RAND

Safety and Justice Program

CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1 ▾

Support RAND

Purchase this document

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore the RAND Safety and Justice Program

View document details

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see RAND Permissions.

This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.

RAND Safety and Justice Program

PREDICTIVE POLICING

The Role of Crime Forecasting in Law Enforcement Operations

Walter L. Perry, Brian McInnis, Carter C. Price,
Susan C. Smith, John S. Hollywood

Supported by the National Institute of Justice

The research described in this report was sponsored by the National Institute of Justice and conducted in the Safety and Justice Program within RAND Justice, Infrastructure, and Environment.

This project was supported by Award No. 2010-IJ-CX-K007, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice.

Library of Congress Cataloging-in-Publication Data is available for this publication.

ISBN: 978-0-8330-8148-3

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND—make a tax-deductible charitable contribution at www.rand.org/giving/contribute.html

RAND® is a registered trademark.

© Copyright 2013 RAND Corporation

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see the RAND permissions page (www.rand.org/pubs/permissions.html).

RAND OFFICES
SANTA MONICA, CA • WASHINGTON, DC
PITTSBURGH, PA • NEW ORLEANS, LA • JACKSON, MS • BOSTON, MA
DOHA, QA • CAMBRIDGE, UK • BRUSSELS, BE
www.rand.org

Summary

Predictive policing is the application of analytical techniques—particularly quantitative techniques—to identify likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions. Several predictive policing methods are currently in use in law enforcement agencies across the United States, and much has been written about their effectiveness. Another term used to describe the use of analytic techniques to identify likely targets is *forecasting*. Although there is a difference between prediction and forecasting, for the purposes of this guide, we use them interchangeably.[1]

Objectives and Approach

Predictive methods allow police to work more proactively with limited resources. The objective of these methods is to develop effective strategies that will prevent crime or make investigation efforts more effective. However, it must be understood at all levels that applying predictive policing methods is not equivalent to finding a crystal ball. For a policing strategy to be considered effective, it must produce tangible results. The objective of this research was to develop a reference guide for departments interested in predictive policing, providing assessments of both the most promising technical tools for making predictions and the most promising tactical approaches for acting on them. More broadly, this guide is intended to put predictive policing in the context of other modern, proactive policing measures.

We approached this task in three ways:

1. We conducted a literature search of academic papers, vendor tool presentations, and recent presentations at conferences, drawing lessons from similar predictive

[1] The most common distinction is that forecasting is objective, scientific, and reproducible, whereas prediction is subjective, mostly intuitive, and nonreproducible. According to this distinction, the methods described in this report are essential forecasting methods. However, the law enforcement community has used *predictive policing* to describe these methods, so it is the term favored here.

techniques used in counterinsurgency and counter–improvised explosive devices operations and related research by the U.S. Department of Defense.

2. We reviewed a number of cases of departments using predictive policing techniques that appear promising.
3. We developed a taxonomy of the different types of operational applications that can be supported using predictive policing.

In many cases, we were able to illustrate how predictive technologies are being used to support police operations through a set of examples and case studies. Although some of the methods are promising and describe the current state of field, they are still more academic than practical. Consequently, this guide can also be viewed as a profile of the state of the art of predictive policing practices and the development of new predictive technologies. As such, it can be considered a baseline document.

A Taxonomy of Predictive Methods

In our assessment of predictive policing, we found that predictive methods can be divided into four broad categories:

1. *Methods for predicting crimes:* These are approaches used to forecast places and times with an increased risk of crime.
2. *Methods for predicting offenders:* These approaches identify individuals at risk of offending in the future.
3. *Methods for predicting perpetrators' identities:* These techniques are used to create profiles that accurately match likely offenders with specific past crimes.
4. *Methods for predicting victims of crimes:* Similar to those methods that focus on offenders, crime locations, and times of heightened risk, these approaches are used to identify groups or, in some cases, individuals who are likely to become victims of crime.

Tables S.1–S.4 summarize each category and show the range of approaches that law enforcement agencies have employed to predict crimes, offenders, perpetrators' identities, and victims, respectively. We found a near one-to-one correspondence between conventional crime analysis and investigative methods and the more recent "predictive analytics" methods that mathematically extend or automate the earlier methods. Conventional methods tend to be heuristic, or mathematically simple. As a result, they are low-cost and can work quite well, especially for analysts facing *low to moderate* data volumes and levels of complexity. In contrast, full-scale predictive analytics require sophisticated analysis methods that draw on *large* data sets. In this context, *large* refers to an amount of data beyond what a single analyst could recall without the assistance of a computer program or similar resources. Conversely, *low to moderate* refers to a data set that is sufficiently small that an analyst could reasonably recall its key facts.

