediction and time horizon.
3. Define the outcome risk categories.
4. Consider the practical implications for a risk-based supervision strategy and ensure adequate resources based on the distribution of risk scores.
5. Choose the predictor variables to be used, based on theoretical, practical and policy considerations.
6. Build a single database file.
7. Estimate the relative costs of false positives and false negatives, keeping in mind that agency leadership must value the relative weight of these inaccuracies.
8. Build an initial model and evaluate the results.
9. Adjust the model to reflect policy-based concerns regarding accuracy and proportional assignment to risk categories; construct additional test models where required.
10. Produce forecasts for offenders already in the agency's caseload.
11. Create the user interface and back-end software to produce live forecasts.
12. Continuously monitor the results of the live forecasts.

Again, it is important to understand that the Philadelphia tool was based on probationers who live in Philadelphia. Needless to say, people in other jurisdictions may be different in key ways — and crime trends vary in different parts of the country and even in different parts of a state. Therefore, a tool that uses random forest modeling must be based on the best available data about the population whose behavior is being predicted.

Finally, say proponents, because random forest modeling can be tailored to specific needs, researchers and practitioners should not limit their thinking to urban probationers, such as those with whom the team worked in Philadelphia. Random forest modeling may prove useful in managing prison populations, for example. Or, said Barnes, perhaps officials in another jurisdiction are interested in looking at the pretrial behavior of people who have merely been charged with an offense.

These would present entirely different environments, of course.

"But," Barnes noted, "the chances are that a jurisdiction has the data to build other kinds of prediction models."

"You just have to make the contact with somebody with reasonable statistical skills, use the database professionals who you almost certainly have already employed, convert the data into a usable format, and go ahead and build the model," he added. "Give it a shot."

About the author: Nancy Ritter is a writer and editor at NIJ.

NCJ 240696

Notes

1. Pew Center on the States, *One in 31: The Long Reach of American Corrections,* Washington, D.C.: Author, 2009, http://www.pewstates.org/uploadedFiles/PCS_Assets/2009/PSPP_1in31_report_FINAL_WEB_3-26-09.pdf (accessed November 27, 2012).

Watch researchers Geoffrey Barnes and Jordan Hyatt discuss the Philadelphia risk assessment tool: http://www.nij.ncjrs.gov/multimedia/video-barnes-hyatt.htm.

Software Tools, Apps and Databases

NIJ has funded research and development that has resulted in more than 50 free or low-cost software tools, apps and databases to assist with investigations and research — and now they are all gathered in one place on NIJ.gov.

Tools available in NIJ's online catalog include:

- **Video Previewer:** A program that assists investigations involving time-consuming video review by quickly processing the video and showing key frames in a PDF.
- **U.S. Y-STR Database:** An online searchable listing of 11- to 17-locus Y-STR haplotypes, which are required to provide a statistical estimate of a match's significance.
- **3D-ID:** A Java program that provides geometric morphometric tools to help assess the sex and ancestry of unidentified cranial remains.
- **CrimeStat 3.4:** A spatial statistical program used to analyze crime locations and identify hot spots.

To find other software tools, apps and databases, browse our catalog at http://www.nij.gov/topics/technology/software-tools.htm.

Ballistic Body Armor: A Closer Look at the Follow-Up Inspection and Testing Program

by Michele R. Coppola

The NIJ program helps ensure that body armor coming off the assembly line meets the requirements of NIJ's standard.

In the mid-1970s, NIJ began developing performance standards for body armor to help provide confidence that law enforcement officers are properly and consistently protected each and every time they face gunfire in the line of duty. Since that time, body armor has been credited with saving the lives of more than 3,100 law enforcement officers.[1]

NIJ's most recent body armor standard — *Ballistic Resistance of Body Armor, NIJ Standard-0101.06,* published in July 2008 — establishes minimum performance requirements and test methods for the ballistic resistance of body armor designed to protect the torso against gunfire (see sidebar, "Revising the Body Armor Standard," on page 15). Although this standard and all other NIJ standards are voluntary — that is, manufacturers are not required to follow them — many public safety agencies require compliance with NIJ standards before they purchase equipment. Through the NIJ Compliance Testing Program (CTP), manufacturers can voluntarily submit equipment samples for testing by NIJ-approved laboratories to determine whether their models comply with a particular standard.

The National Law Enforcement and Corrections Technology Center System's National Center (NLECTC-National) oversees NIJ's body armor conformity assessment efforts.

With funding from NIJ, NLECTC-National administers two distinct phases of conformity assessment through the CTP. Phase 1 involves documenting the design of an armor model and testing up to 28 samples voluntarily submitted by manufacturers to verify that the model meets the standard's minimum performance requirements. Models that meet the standard are added to the Compliant Products List, which can be found at http://www.nij.gov/topics/technology/body-armor/compliant-ballistic-armor.htm.[2]

But how do we ensure that recently manufactured body armor is constructed similarly to samples that were previously tested and deemed compliant with the NIJ standard?

This is where phase 2 comes in. NLECTC-National began implementing the second part of the conformity assessment effort, called the Follow-Up Inspection and Testing Program, in 2010. The program subjects new armor samples to additional ballistic testing and compares the construction of newly made armor with samples evaluated in phase 1, providing confidence that body armor coming off the assembly line is manufactured consistently and performs in accordance with NIJ standards.

"The follow-up program provides an additional set of eyes and ears into the manufacturing process," said Lance Miller, NLECTC-National director. "We want to ensure that the men and women who wear these vests on a daily basis have as much confidence in these products as we can possibly give them."

How the Program Works

The follow-up program applies to armor models deemed compliant

> "We want to ensure that the men and women who wear these vests on a daily basis have as much confidence in these products as we can possibly give them."

with the 2008 NIJ standard. Each month, the CTP staff reviews the number of models a manufacturer currently has on the Compliant Products List that have not been inspected within the past 10 months and prepares a list of models and manufacturers for follow-up inspection.

Independent inspectors conduct surprise inspections at these manufacturers' locations. If a manufacturer does not agree to this follow-up inspection and testing, its armor will not remain on the Compliant Products List.

The inspectors randomly select two newly manufactured vests for each identified model of interest and send them to NIJ-approved, accredited laboratories for testing. The laboratories send the test results to Underwriters Laboratories, an independent, not-for-profit testing and certification organization, for processing. Meanwhile, the laboratories send the vests to CTP staff, who also inspect the armor's construction. Both steps — the laboratory testing and the inspection of construction — help ensure that the manufacturer has built the newly manufactured vest the same way as vests previously submitted for phase 1 testing.

What Inspectors Found

Inspectors conducted their first follow-up inspection in September 2010. Through August 2012, they had visited 90 manufacturing locations in five countries (United States, Canada, Mexico, Colombia and the People's Republic of China) and tested 222 body armor models, according to Jamie Phillips, NLECTC-National conformity assessment coordinator. Of those models, five sustained multiple perforations during laboratory testing. Subsequently, the manufacturers issued recalls and replaced more than 1,750 fielded

Revising the Body Armor Standard

NIJ anticipates that a Special Technical Committee will begin revising the ballistic-resistant body armor standard in 2013. As a first step, NIJ has held workshops to obtain comments and suggestions from manufacturers of body armor. It also held a "needs and requirements" meeting, during which officers identified the operational environments in which they work, missions and roles performed while wearing armor, and other equipment that may be affected while wearing armor.

armors to ensure that practitioners had effective ballistic body armor that complied with the NIJ standard. The manufacturers also took corrective actions to fix what was causing the perforations.

"Staff at NIJ and NLECTC have worked actively with manufacturers to identify the root cause of these performance issues," explained Miller. "In cases where it was a significant issue, manufacturers voluntarily took immediate action to recall and replace units or take some sort of corrective action out in the field."

To date, inspectors have discovered eight models with major variations in construction that could affect ballistic performance. For example, in one case, the number of layers in the follow-up testing vest samples differed from those in the original samples; in another, leaking covers allowed water to penetrate to the ballistic panel. Inspectors also identified

"We view the standard itself as a living, breathing document that is flexible and can adapt to changing trends in the industry and new testing methods."

33 models with minor variations in construction that would not affect ballistic performance. In response, manufacturers worked with the CTP team to implement quality-control improvements at several locations to prevent these and other variations in construction.

Moving Forward Together

Inherent to the follow-up inspection process is increased communication between body armor manufacturers and the assessors.

"The program provides an opportunity to work more closely with manufacturers to ensure that fielded armor is more likely to comply with the NIJ standard," Phillips said. "It allows manufacturers to express their concerns, and we, in turn, are able to explain the reasons behind our decisions and how those decisions support the law enforcement community as a whole."

NIJ does not anticipate major changes to the follow-up process, but staff will continue to explore opportunities for improvement. "I think we view it in the same light as the entire Compliance Testing Program," Miller noted. "We view the standard itself as a living, breathing document that is flexible and can adapt to changing trends in the industry and new testing methods, and I don't see the follow-up program as any different."

"We obviously have learned much," he added. "As we continue this dialogue with manufacturers, we continue to learn more about the body armor manufacturing processes and how quality management in that industry works. And as we learn more, we will adapt the program."

NIJ's Body Armor Challenge

Most law enforcement agencies replace their body armor every three to five years — the typical length of the manufacturer's warranty. However, scientists suspect that the ballistic performance of an individual vest may vary due to a variety of physical, chemical and environmental factors.

In September 2012, NIJ issued the first Department of Justice Challenge, asking scientists, inventors and innovators to submit creative ways to test the performance and usability of body armor without destroying it. The goal of the challenge is to empower those who depend on this critical safety equipment to make informed decisions based on solid scientific evidence regarding the ballistic performance of the body armor they use. Winners of the first phase of this multiphase challenge are expected to be announced in March 2013. To read the challenge, visit http://www.nij.gov/funding/2012/body-armor-challenge.htm.

About the author: Michele R. Coppola is a senior writer and editor at the National Law Enforcement and Corrections Technology Center System's National Center.

NCJ 240697

For more information:

- Visit http://www.justnet.org/body_armor/index.html or contact NIJ Program Manager Michael O'Shea at michael.oshea@usdoj.gov.