Table S.1 summarizes predictive policing methods related to predicting crimes. As the table shows, conventional approaches start with mapping crime locations and determining (using human judgment) where crimes are concentrated ("hot spots"). The corresponding predictive analytics methods start, at the most basic level, with regression analyses and extend all the way to cutting-edge mathematical models that are the subjects of active research.

Table S.2 summarizes methods to identify individuals at high risk of offending in the future. The bulk of these methods relate to assessing individuals' risk. Here, conventional methods rely on clinical techniques that add up the number of risk factors to create an overall risk score. The corresponding predictive analytics methods use regression and classification models to associate the presence of risk factors with a percent chance that a person will offend. Also of interest are methods that identify criminal groups (especially gangs) that are likely to carry out violent assaults on each other in the near future. Hence, these methods can also be used to assess the risk that an individual will become a victim of crime.

Table S.3 summarizes methods used to identify likely perpetrators of past crimes. These approaches are essentially real-world versions of the board game Clue™: They use available information from crime scenes to link suspects to crimes, both directly and by processes of elimination. In conventional approaches, investigators and analysts

Table S.1
Law Enforcement Use of Predictive Technologies: Predicting Crimes

Problem	Conventional Crime Analysis (low to moderate data demand and complexity)	Predictive Analytics (large data demand and high complexity)
Identify areas at increased risk		
Using historical crime data	Crime mapping (hot spot identification)	Advanced hot spot identification models; risk terrain analysis
Using a range of additional data (e.g., 911 call records, economics)	Basic regression models created in a spreadsheet program	Regression, classification, and clustering models
Accounting for increased risk from a recent crime	Assumption of increased risk in areas immediately surrounding a recent crime	Near-repeat modeling
Determine when areas will be most at risk of crime	Graphing/mapping the frequency of crimes in a given area by time/date (or specific events)	Spatiotemporal analysis methods
Identify geographic features that increase the risk of crime	Finding locations with the greatest frequency of crime incidents and drawing inferences	Risk terrain analysis

Table S.2
Law Enforcement Use of Predictive Technologies: Predicting Offenders

Problem	Conventional Crime Analysis (low to moderate data demand and complexity)	Predictive Analytics (large data demand and high complexity)
Find a high risk of a violent outbreak between criminal groups	Manual review of incoming gang/ criminal intelligence reports	Near-repeat modeling (on recent intergroup violence)
Identify individuals who may become offenders:	Clinical instruments that summarize known risk factors	Regression and classification models using the risk factors
Probationers and parolees at greatest risk of reoffending		
Domestic violence cases with a high risk of injury or death		
Mental health patients at greatest risk of future criminal behavior or violence		

Table S.3
Law Enforcement Use of Predictive Technologies: Predicting Perpetrator Identities

Problem	Conventional Crime Analysis (low to moderate data demand and complexity)	Predictive Analytics (large data demand and high complexity)
Identify suspects using a victim's criminal history or other partial data (e.g., plate number)	Manually reviewing criminal intelligence reports and drawing inferences	Computer-assisted queries and analysis of intelligence and other databases
Determine which crimes are part of a series (i.e., most likely committed by the same perpetrator)	Crime linking (use a table to compare the attributes of crimes known to be in a series with other crimes)	Statistical modeling to perform crime linking
Find a perpetrator's most likely anchor point	Locating areas both near and between crimes in a series	Geographic profiling tools (to statistically infer most likely points)
Find suspects using sensor information around a crime scene (GPS tracking, license plate reader)	Manual requests and review of sensor data	Computer-assisted queries and analysis of sensor databases

do this largely by tracing these links manually, with assistance from simple database queries (usually, the names, criminal records, and other information known about the suspects). Predictive analytics automate the linking, matching available "clues" to potential (and not previously identified) suspects across very large data sets.