Notes

1. DuPont, "IACP/DuPont Kevlar Survivors' Club," Wilmington, Del.: Author, 2012. Retrieved December 3, 2012, from http://www2.dupont.com/personal-protection/en-us/dpt/article/kevlar-survivors-club.html.
2. There is a similar list for stab-resistant body armor: http://www.nij.gov/nij/topics/technology/body-armor/compliant-stab-armor.htm.

Learn more about body armor: http://www.nij.gov/topics/technology/body-armor/welcome.htm.

Watch how body armor works to ensure that it keeps officers safer: http://nij.ncjrs.gov/multimedia/video-body-armor-officer.htm.

See how procurement officials can find the right vest for each officer: http://nij.ncjrs.gov/multimedia/video-body-armor-purchasing.htm.

THE NIJ CONFERENCE

Looking Back to See the Future of Prison Downsizing

Recent declines in U.S. prison populations have caused many reformers to suggest that America's experiment with mass incarceration is ending. But current prison downsizing policies may well backfire if we fail to heed the lessons learned from the intermediate sanctions movement of the 1990s. Delivering the keynote address at the 2012 NIJ Conference, Joan Petersilia summarized these lessons and discussed why we must consider them if we want to reverse — for good — four decades of prison expansion.

- Watch Petersilia's keynote address at http://nij.ncjrs.gov/multimedia/video-nijconf2012-laub-petersilia.htm.

Two plenary sessions from the 2012 NIJ Conference are also available for viewing:

- **Game Change:** How Researcher-Practitioner Partnerships Are Redefining How We Study Crime at http://nij.ncjrs.gov/multimedia/video-nijconf2012-opening-plenary.htm.
- **Protecting Our Protectors:** Using Science to Improve Officer Safety and Wellness at http://nij.ncjrs.gov/multimedia/video-nijconf2012-closing-plenary.htm.

The Grant Progress Assessment Program: Looking Back on Success and Moving Forward

Several people contributed to this article, including Patricia Kashtan, Jolene Hernon and Beth Pearsall.

The NIJ program helped educate grantees on how to properly administer their grants.

Every year, NIJ awards several hundred grants totaling millions of dollars to state and local laboratories to help them improve their ability to conduct accurate and timely forensic testing.

As administrator of the funds, NIJ must monitor and assess how the laboratories spend these taxpayer dollars and then report to Congress on what it finds. To help accomplish its monitoring goals, in 2005 NIJ instituted a program to ensure that grantees (i.e., the laboratories) were correctly documenting their efforts and spending funds according to congressional and NIJ guidelines.

The Grant Progress Assessment program provided free, external assessments to grant recipients from mid-2005 to 2011. During that time, the Grant Progress Assessment staff examined more than 2,300 awards worth a total of more than $1 billion.

When NIJ suspended the program in September 2011 due to budget constraints, all the players — NIJ, assessors and laboratories — agreed that the program had been a great success. This article documents the lessons learned.

Goals of the Grant Progress Assessment Program

The program's purpose was twofold:

- To assist NIJ in its administrative oversight of forensic science awards
- To educate grantees on proper grant administration

The program gave NIJ a comprehensive overview of the awards process as well as in-depth data about grantee performance. In addition, it gave the laboratories objective, professional feedback about how they managed their funds. Through the assessments, laboratories came to better understand the program's requirements and special conditions. Over time, laboratory staff were able to identify potential issues and resolve them before they became problems. They also improved their ability to identify successes that they could use to secure future funding.

NIJ used the Grant Progress Assessment program to monitor several kinds of forensic science grants:

- Forensic DNA Backlog Reduction Program
- Capacity Enhancement Program
- Convicted Offender and/or Arrestee DNA Backlog Reduction Program
- Solving Cold Cases with DNA
- Paul Coverdell Forensic Science Improvement Grants Program

How the Program Worked

On a set two-year cycle, a trained assessor or team of assessors — many of them DNA laboratory managers themselves — visited each laboratory that had an open forensic science grant.

Using a checklist, the assessors reviewed the status of the laboratory's grant and assessed the grantee's use of federal funds to increase the laboratory's capabilities and capacities — all at no cost to the agencies. Generally, the assessors spent two to five days onsite with the laboratory staff.

> "You get the quality you inspect, not the quality you expect."

The checklist included everything from budgets to performance measurement data to deliverables.

Assessment Led to Education

The assessments were conducted primarily to ensure that federal funds were used appropriately, but they had the added benefit of teaching grantees — many of whom had never received federal funding — how to understand and comply with the rigorous requirements of their award.

When the Grant Progress Assessment program began, the assessors reported that they were as much educators as evaluators. During initial site visits, they often found records in disarray, and subsequently they spent a large portion of the visit showing grantees how to set up procedures for documenting, organizing and reporting the data required for grant compliance. As one NIJ staffer explained, "You get the quality you inspect, not the quality you expect."

Common findings from the early years of the Grant Progress Assessment program included the following:

- Incorrect information on the grant and the financial point of contact
- Late progress and financial reports
- Purchases that had been made outside of the proposed timeline

Resources for Managing Laboratory Grants

A number of resources are available to help laboratories continue to comply with the terms of their grant and ensure that public funds are spent in a fiscally responsible manner. These resources are described below:

For active grantees:

- *Grant Management System User Guide:* Provides step-by-step instructions and screenshots to help grantees manage their funds in compliance with federal requirements and deadlines. Visit http://www.ojp.usdoj.gov/training/pdfs/gmsissues.pdf.
- Grants Financial Management Online Training Module: Offers 24 training modules on the basics of federal grants management, including financial management systems, administrative rules, subawards, reporting requirements, financial monitoring and audit requirements, and award closeout. Visit http://gfm.webfirst.com.

For the general public:

- *Office of Justice Programs Financial Guide:* The primary reference manual for award recipients. Contains compilations of laws, rules and regulations that affect the financial and administrative management of awards. Visit http://www.ojp.usdoj.gov/financialguide/index.htm.
- NIJ Solicitations: Contains information about applying for and managing funding from NIJ. Visit http://www.nij.gov/funding/welcome.htm.

- Unallowable expenditures
- Lack of financial control when it came to commingling of funds
- Expenditures from a category not in the approved budget or made without prior approval
- Activity in the progress reports that was not consistent with the proposal

(For more on assessment findings, see sidebar, "Typical Assessment Findings," on this page.)

By the second or third year that they visited laboratories, assessors began seeing a dramatic improvement in reporting and compliance. Grantees were more organized; they had proper systems in place. They knew what questions the assessors were going to ask — for instance, about the validity of equipment purchases or whether agency asset numbers and serial numbers matched — and were ready with the answers and supporting documentation. As the program advanced, its impact became even more apparent — adverse findings became fewer and fewer and compliance became the norm.

Typical Assessment Findings

These are actual findings discovered and reported by Grant Progress Assessment assessors and reported to NIJ:

- Grant funds were being used to support activity that did not reflect the goals outlined in the proposal.
- No progress had been made in producing deliverables 18 months into the award.
- Grant-funded equipment was not being used.
- The grantee's records did not separate cases worked with grant funds and those worked with regular salary funds.
- The maximum daily rate provided to contractors exceeded the maximum allowable daily rate of $450.
- The number of cases reviewed listed in progress reports counted the same cases for different project periods.
- The grantee used funds for a conference unrelated to DNA backlog analysis.
- Funds in the consultants/contract budget category were encumbered prior to the award start date.
- The grantee issued a sole source purchase order before receiving approval.
- The grantee used funds from the consultant category, which had a zero budget.

Lessons Learned

Perhaps the biggest lesson learned by participating grantees was: Be organized!

- Keep copies of application documents.
- Keep receipts and documents related to purchases, including why the purchase was needed.
- Organize files.
- Follow instructions for reporting performance metrics.
- Perform frequent self-checks to ensure that procedures and activities continue to comply with NIJ's rigorous guidelines.

As for NIJ, a significant discovery from the Grant Progress Assessment program was that what was often thought by assessors to be a commonplace answer to a question was not always the commonplace answer in the mind of the grantee. During site visits, assessors played a critical role in bridging the gap in perceptions between the two sides.

The formalized review process introduced by the Grant Progress Assessment program helped both NIJ and grant recipients achieve their goals of ensuring that grants were being used to achieve the goals and objectives set out by Congress: Improve the capacity of crime laboratories to solve crimes.

About the authors: Patricia Kashtan is a program operations specialist in NIJ's Office of Investigative and Forensic Sciences. Jolene Hernon is Director of NIJ's Office of Communications. Beth Pearsall is the managing editor of the *NIJ Journal.*

NCJ 240698

For more information:

- Learn more about how the Grant Progress Assessment program worked at http://www.nij.gov/topics/forensics/lab-operations/capacity/grant-progress-assessments/welcome.htm.

Kristen M. Zgoba Wins the Peter P. Lejins Research Award

by Marilyn C. Moses

NIJ congratulates Kristen M. Zgoba, winner of the 2013 Peter P. Lejins Research Award, for her steadfast commitment to criminal justice research.

The Peter P. Lejins Research Award is the highest honor bestowed upon a corrections researcher. "The Lejins Award is given to an individual who has produced significant research for the correctional community and has demonstrated personal commitment and contribution to improve the profession of corrections," explained James Gondles, Executive Director of the American Corrections Association, which bestows the award.

Kristen M. Zgoba received her first NIJ research grant in 2006, just two years after earning her doctorate from Rutgers University. With the grant, she established prevalence rates of sex offenses before and after Megan's Law and compared recidivism rates among sex offenders who were subject to the law with rates for those who were not. This seminal work has become a benchmark for future research in the field. (To read more about the study, see https://www.ncjrs.gov/pdffiles1/nij/225402.pdf.)

Since that first NIJ award, Zgoba has contributed an impressive body of knowledge about sex offenses and offenders, including characteristics of sex offenders, predictive validity of risk assessment tools (such as the Static–99R), recidivism trajectories, and collateral consequences of sex offender notification and residency restriction laws.

Zgoba is currently supervisor of research and evaluation at the New Jersey Department of Corrections, where she focuses on sex offender management, offender recidivism and correctional health care, among other things. She developed a five-year research initiative and grant framework for her department and served as the statistical liaison for the "Another [Second] Chance" Initiative under former Governor Corzine's Crime Plan. She also is an adjunct professor at Rutgers.

"Her skills are top-notch, her grasp of the importance and uses of research in practical ways is unsurpassed, and her contributions are some of the most important in corrections departments in the nation," said Richard Tewksbury, 2006 recipient of the Lejins Award. "Dr. Zgoba has earned this distinction."

About the author: Marilyn C. Moses is a social science analyst in NIJ's Justice Systems Research Division.

NCJ 240699

Check out NIJ's Web page on sex offender research: http://nij.gov/nij/topics/corrections/community/sex-offenders/welcome.htm.

Listen to Zgoba and fellow panelists discuss "Sex Offenders in the Community: Post-Release, Registration, Notification and Residency Restrictions" at the 2010 NIJ Conference: http://nij.ncjrs.gov/multimedia/audio-nijconf2010-sex-offenders.htm.

Photo Credit: Christopher Onstott

Sex Offenders Monitored by GPS Found to Commit Fewer Crimes

by Philip Bulman

An NIJ-sponsored research project examines the impact that GPS monitoring has on the recidivism rates of sex offenders in California.

A study of California high-risk sex offenders on parole found that those placed on GPS monitoring had significantly lower recidivism rates than those who received traditional supervision.

Researchers examined the effectiveness of using GPS to monitor high-risk sex offenders placed on parole in California (see sidebar, "Using GPS to Monitor Sex Offenders," on page 23). The NIJ-sponsored study included 516 high-risk parolees who had been released from prison between January 2006 and March 2009. Half of the parolees wore GPS monitoring devices in addition to receiving traditional parole supervision, which involves regular contact by parole agents and weekly sex-offender treatment classes ("GPS group"); the other half received only traditional parole supervision ("traditional group"). Researchers tracked each parolee for one year following his initial parole date.

The study involved:

- An outcome evaluation to assess both the cost of the GPS program and its effectiveness in reducing the criminal behavior of high-risk sex offender parolees.
- A process evaluation to assess the program's design and implementation.

The researchers collected information from the state's data management system and examined official arrest records, parole supervision records, GPS monitoring data and state cost information. In addition, they conducted a survey of roughly 1,000 California Department of Corrections and Rehabilitation (CDCR) parole officers. The survey included questions about the GPS monitoring system, caseloads, program staffing and screening of high-risk sex offender parolees.

GPS More Expensive, but Also More Effective

The researchers found that parolees in the traditional group — those not placed on GPS monitoring — committed new crimes and had their parole revoked more often than did parolees in the GPS group. In addition, the traditional group returned to custody at a rate 38 percent higher than the GPS group.

The cost analysis showed that in California, monitoring parolees using GPS costs approximately $35.96 a day per person, while the cost of traditional supervision is about $27.45 a day. The GPS program is more expensive but more effective. Although the GPS program costs $8.51 more per day than traditional supervision, the GPS approach produced a decrease of 12 percentage points in arrests for any offense (from approximately 26 percent to 14 percent). In addition, offenders who were monitored using GPS complied with the terms of their parole at higher rates than did offenders on traditional parole.

The cost of California's GPS monitoring is lower than the cost of moving parolees to "indefinite civil commitment," which entails sending sex offenders whose prison sentences are over, but who are believed to be too dangerous to release into the community, directly from confinement in prison to confinement in dedicated institutions. Such civil confinement programs can cost an average of more than $100,000 a year per person because of the programming that must be provided.

Recommendations

The researchers also examined the degree to which the GPS program delivered services as designed. Through a process evaluation that looked at responses from the survey of parole agents as well as GPS monitoring data, they found that CDCR had developed a protocol for the GPS program and largely followed that protocol while implementing the program. Based on their findings, the researchers made several recommendations.

Reexamine the identification of high-risk sex offenders. To identify high-risk populations, California currently uses the standardized Static-99 risk instrument, which measures "static" factors that do not change over time (see related article, "Predicting Recidivism Risk: New Tool in Philadelphia Shows Great Promise," on page 4). However, in the survey of parole agents, nearly half of respondents said that the Static-99 does a poor job of identifying high-risk sex offenders. The researchers noted that the current risk instrument may predict

Using GPS to Monitor Sex Offenders

GPS monitoring uses satellites to calculate an offender's physical position. The offender wears a tamper-resistant bracelet — typically worn around the ankle — that receives transmissions from the satellites and calculates the offender's location. In "passive" monitoring systems, this information is stored and transmitted at appointed times to a monitoring station. In "active" systems, information on the individual's location transmits to a monitoring station in near real time, allowing the station to alert officers immediately when a violation occurs. Both systems allow exclusion zones (such as schools or other places where children congregate) or inclusion zones (such as a workplace) and provide information on when and where an individual has been throughout the day.

In California, sex offenders designated as high-risk are placed on actively GPS-monitored caseloads, while non-high-risk sex offenders are on passively GPS-monitored caseloads. However, in the state, information in both caseload types is received at near-real-time intervals. The difference is that information in the active system is reviewed more frequently than information in the passive system. Vendor-operated monitoring centers track this information and email daily reports to parole agents that detail all of the activity recorded by the GPS device. The centers also send an immediate alert notification to agents via text message whenever the GPS device records an inclusion/exclusion zone violation, tampering with the strap, a low battery, a cell communication gap or no GPS communication.

recidivism, but those convicted of noncontact offenses such as exhibitionism pose less of a threat than do rapists and child molesters. Thus, the researchers recommend using a system that accounts for the different recidivism risks among offenders and the varying threats to public safety.

Monitor attendance at treatment classes. CDCR mandates that high-risk sex offender parolees attend weekly treatment classes. However, the researchers found a disconnect between parole agents and service providers in terms of tracking treatment attendance; 100 parolees had no record of attending treatment during the study period. Further, in the survey of parole agents, only 75 percent of agents said that their parolees attended treatment at least once a week. Previous research indicates that the meticulous monitoring of sex offender treatment is an important facet of sex offender supervision and that sex offenders who stop attending treatment have higher recidivism rates. The researchers recommend that parole authorities strictly monitor and enforce weekly class attendance.

Use graduated sanctions that balance cost and risk. Instantly sending someone back to prison for a minor violation is costly. GPS supervision costs $35.96 daily, whereas the cost of keeping someone in a California prison is about $129 per day. The researchers recommend that CDCR — rather than issuing blanket parole revocations and sentencing violators to go back to prison for a few months at a time — employ a graduated sanctions system for dealing with parole violations. Such a system weighs the gravity of the offense against the need to preserve public safety, thereby increasing the likelihood that a parolee with a serious violation is incarcerated, while one who presents less danger is still sanctioned but in a less restrictive, less costly manner (for example, by imposing a home curfew on the offender). The researchers added that California is in the process of piloting a new, structured decision-making system for dealing with parole violations, which will allow parole agents to scientifically weigh an offender's risk level and the benefits of alternatives to prison as part of their decision-making process.

Mandate the use of zones. The researchers also found that parole agents were neglecting to use inclusion and exclusion zones. Such zones are intended to keep parolees either within certain areas, such as home and work, or away from certain

History of GPS Monitoring Policies in California

The California Department of Corrections and Rehabilitation (CDCR) began using GPS to supervise sex offenders in the community in 2005, when it began a two-year pilot program involving 500 GPS devices. The goals of the pilot were to give corrections officials experience with GPS monitoring and to resolve as many implementation issues as possible before expanding the program.

However, the passage of California Proposition 83 — better known as Jessica's Law — in November 2006 forced the department to quickly expand its efforts; Jessica's Law required that all sex offenders be placed on GPS monitoring for life. The law also:

- Forbade sex offenders from living within 2,000 feet of any school or park where children congregate.
- Increased sentences for some sex crimes, including life sentences for some offenses against children.
- Changed the criteria for sexually violent predators, increasing the number of sex offenders who are eligible for a civil commitment for treatment instead of being released on parole.
- Made CDCR parole officers responsible for enforcing the terms and conditions of Jessica's Law while a parolee is under the state's supervision.

By April 2008, CDCR had equipped its high-risk sex offender population of 2,500 with ankle monitors. By the end of 2008, the department had fully implemented the program by equipping 2,300 non-high-risk offenders with monitors, bringing the total to 4,800. This total was nearly three times as many as that in Florida, which has the second-largest use of GPS units.

As of August 2011, almost 10,000 sex offenders were on parole in California. About 7,000 of them were living in the community, with roughly 99 percent being monitored by GPS technology.

places, such as schools or parks that attract many children. In the process evaluation, the researchers found that only 60 percent of parole officers always or often discussed the limits of inclusion zones, and only half discussed exclusion zone limits. The researchers argue that the use of zones may be the most important GPS tool because the application of zones allows parole officers to be alerted to specific offender movements. Thus, they recommend making the zones compulsory.

Use a monitoring center to screen alerts. The large majority of parole agents (89 percent) reported in the survey that GPS monitoring was more time-intensive than traditional supervision. Until fairly recently, officers were receiving alerts when offenders tampered with the GPS device or committed other detectable violations. These alerts might also have included incidents such as an offender being in the basement of a building, sounding an "alarm" simply because someone was out of reach of the GPS monitoring system for a few minutes. From January 2009 until December 2010, paroled California sex offenders generated 1.5 million alert notifications. The researchers noted that according to an internal CDCR document, officers spend 44 percent of their time monitoring movements by GPS and only 12 percent of their time in the field.

To help remove the burden on agents of responding to "minor" alerts, California switched to a centralized monitoring system in 2011. Under the new system, two vendor-operated centers screen the thousands of GPS alerts that agents receive each month and respond to the more technical alerts, such as a battery that has run too low. The centers forward alerts that are more serious to parole officers, allowing officers to focus more closely on direct supervision and on responding to real threats to community safety.

Limit caseload to 20. GPS increases the information that officers receive about parolees, but reviewing this information is time-consuming and reduces the time available for direct supervision. As noted, agents spend only about 12 percent of their time in the field. According to the researchers, the best way to ensure that parole agents have sufficient time to directly supervise offenders is to limit the caseloads of GPS parole agents. In fact, the researchers found in their outcome evaluation that the size of the caseload was correlated with parole violations and with parolees returning to custody. Therefore, the researchers recommend smaller caseloads of no more than 20 people per officer.

Issues and Concerns

Finally, the researchers noted that GPS monitoring is not a panacea. The systems can give false positives for violations. For example, sometimes a monitored offender "disappears" simply because he is in an underground location, forgets to recharge the battery that powers the system, or even decides to go to sleep under an electric blanket that disrupts the GPS signal. In these instances, the system would send an alert even though no criminal activity was taking place.

The researchers also pointed out other possible limitations of their work. The study lasted only a year, and results may vary over longer periods. In addition, at least one previous study found that once the GPS monitoring ends, offenders who had been monitored by GPS do just as poorly as other offenders.

Researchers at Development Services Group, Inc., of Bethesda, Md., and Old Dominion University in Norfolk, Va., conducted the study of California parolees. Read their final report to NIJ, *Monitoring High-Risk Sex Offenders With GPS Technology: An Evaluation of the California Supervision Program,* available at https://www.ncjrs.gov/pdffiles1/nij/grants/238481.pdf.

About the author: Philip Bulman is a writer and editor at NIJ.

NCJ 240700

For more information:

- To learn more about the Offender Tracking Standard under development at NIJ, visit http://www.nij.gov/nij/topics/corrections/community/monitoring-technologies/tracking-standards.htm.

Watch an interview with George Drake, community corrections program manager at the Corrections Technology Center of Excellence, on making community supervision safer through electronic monitoring: http://nij.ncjrs.gov/multimedia/video-nijconf2011-drake.htm.

Ending Modern-Day Slavery: Using Research to Inform U.S. Anti-Human Trafficking Efforts

by Maureen Q. McGough

NIJ study examines the challenges facing the criminal justice system when combating human trafficking.

Trafficking in persons is modern-day slavery and exists in virtually every country in the world — and the United States is no exception.[1] Almost 150 years after the 13th Amendment abolished slavery and involuntary servitude, there are still men, women and children enslaved into labor and commercial sexual exploitation in the U.S. (see sidebar, "Understanding Modern-Day Slavery," on page 27).

In recent years the worldwide human trafficking problem has attracted significant political and social attention. Awareness-raising initiatives such as the United Nation's Blue Heart Campaign[2] encourage involvement and action to fight human trafficking on a global scale. In the U.S., the Department of Homeland Security's Blue Campaign unites anti-human trafficking programs and offers resources for law enforcement and the public to help raise awareness and provide much-needed training.[3]

Despite growing awareness of the issue and an influx of resources from such influential bodies as the United Nations and other intergovernmental organizations, foundations, non-governmental organizations and the U.S. government, the field is still hampered by its inability to measure the size and scope of trafficking.[4]

The data used to estimate the prevalence of human trafficking in the U.S. are lacking in scope and quality at the federal, state and local levels.[5] The lack of reliable data and a dependence on inadequate evidence have fueled disagreement among anti-human trafficking movements in this country, and some researchers have criticized the issue as unsubstantiated and estimates of the problem as dubious.[6] Recent estimates of people trafficked into the U.S. each year, for example, have varied widely from a low of approximately 14,500 to a high of approximately 50,000.[7]

Unfortunately, challenges also exist in gauging the effectiveness of the criminal justice system's response. Rates of identification, investigation and prosecution are of limited value in determining the effectiveness of U.S. responses to human trafficking because the data supporting prevalence estimates are unreliable.

Research can play an invaluable role in understanding the criminal justice system's ability to respond to trafficking and in identifying obstacles that hinder current efforts. The need for robust research is all the more pressing given restricted budgets and declining resources. At a time when governments increasingly are looking to use evidence-based practices, policymakers and practitioners are looking to the research community to produce the data needed to analyze the impact of anti-trafficking efforts.

The problem can be cyclical — without accurate estimates of the prevalence of human trafficking, it can be difficult to know how to allocate resources to study the issue. The U.S. State Department's annual compendium of countries' anti-human trafficking efforts, the *Trafficking in Persons Report,* recognizes this data deficiency and recommends that the U.S. improve the data and analysis of human trafficking cases at the state and local level.[8] NIJ has funded a number of projects to improve data collection and analysis of the issue. This article discusses one recent study that looked specifically at the challenges facing state and local criminal justice systems.

> "[The stories of human trafficking victims] remind us what kind of inhumane treatment we are capable of as human beings. They are living, breathing reminders that the war against slavery remains unfinished."
>
> —U.S. Secretary of State Hillary Clinton, at the release of the U.S. State Department's *2012 Trafficking in Persons Report*, June 19, 2012

A Hidden Crime

The nature of human trafficking helps keep this crime hidden. Captors often closely guard their victims, leaving them isolated with little to no freedom of movement.[9] They restrict victims' contact with the outside world.[10] Domestic servants remain "invisible" in private homes, and private businesses can serve as fronts for trafficking operations.[11] Many victims face language barriers that prevent them from seeking help. Additionally, international victims who enter the U.S. may be uncertain of their immigration status and thus less inclined to work with authorities.

Understanding Modern-Day Slavery

Most countries banned "chattel slavery" — one person owning another person as property — in the 1800s. Despite this, slavery continues in the modern day. Although owning slaves used to be a major investment formalized through legal documents, today's slaves are held through debt bondage, indentured servitude or other forms of control.

For more than a decade, the phrase "human trafficking" has been used to describe the act of holding a person in forced service — the very definition of slavery. The term can cause confusion, however, because it implies that traffickers always transport victims across borders; in actuality, victims can also be held in their own homes. Experts maintain that when considering the issue of human trafficking, it is important to do so in an accurate context — acknowledging that trafficking is modern slavery and that trafficked persons are slaves.

Furthermore, victims, for a variety of reasons, do not always identify themselves as such. Human trafficking victims suffer tragic psychological trauma and may experience Stockholm syndrome, generating positive feelings and gratitude toward their captors for perceived favors or even for being allowed to live.[12] Law enforcement commonly lacks training to identify these signs of trauma, making it difficult for them to sever the controlling bond that captors have over their victims and decreasing the likelihood that victims will cooperate.[13] Even if victims identify themselves as such and are aware of their rights, they still might hesitate to report their victimization out of fear of reprisal from the trafficker, lack of trust in law enforcement or fear of deportation.[14]

"[Human trafficking] ought to concern every person because it is a debasement of our common humanity. It ought to concern every community because it tears at our social fabric. It ought to concern every business because it distorts markets. It ought to concern every nation because it endangers public health and fuels violence and organized crime."

—President Barack Obama,
remarks at the Clinton Global Initiative
Annual Meeting, September 25, 2012

Challenges at the State and Local Levels

Since the passage of the Trafficking Victims Protection Act in 2000, 49 states have enacted legislation that criminalizes human trafficking and empowers state and local law enforcement — often the first responders to interact with victims — to investigate these cases without depending on federal authorities and to prosecute human trafficking cases in state courts.[15]

Increased involvement of state and local law enforcement is critical because they handle the bulk of criminal cases in the United States. Even before the passage of state anti-trafficking legislation, federal law enforcement requested that state and local officers "be the eyes and ears for [federal law enforcement in] recognizing, uncovering and responding to circumstances that may appear to be routine street crime, but may ultimately turn out to be a human trafficking case."[16] In fact, in a survey of state and local law enforcement personnel, 32 percent of respondents indicated that they identified many of their human trafficking cases when they were investigating other crimes.[17]

Despite this increased involvement, reports show that fewer trafficking cases have been identified and prosecuted than would be expected given current estimates.[18] This has led to speculation that either incidents of human trafficking are significantly overestimated or government officials and law enforcement agencies are not effectively confronting the problem.[19]

NIJ funded Amy Farrell and her colleagues at Northeastern University and researchers at the Urban Institute to examine the challenges facing state and local criminal justice systems when investigating and prosecuting human trafficking cases. The researchers conducted a 12-site study that included in-depth interviews with 166 practitioners from federal, state and local law enforcement; state and federal prosecutors; victim service providers; and other stakeholders. The researchers also analyzed data from 140 closed human trafficking case files[20] to determine which characteristics of human trafficking cases attract local law enforcement's attention and predict adjudicatory outcomes.[21] Although the study is not nationally representative, the findings can help us understand why the number of human trafficking cases is lower than estimates of the problem might predict. Here is what the researchers found:

Identification challenges

The study confirmed that identifying victims is particularly challenging because perpetrators hide and move their victims. The interviews also revealed that the cultural and

The Prevalence of Labor Trafficking in the United States

In the NIJ-funded study discussed in the main article, researchers found that the majority of cases identified by law enforcement involved sex trafficking. Only 11 percent of cases were labor trafficking cases; cases with both labor and sex trafficking made up an additional 4 percent. Federal and state data indicate that more investigations and prosecutions take place for sex trafficking than for labor trafficking.

Notably, however, in the U.S. State Department's *2012 Trafficking in Persons Report,* victim services providers in the U.S. reported assisting significantly more foreign-national labor trafficking victims than sex trafficking victims.[1] Concurrently, the Department of Health and Human Services has seen a steady rise in labor trafficking victims, and nongovernmental organizations have reported increasing instances of traveling sales crews and peddling rings using child and adult forced labor in the U.S.[2]

Some research suggests that labor trafficking victims are harder to identify than sex trafficking victims, given that international victims may be mistaken for smuggled immigrants. Further, the victimization of labor trafficking victims (many of whom are male) may be seen as less compelling than that of sex trafficking victims (many of whom are young women). In the NIJ-funded study discussed in the main article, researchers found that police and prosecutors were commonly unfamiliar with labor laws and regulations and lacked the infrastructure to identify instances of labor trafficking in various workplace settings.

Empirical research follows the same tendency to focus on sex trafficking. In an NIJ-funded bibliography of research literature on human trafficking, researchers found that the majority of articles addressed sex trafficking. Indeed, out of 39 articles, only four dealt with trafficking for labor exploitation or domestic servitude.[3]

The lack of knowledge about the scope and scale of labor trafficking in this country is particularly concerning given U.S. agriculture's heavy reliance on migrant laborers. To shed light on the issue, NIJ funded a study of migrant laborers in San Diego County; the study used respondent-driven sampling to produce statistical estimates of labor trafficking in the area. The study found that labor trafficking victimization appeared to be rampant among unauthorized Spanish-speaking immigrant workers in the county, with an estimate that more than 30 percent of this target population were labor trafficking victims.[4]

If the numbers coming out of San Diego County are any indication of prevalence in other parts of the country, there is a significant, immediate need for a greater understanding of the scope, scale and methods of labor trafficking on a national level to support and inform critical anti-trafficking efforts. Accordingly, NIJ plans to focus forthcoming solicitations (dependent on funding availability) on the prevalence and methods of labor trafficking in the U.S.

For more information:

- See http://www.nij.gov/nij/topics/crime/human-trafficking/welcome.htm.

Notes

1. U.S. Department of State, *Trafficking in Persons Report,* Washington, D.C.: Author, 2012, p. 360, available at http://www.state.gov/j/tip/rls/tiprpt/2012.
2. Ibid.
3. Goździak, Elżbieta M., and Micah N. Bump, "Data and Research on Human Trafficking: Bibliography of Research-Based Literature," Final report to the National Institute of Justice, grant number 2007-VT-BX-K002, October 2008, NCJ 224392, p. 7, available at https://www.ncjrs.gov/pdffiles1/nij/grants/224392.pdf.
4. Zhang, Sheldon X., "Looking for a Hidden Population: Trafficking of Migrant Laborers in San Diego County," Final report to the National Institute of Justice, grant number 2009-IJ-CX-0011, November 2012, NCJ 240223, available at https://www.ncjrs.gov/pdffiles1/nij/grants/240223.pdf.

organizational characteristics of police agencies can hinder efforts to identify victims and that local law enforcement and communities generally do not make combating human trafficking a priority.[22] This often means fewer resources for training and staff for labor-intensive investigations.

The cases the researchers reviewed had primarily been identified through reactive approaches, illustrating that officers generally wait for victims to self-identify or for community tips about potential victimization to be received before they launch investigations. The researchers found that law enforcement uniformly lamented the lack of identified labor trafficking cases, suggesting that although officers believe labor trafficking is occurring in local communities, they have not received information about these cases. The use of proactive strategies to identify trafficking cases was uncommon and rarely involved cooperation between law enforcement and prosecutors.

Investigation challenges

Researchers in the Northeastern University/Urban Institute study reviewed 140 closed trafficking case files to determine whether the evidence supported a charge of human trafficking as outlined in the Trafficking Victims Protection Act and its reauthorizations.[23] The review found evidence of human trafficking in the majority of the cases. Among the indicators found were threatening to harm or actually physically or nonphysically harming the victim; demeaning and demoralizing the victim; dominating, intimidating and controlling the victim; and disorienting and depriving the victim of alternatives.[24] However, few suspects were actually charged with human trafficking offenses.[25] Even when they received specific human trafficking cases, prosecutors were more likely to prosecute using laws with which they, judges and juries were more familiar, such as promoting prostitution, kidnapping or fraud.

> Although we still need reliable estimates of nationwide prevalence, it is clear that human trafficking occurs on a large scale within U.S. borders.

Interviews with practitioners confirmed what other research has shown: Victims were reluctant to cooperate with investigations because they either feared retaliation from their trafficker or distrusted law enforcement. In some cases, because jurisdictions lacked specialized services for trafficking victims, such as secure housing, law enforcement officers arrested victims to keep them from returning to their traffickers or to help them feel safe from pimps and thereby encourage the victims' cooperation in the investigation. Although the arrest was often for the victim's protection, it essentially resulted in the victim being treated like a suspect. These victims may feel revictimized and experience the same negative emotions they experienced in the trafficking situation, thus adding to often pre-existing distrust of law enforcement.[26]

Specialized services for trafficking victims are all the more critical because victims' loyalty to their traffickers may stem, in part, from feelings that they have no practical alternatives to their current situation.

One of the law enforcement practitioners interviewed explained the potential impact that additional resources could have on combating trafficking by providing victims with a viable alternative:

> We have nothing to say, "Hey, I can put you up in … this place. And I can help you get an education. And I can help you get a job. And I can help you take care of your kids." You know, we don't have that. If I had that, man … we could stop prostitution.

Interviews with law enforcement officers also revealed that some officers had negative stereotypes about the people commonly found to be victims of human trafficking, especially those involved with prostitution and those with drug addictions. Some reported the use of derogatory terms for victims, and one officer said, "Victims are often unreliable, often addicted to drugs. It's probably easier to prosecute homicides because the victims are dead."

Law enforcement commonly lacked training on how to investigate human trafficking cases. The researchers found that in many study sites, vice investigators were conducting human trafficking investigations using standard vice investigation strategies geared toward drug and gambling crimes, further reducing the likelihood of a successful trafficking investigation.

Additionally, many trafficking cases are cross-jurisdictional, and agencies reported that cases often fell apart when they lacked the resources or institutional support to gather evidence or conduct interviews in other states. Importantly, officers indicated that they could not dedicate time and resources to investigate cases they felt would not result in prosecution.

Therefore, the reluctance of prosecutors to file charges in human trafficking cases created a negative feedback loop, in some instances diminishing an investigator's determination to identify and investigate these challenging cases.

Prosecution challenges

Because state statutes on human trafficking are relatively new, there is a lack of precedent and case law, and prosecutors operate with little or no guidance on prosecutorial techniques or other resources, such as sample jury instructions. Many prosecutors interviewed in the study by Farrell and colleagues said they were concerned about losing high-profile cases (and damaging their reputation), and so they prosecuted cases using a charge other than human trafficking. A local prosecutor said, "[T]hat was sort of the unwritten policy of the office: 'Why bother with this goofy human trafficking statute, just charge other crimes that you are more comfortable with and that you have used in the past.'" Prosecutors also reported that victims were often reluctant to testify or did not seem credible, and most state and local agencies lacked the institutional infrastructure (such as a specialized human trafficking unit) to support prosecution.

Although few of the cases studied resulted in suspects being charged for human trafficking offenses, offenders were held accountable in 69 percent of the cases, in that they were prosecuted for different offenses, such as rape, kidnapping or pandering. However, this poses an obvious problem in crime reporting. For example, the 2008 reauthorization of the Trafficking Victims Protection Act mandates that the FBI collect information about human trafficking offenses through the Uniform Crime Reporting (UCR) program. But unless state and local law enforcement routinely investigate human trafficking cases as such, crime data reported through the UCR will inevitably undercount instances of human trafficking. Underreporting could be more harmful than no reporting at all, particularly when agencies tie funding decisions to what the crime data show are the most prevalent problems.

The Road Ahead

Human trafficking is believed to be a growing crime, fueled by low risk and the potential for high monetary gain. Although we still need reliable estimates of nationwide prevalence, it is clear that human trafficking occurs on a large scale within U.S. borders.[27]

Identifying, understanding and combating this inherently covert form of modern-day slavery requires robust research to shed light on an otherwise obscure problem. To further this body of research, NIJ recently awarded grants to study:

- The underreporting of sex trafficking victims who are minors
- The role that gangs play in sex trafficking
- Effective counter-trafficking legislation, law enforcement processes and demand-reduction strategies
- Strategies to stabilize and integrate adult survivors of human trafficking in the U.S.

About the author: Maureen Q. McGough is an attorney and international research partnership specialist in NIJ's Office of Research Partnerships.

NCJ 240701

Listen to Amy Farrell discuss the study's findings at the 2012 NIJ Conference: http://nij.ncjrs.gov/multimedia/video-nijconf2012-farrell.htm.

For more information:

- Learn more about the study and NIJ's ongoing research on human trafficking: http://nij.gov/topics/crime/human-trafficking/welcome.htm.

Notes

1. See generally U.S. Department of State, *Trafficking in Persons Report*, Washington, D.C.: Author, 2012, available at http://www.state.gov/j/tip/rls/tiprpt/2012.
2. For more information on the Blue Heart Campaign, visit http://www.unodc.org/blueheart/en/about-us.html.
3. For more information on the Department of Homeland Security's Blue Campaign, visit http://www.dhs.gov/blue-campaign.
4. The Trafficking Victims Protection Act (TVPA) defines a severe form of human trafficking as, "A commercial sex act induced by force, fraud or coercion, or in which the person induced to perform such act has not attained 18 years of age; or the recruitment, harboring, transportation, provision or obtaining of a person for labor or services, through the use of force, fraud or coercion for the purpose of subjection to involuntary servitude, peonage, debt bondage or slavery" (TVPA, Section 103, 8a and b).
5. Farrell, Amy, Jack McDevitt, Noam Perry, Stephanie Fahy, Kate Chamberlain, William Adams, Colleen Owens, Meredith Dank, Michael Shively, Ryan Kling, and Kristin Wheeler, "Review of Existing Estimates of Victims of Human

Trafficking in the United States and Recommendations for Improving Research and Measurement of Human Trafficking," Final report to ATEST, 2010, pg. vii and 9.

6. Weitzer, Ronald, "The Social Construction of Sex Trafficking: Ideology and Institutionalization of a Moral Crusade," *Politics & Society* 35 (2010): 447-448, available at http://myweb.dal.ca/mgoodyea/Documents/Migration%20studies/The%20social%20construction%20of%20sex%20trafficking%20Weitzer%20Pol%20Soc%202007%20%2035(3)%20%20447-475.pdf.

7. Clawson, Heather J., Nicole Dutch, Amy Solomon, and Lisa Goldblatt Grace, *Human Trafficking Into and Within the United States: A Review of the Literature,* Washington, D.C.: U.S. Department of Health and Human Services, 2009, available at http://aspe.hhs.gov/hsp/07/HumanTrafficking/LitRev.

8. *Trafficking in Persons Report,* 12.

9. Clawson, Heather J., and Nicole Dutch, *Identifying Victims of Human Trafficking: Inherent Challenges and Promising Strategies from the Field,* Washington, D.C.: U.S. Department of Health and Human Services, 2008, available at http://aspe.hhs.gov/hsp/07/humantrafficking/IdentVict/ib.htm.

10. Ibid.

11. Clawson, Heather, Mary Layne, and Kevonne Small, "Estimating Human Trafficking into the United States: Development of a Methodology," Final report to the National Institute of Justice, task order number 2004T0178, September 2006, NCJ 215475, available at https://www.ncjrs.gov/pdffiles1/nij/grants/215475.pdf.

12. See Office of Refugee Resettlement, "Fact Sheet: Labor Trafficking," U.S. Department of Health and Human Services, 2012, available at http://www.acf.hhs.gov/programs/orr/resource/fact-sheet-labor-trafficking-english, and Fitzgerald, Laura, "Stockholm Syndrome," *Time Magazine* (August 2009), available at http://www.time.com/time/magazine/article/0,9171,1920301,00.html.

13. Walker-Rodriguez, Amanda, and Rodney Hill, "Human Sex Trafficking," *Law Enforcement Bulletin* (March 2011), available at http://www.fbi.gov/stats-services/publications/law-enforcement-bulletin/march_2011/human_sex_trafficking.

14. Newton, Phyllis J., Timothy M. Mulcahy, and Susan E. Martin, "Finding Victims of Human Trafficking," Final report to the National Institute of Justice, grant number 2007-VT-BX-0001, October 2008, NCJ 224393, available at https://www.ncjrs.gov/pdffiles1/nij/grants/224393.pdf. See also Office of Refugee Resettlement, *Fact Sheet: Identifying Victims of Human Trafficking,* U.S. Department of Health and Human Services, 2012, available at http://www.acf.hhs.gov/programs/orr/resource/fact-sheet-identifying-victims-of-human-trafficking.

15. Farrell, Amy, Jack McDevitt, Rebecca Pfeffer, Stephanie Fahy, Colleen Owens, Meredith Dank, and William Adams, "Identifying Challenges to Improve the Investigation and Prosecution of State and Local Human Trafficking Cases," Final report to the National Institute of Justice, grant number 2009-IJ-CX-0015, April 2012, NCJ 238795, available at https://www.ncjrs.gov/pdffiles1/nij/grants/238795.pdf.

16. Ibid., 3.

17. Clawson, Heather, Nicole Dutch, and Megan Cummings, "Law Enforcement Response to Human Trafficking and the Implications for Victims: Current Practices and Lessons Learned," Final report to the National Institute of Justice, grant number 2004-WG-BX-0088, December 2006, NCJ 216547, p. iv, available at https://www.ncjrs.gov/pdffiles1/nij/grants/216547.pdf.

18. Farrell 2012, 1.

19. Ibid.

20. A "human trafficking case file" included cases investigated as human trafficking and prosecuted locally as such, investigated as human trafficking but prosecuted locally as a different crime, locally investigated as trafficking but prosecuted federally, locally investigated as trafficking but never prosecuted, or prosecuted locally as trafficking but originally identified as a different crime.

21. Farrell 2012, 19.

22. Notably, a different survey found that between 73 percent and 77 percent of local law enforcement leaders think that human trafficking is rare or nonexistent in their community, further explaining the lack of prioritization of trafficking. Farrell, Amy, Jack McDevitt, and Stephanie Fahy, "Understanding and Improving Law Enforcement Responses to Human Trafficking: Final Report," Final report to the National Institute of Justice, grant number 2005-IJ-CX-0045, December 2008, NCJ 222752, available at https://www.ncjrs.gov/pdffiles1/nij/grants/222752.pdf.

23. The TVPA (and its reauthorizations) define a human trafficking victim as a person induced to perform labor or a commercial sex act through force, fraud or coercion.

24. Farrell 2012, 306.

25. Seven percent of the reviewed cases resulted in a sex trafficking charge, 9 percent in a sex trafficking of a minor charge, and 2 percent in a labor trafficking charge.

26. Farrell 2012, 121.

27. Hopper, Elizabeth, "Underidentification of Human Trafficking Victims in the United States," *Journal of Social Work Research and Evaluation* 5 (2004): 125-136. See generally Clawson, Layne and Small 2006; U.S. Department of State, *Trafficking in Persons Reports,* Washington, D.C.: Author, 2001-2012, available at http://www.state.gov/j/tip/rls/tiprpt/index.htm; and Zhang, Sheldon X., "Looking for a Hidden Population: Trafficking of Migrant Laborers in San Diego County," Final report to the National Institute of Justice, grant number 2009-IJ-CX-0011, November 2012, NCJ 240223, available at https://www.ncjrs.gov/pdffiles1/nij/grants/240223.pdf.

National Missing and Unidentified Persons System
NamUs

NamUs Receives Paul H. Chapman Award

The Foundation for Improvement of Justice, Inc., bestows the Paul H. Chapman Award on organizations whose accomplishments improve local, state and federal systems of justice in the U.S. Chuck Heurich, program manager in NIJ's Office of Investigative and Forensic Sciences, accepted the award on behalf of NIJ and the National Missing and Unidentified Persons System (NamUs) team in August 2012.

Launched in 2009, NamUs is the nation's first publicly searchable online repository of missing and unidentified decedent records. It offers databases to medical examiners, coroners, law enforcement officials and the general public.

The NamUs team was nominated by a woman whose childhood friend disappeared more than three decades ago. The woman's friend, Mary "Bobo" Shinn, a real estate agent, was called to show a house in Magnolia, Ark., on July 20, 1978. She has not been seen or heard from since — and her case is one of more than 16,000 cases currently in NamUs.

- To learn more about NamUs, visit http://www.namus.gov.
- Read about how NamUs has helped agencies solve cases at http://www.nij.gov/journals/264/solving.htm.

The Pitfalls of Prediction

by Greg Ridgeway

The criminal justice system should take advantage of the latest scientific developments to make reliable predictions.

Prediction is common in everyday life. We make predictions about the length of our morning commute, the direction of the stock market, and the outcomes of sporting events. Most of these common-sense predictions rely on cognitive shortcuts — or heuristics — that shape our expectations of what is likely to occur in the future. But these heuristics are not necessarily true; they rely on cognition, memory and sensory impressions rather than a balanced analysis of facts. Consequently, they can result in biased predictions.

The challenge of predicting the future has always been at the heart of the criminal justice system. Judges weigh the risks of releasing offenders to probation, police agencies try to anticipate where officers should be deployed to prevent future crime, and victims wrestle with the uncertain odds of being revictimized.

There is a long history of research on prediction in criminology and criminal justice, and two developments are helping the criminal justice system improve its ability to make reliable, scientific predictions. First, more and more jurisdictions are accumulating rich data and are getting better at linking across their data sources. Second, a growing set of sophisticated analytic prediction tools is available to help agencies make decisions about future events, unknown risks and likely outcomes.

Practitioners can now combine expert assessment with data-driven prediction models to discern how much risk a probationer poses,

Editor's Note: *This article was presented to seven law enforcement agencies that were developing predictive policing programs.*

determine whether a pair of illicit drug transactions signals the emergence of a drug market, or project whether crime will increase or decrease during the next month. More and more, police departments are using forecasting tools as a basis for formal predictive policing efforts; these statistical prediction methods inform their prevention strategies so they can anticipate rather than react to crime.[1] (See sidebar, "NIJ's Role in Predictive Policing," on this page.)

Although the science of prediction continues to improve, the work of making predictions in criminal justice is plagued by persistent shortcomings. Some stem from unfamiliarity with scientific strategies or an over-reliance on timeworn — but unreliable — prediction habits. If prediction in criminal justice is to take full advantage of the strength of these new tools, practitioners, analysts, researchers and others must avoid some commonplace mistakes and pitfalls in how they make predictions.

NIJ's Role in Predictive Policing

Law enforcement work is frequently reactive: Officers respond to calls for service, control disturbances and make arrests. But law enforcement work is becoming increasingly proactive: Departments combine data with street intelligence and crime analysis to understand why a problem arises and predict what might happen next if they take certain actions.

NIJ is supporting predictive policing efforts in a number of ways:

- **Predictive policing symposiums.** NIJ convened two symposiums at which researchers, practitioners and law enforcement leaders developed and discussed the concept of predictive policing and its impact on crime and justice. Read summaries of both sessions at http://www.nij.gov/topics/law-enforcement/strategies/predictive-policing/symposium/welcome.htm.
- **Predictive policing grants.** The Chicago and Shreveport police departments are using grants to explore data-driven policing strategies. In Phase 1, they received funding to identify a problem and develop predictive policing strategies to solve it. In Phase 2, they were awarded additional funding to implement and evaluate the strategies. For more on these grants, see http://www.nij.gov/nij/topics/law-enforcement/strategies/predictive-policing/symposium/discussion-demonstrations.htm.

For more information:

- To learn more about predictive policing in general, read the *NIJ Journal* article "Predictive Policing: The Future of Law Enforcement?" at http://www.nij.gov/journals/266/predictive.htm.

Pitfall #1: Trusting Expert Prediction Too Much

Using data and computers to predict or help experts predict shows promise, but the pace of adoption has not matched that promise. Why? Perhaps we trust ourselves more than we trust machines.

For example, more than 30 years ago, Stanford scientists developed a pathbreaking, computer-based medical expert system that could synthesize patient features and therapeutic options.[2] The system, called MYCIN, outperformed practitioners in selecting the right antibiotic treatments. Despite MYCIN's demonstrated success and similar kinds of computer-based prediction successes, we still do not see these systems being used in our doctors' offices. Some researchers have found that physicians have "a high regard for their own decision-making ability and are afraid of any competition from computers."[3]

So how do experts and machines compare in their ability to predict in the justice system?

Consider this example: A panel of 83 experts — law professors, deans of law schools and others who had practiced before or clerked at the U.S. Supreme Court — set out to predict how the U.S. Supreme Court would vote on the 68 upcoming cases on the 2002 docket. Based on their knowledge of the justices and the ins and outs of the court, they correctly predicted how the Supreme Court would vote on 59 percent of the cases.

Researchers used a computer program to make the same prediction. The computer analyzed 628 previous Supreme Court cases and generated data-derived rules.[4] The researchers created a decision-tree prediction model based on a simple set of these rules.

Figure 1. Decision Tree for Supreme Court Justice Sandra Day O'Connor

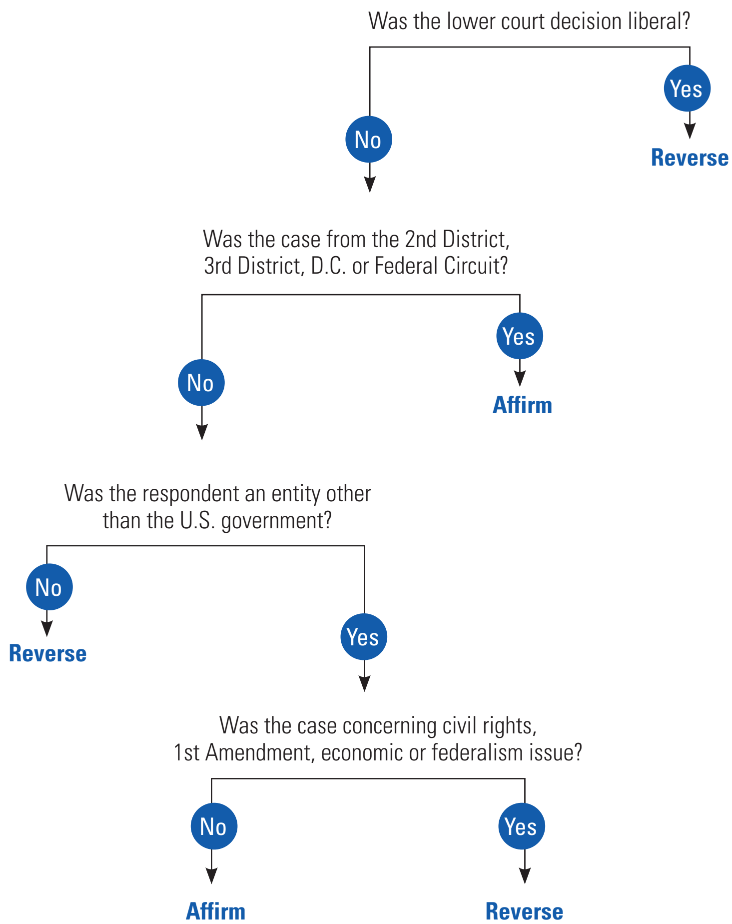


Figure 1 shows the decision tree for Justice Sandra Day O'Connor. Based on a simple set of rules — such as whether the lower court decision was liberal — the model was able to predict how Justice O'Connor would decide 70 percent of the cases in 2002. Using similar decision trees for the other eight justices,[5] the model correctly predicted the majority opinion in 75 percent of the cases, substantially outperforming the experts' 59 percent. The experts lost out to a machine that had a few basic facts about the cases.

So what can we take away from this example? It should lead us to question — but not necessarily dismiss — the predictions of experts, including ourselves. Of course, not all cases afford us the data to build predictive models. But if we have data that we can use to construct predictive models, then we should build the models and test them even if our expert detectives, probation officers and others in the field indicate that they already know how to predict. They may be as surprised as the expert panel was in the Supreme Court example.

Pitfall #2: Clinging to What You Learned in Statistics 101

If your knowledge of prediction is limited to what gets covered in introductory statistics courses, you are probably unfamiliar with the prediction model used above. Instead, you most likely learned how to check model assumptions and carefully test hypotheses. But when it comes to prediction, the rules are different and rather simple: Are the predictions accurate, and can you get them when you need them? You can judge the quality of a specific prediction model by considering the following:

Performance criteria. Do the model's goals and constraints match the intended use? Methods that are good at predicting, for example, whether an injury will result from a mission are not necessarily the same as those that are good at predicting the number of days an officer will be out with that injury. If you are planning a tactical unit's staffing, it is important for you to know the expected person-hours that will be lost to injuries. Thus, using a model that can accurately predict only whether an injury will occur — and not how long an officer will be out — would be insufficient.

Accuracy. Can the model make accurate forecasts? More specifically, the implemented model should be better at prediction than the agency's current practice. For example, if cops are allocating resources to neighborhoods where they think crime will spike, then going forward we should test whether the prediction model is better at selecting those neighborhoods. If a probation officer is assigning remote monitoring anklets to DUI probationers, then we should test whether the prediction model is better at picking which DUI probationers will reoffend in the next six months. For a prediction model to be useful, it must outperform practice as usual.

Computation time. Can you apply the prediction model in a reasonable amount of time? Some models can be computationally intensive to run and use. There is little point in using a model that cannot produce predictions in time for them to be useful.

Handling mixed data types. Can the prediction model manage and properly interpret numbers, dates and times, geography, text, and missing values — which datasets almost always have?

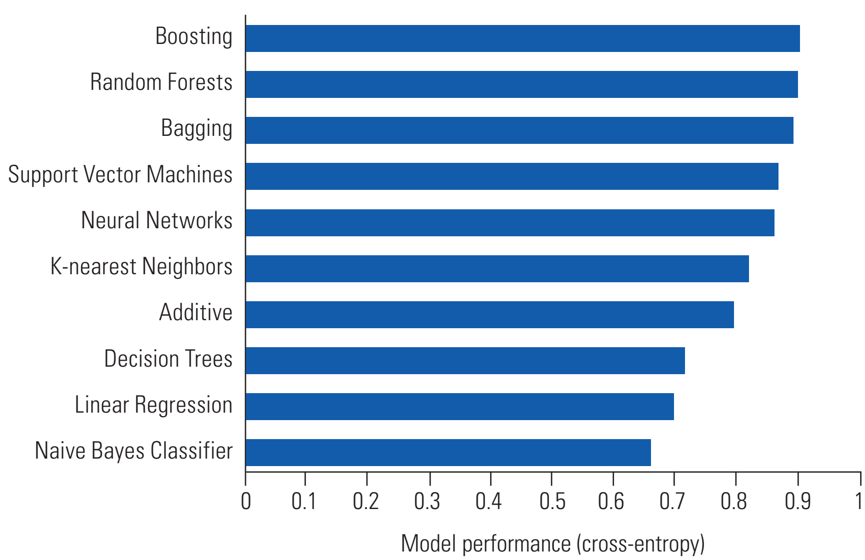
Interpretability. Can a person understand why the prediction model makes the predictions it does? We would prefer to be able to understand the reasoning behind a prediction. However, if getting transparency requires using a model that is *less* accurate in predicting, say, when and where a gang retaliation shooting will take place, then a more transparent model might not be worth the cost. This issue will be discussed further under Pitfall #5.

Pitfall #3: Assuming One Method Works Best for All Problems

In 2006, researchers examined how the most commonly used prediction methods performed head-to-head.[6] They looked at 11 datasets covering a variety of prediction tasks and measured each method's accuracy. The researchers found that the more modern methods of boosting and random forests consistently performed best, whereas linear regression — well over 70 years old and by far the most widely used method — did not fare well. (See Figure 2.) Note that decision trees, the method used in the Supreme Court example, is also near the bottom of the list, suggesting that even better accuracy in predicting case outcomes is possible. The University of Pennsylvania team working with Philadelphia's Adult Probation and Parole Department to predict probationers at high risk of violent crime opted for random forests. (See "Predicting Recidivism Risk: New Tool in Philadelphia Shows Great Promise" on page 4.)

Prediction can play a major role in the criminal justice system. Even small improvements in where police are assigned, which cold cases receive more attention, or which probationers receive more intense supervision can result in performance and efficiency gains.

Figure 2. Comparison of 10 Widely Used Prediction Methods



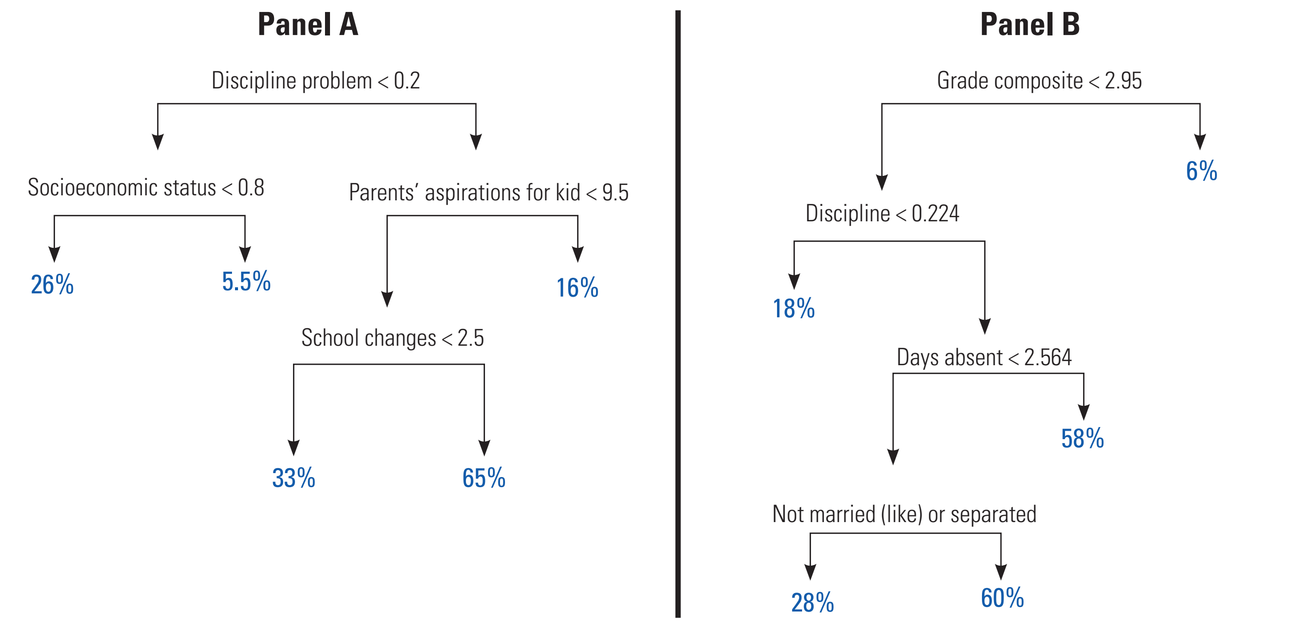
For more information on these prediction models, see Caruana, Rich, and Alexandru Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms Using Different Performance Metrics," in *Proceedings of the 23rd International Conference on Machine Learning,* New York: Association for Computing Machinery, 2006: 161-168.

However, the researchers who compared these prediction methods also found that the best-performing method for any particular dataset varied. This means that analysts cannot fall in love with a single model — depending on the particular prediction problem, their preferred method might not be the best fit.

Pitfall #4: Trying to Interpret Too Much

Practitioners tend to favor decision-tree models like the one used in the Supreme Court example because

Figure 3. Example of Decision-Tree Models Fit to Two Samples from NELS88 Dataset



they offer transparency. One can, after all, trace the pathways through the tree. And the Justice O'Connor tree, based on a set of simple rules, provides a compact, easy-to-follow story.

But not all trees are as straightforward — they can have many branches, the path may not be easy to follow, and the rules can be quite sensitive to small changes in the dataset.

For example, we can create a decision-tree model to predict student dropout risk among 16,000 students in the 1988 National Education Longitudinal Study (NELS88). If we randomly split the data on students into two halves, each with 8,000 students, and fit decision-tree models predicting dropout risk to each half, the resulting trees will look like those in Figure 3.

We arrive at very different interpretations about the reasons behind student dropout. Looking at the first tree (Panel A), we might conclude that discipline problems are the most important factor. When we look at the second tree (Panel B), it seems that grades are most important. Incidentally, the two decision trees had identical predictive accuracy.

The lesson is this: Although it is tempting to try to interpret results, the tree's structure is actually quite unstable. Instead, users should focus on the accuracy of the predictions. In some ways, this is analogous to using a watch — you expect it to give you the time accurately even if you do not completely understand how it works.

Pitfall #5: Forsaking Model Simplicity for Predictive Strength — or Vice Versa

Earlier, we noted that we would prefer to have a more interpretable model than a less interpretable one. Unfortunately, there is often a tradeoff, with more interpretability coming at the expense of more predictive capacity. But it is crucial that predictive models are designed for those who are going to use them, and in some cases, being able to interpret results is more important than achieving greater predictive capability.

Take, for instance, the Los Angeles Police Department's (LAPD's) effort to identify new recruits.[7] The LAPD did not know why some candidates made it through the recruiting process and became officers and others did not — and thus, it did not know whether it was using its resources efficiently.

To help the LAPD predict which recruits had a better chance of becoming officers, researchers developed a priority score based on a few easily collected facts about each candidate. The score rated how likely that candidate was to join the department. Recruiters could then usher these viable candidates through the process more quickly.

Looking at LAPD data on former recruits, the researchers found that three factors were critically important:

- Whether the candidates had "issues" identified in a preliminary background questionnaire that could disqualify them from service (e.g., criminal, financial, driving and drug history)
- Level of education
- Where they lived

They developed the point system in Table 1.

Table 1. Priority Score Point System

Preliminary Background	Does not qualify because has too many issues	Potentially qualifies, has some issues	Qualifies with no issues
	+0	+13	+22
Education	Has GED	Has high school	Has some college
	+0	+4	+8
Residence	Lives outside California	Lives in California	Lives in Los Angeles County
	+0	+2	+5

Under the system, if candidates have too many background issues, they do not qualify for service and receive 0 points. Most candidates have some issues but not enough to disqualify them, and they receive 13 points. Some have no issues and receive 22 points. The other two factors, education and residence, follow a similar point structure. Most candidates have high school diplomas (4 points), and most live in Los Angeles County (5 points). A candidate's priority score is the sum of the points for their preliminary background, education and residence.

The model predicts that a candidate who has a total of 22 points — for example, a recruit with some issues, a high school degree and residency in Los Angeles County — has a 20 percent chance of joining the LAPD. Candidates who have no issues and some college and live in Los Angeles County score the maximum 35 points; according to the model, they have a 43 percent chance of joining the department. By separating these highly viable candidates from the rest of the recruits, this system allows the LAPD to prioritize candidates and more efficiently allocate its recruiting resources. And because the model is simple to understand and simple to implement, the LAPD recruiting team used it.

This simplicity gets at the important issue: A decent transparent model that is actually used will outperform a sophisticated system that predicts better but sits on a shelf. If the researchers had created a model that predicted well but was more complicated, the LAPD likely would have ignored it, thus defeating the whole purpose.

Pitfall #6: Expecting Perfect Predictions

When using prediction models, managing expectations and focusing on the big picture are critical. Predictions will not be perfect, but the ultimate goal is to improve *overall* efficiency.

In the LAPD example, a highly viable candidate has a 43 percent chance of joining the force. This means that 57 percent of highly viable recruits drop out. Invariably, a candidate given a high viability score will fail miserably in the process. Because of this, some will say that doing business the old way is a better strategy than using the prediction model. But a handful of misclassified candidates should not overshadow the gains made in recruiting efficiency.

Predictive policing offers another example. It holds the promise of anticipating where crimes will occur so that police can prevent those crimes. However, prevention activities prompted by prediction models are poised to disappoint some. Consider a model that anticipates the time and place of the next retaliatory gang shooting almost perfectly, coupled with a model that directs officers to the right place at the right time. In such an ideal situation, the predicted shooting will never materialize. Naturally, those in the field will question why they were deployed to this place and not to another place with more pressing problems.

Pitfall #7: Failing to Consider the Unintended Consequences of Predictions

Prediction models can have unintended consequences that must be anticipated. Take, for instance, the LAPD recruiting example. The goal of the prediction model was to help the department improve its recruiting process. However, the LAPD is under a 30-year-old court order to meet diversity targets, such as having women make up 25 percent of new recruits. The priority point system could undermine the LAPD's ability to comply with the court order if, for example, prioritizing those with some college reduced the number of minority recruits.

The researchers who developed the point system considered this unintended consequence and noted that a small change could not only avoid the problem but also actually help the LAPD achieve compliance. They determined that if female applicants received an additional 7 points, then the system would be tuned so the department would meet its goal for recruiting women and its racial diversity goals because minority candidates were more likely to be women. Although this change reduces predictive accuracy because it places priority on some candidates with a lower chance of joining, it optimizes recruiting resources subject to the department's diversity goals.

The Power of Prediction

Prediction can play a major role in the criminal justice system. Even small improvements in where police are assigned, which cold cases receive more attention, or which probationers receive more intense supervision can result in performance and efficiency gains.

However, if the criminal justice system is going to reap such gains by using prediction models, it must seek to avoid the pitfalls that are so often a part of prediction.

About the author: Greg Ridgeway is acting director of NIJ.

NCJ 240702

Watch Greg Ridgeway talk about predictive policing at the 2010 NIJ Conference: http://nij.ncjrs.gov/multimedia/video-nijconf2010-ridgeway.htm.

Notes

1. Bratton, William J., and Sean W. Malinowski, "Police Performance Management in Practice: Taking COMPSTAT to the Next Level," *Policing* 2(3) (2008): 259-265.
2. Yu, Victor L., Lawrence M. Fagan, Sharon M. Wraith, William J. Clancey, A. Carlisle Scott, John Hannigan, Robert L. Blum, Bruce G. Buchanan, and Stanley N. Cohen, "Antimicrobial Selection by a Computer: A Blinded Evaluation by Infectious Disease Experts," *Journal of the American Medical Association* 242(12) (1979): 1279-1282.
3. Engle, Jr., Ralph L., and Betty J. Flehinger, "Why Expert Systems for Medical Diagnosis Are Not Being Generally Used: A Valedictory Opinion," *Bulletin of the New York Academy of Medicine* 63(2) (1987): 193-198. Heathfield is more optimistic but echoes that "theoretical and technical limitations are not the major barriers to the successful implementation … but rather more complex professional and organizational issues are at stake." Heathfield, Heather, "The Rise and 'Fall' of Expert Systems in Medicine," *Expert Systems* 16(3) (1999): 183-188.
4. Ruger, Theodore W., Pauline T. Kim, Andrew D. Martin, and Kevin M. Quinn, "The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking," *Columbia Law Review* 104 (2002): 1150-1210.
5. The scores for all the justices were as follows: O'Connor: 70 percent; Ginsburg: 55 percent; Breyer: 57 percent; Stevens: 61 percent; Souter: 61 percent; Kennedy: 71 percent; Scalia: 73 percent; Rehnquist: 76 percent; and Thomas: 75 percent.
6. Caruana, Rich, and Alexandru Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms Using Different Performance Metrics," in *Proceedings of the 23rd International Conference on Machine Learning,* New York: Association for Computing Machinery, 2006: 161-168.
7. Lim, Nelson, Carl F. Matthies, Greg Ridgeway, and Brian Gifford, *To Protect and to Serve: Enhancing the Efficiency of LAPD Recruiting.* Santa Monica, Calif.: RAND Corporation, 2009.

The National Institute of Justice is
the research, development and evaluation
agency of the U.S. Department of Justice.
NIJ's mission is to advance scientific research,
development and evaluation to enhance
the administration of justice
and public safety.

The National Institute of Justice is a component
of the Office of Justice Programs, which also
includes the Bureau of Justice Assistance;
the Bureau of Justice Statistics; the Community
Capacity Development Office; the Office for
Victims of Crime; the Office of Juvenile Justice
and Delinquency Prevention; and the Office of Sex
Offender Sentencing, Monitoring, Apprehending,
Registering, and Tracking (SMART).

Photo Sources: Christopher Onstott,
Mike Danko, Palladian Partners, Inc.,
iStock, ThinkStock and Veer.

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

Washington, DC 20531

Official Business
Penalty for Private Use $300

NCJ 240695

PRESORTED STANDARD
POSTAGE & FEES PAID
DOJ/NIJ
PERMIT NO. G-91

ISSUE NO. 271

NIJ