Table S.4 summarizes methods to identify groups—and, in some cases, individuals—who are likely to become victims of crime. These methods mirror those used to predict where and when crimes will occur, as well as some of the methods used to predict who is most likely to commit crimes. Predicting victims of crime requires

Table S.4
Law Enforcement Use of Predictive Technologies: Predicting Crime Victims

Problem	Conventional Crime Analysis (low to moderate data demand and complexity)	Predictive Analytics (large data demand and high complexity)
Identify groups likely to be victims of various types of crime (vulnerable populations)	Crime mapping (identifying crime type hot spots)	Advanced models to identify crime types by hot spot; risk terrain analysis
Identify people directly affected by at-risk locations	Manually graphing or mapping most frequent crime sites and identifying people most likely to be at these locations	Advanced crime-mapping tools to generate crime locations and identify workers, residents, and others who frequent these locations
Identify people at risk for victimization (e.g., people engaged in high-risk criminal behavior)	Review of criminal records of individuals known to be engaged in repeated criminal activity	Advanced data mining techniques used on local and other accessible crime databases to identify repeat offenders at risk
Identify people at risk of domestic violence	Manual review of domestic disturbance incidents; people involved in such incidents are, by definition, at risk	Computer-assisted database queries of multiple databases to identify domestic and other disturbances involving local residents when in other jurisdictions

identifying at-risk groups and individuals—for example, groups associated with various types of crime, individuals in proximity to at-risk locations, individuals at risk of victimization, and individuals at risk of domestic violence.

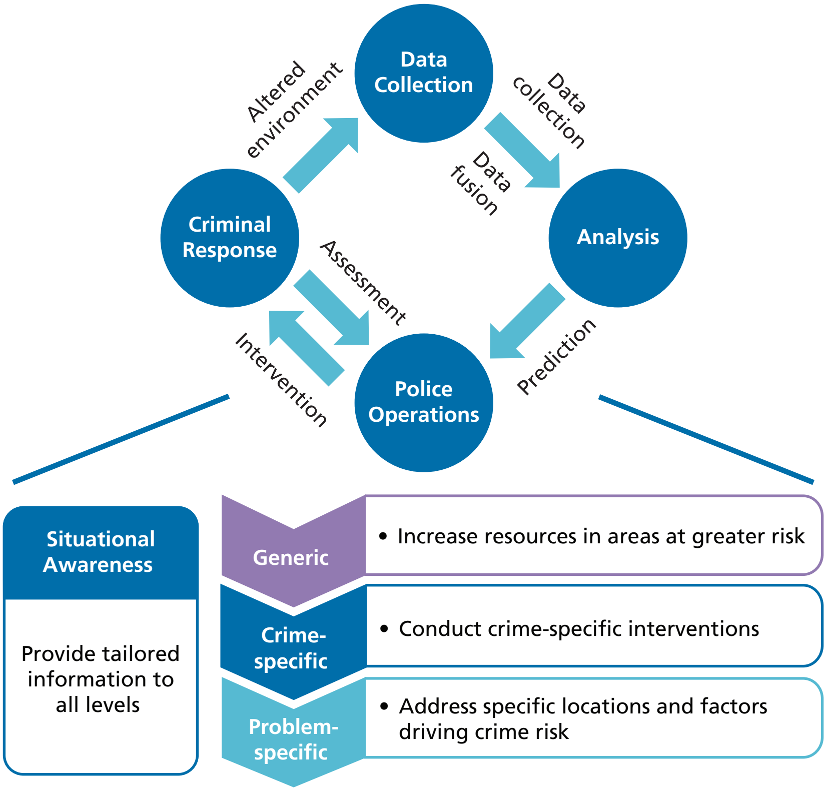
Prediction-Led Policing Process and Prevention Methods

Making "predictions" is only half of prediction-led policing; the other half is carrying out interventions, acting on the predictions that lead to reduced crime (or at least solve crimes). What we have found in this study is that predictive policing is best thought of as part of a *comprehensive business process*. That process is summarized in Figure S.1. We also identified some emerging practices for running this business process successfully through a series of discussions with leading predictive policing practitioners.

At the core of the process shown in Figure S.1 is a four-step cycle (top of figure). The first two steps are collecting and analyzing crime, incident, and offender data to produce predictions. Data from disparate sources in the community require some form of data fusion. Efforts to combine these data are often far from easy, however.

The third step is conducting police operations that intervene against the predicted crime (or help solve past crimes). The type of intervention will vary with the situation and the department charged with intervening. Figure S.1 shows three broad types of interventions (lower right of figure). They are, from simplest to most complicated,

Figure S.1
The Prediction-Led Policing Business Process



RAND *RR233-S.1*

generic intervention, crime-specific intervention, and problem-specific intervention. In general, we hypothesize that the more complicated interventions will require more resources, but they will be better tailored to the actual crime problems—and get better results. Regardless of the type of intervention, those carrying it out need information to execute the intervention successfully. Thus, providing information that fills the need for *situational awareness* among officers and staff is a critical part of any intervention plan.

The interventions lead to a criminal response that ideally reduces or solves crime (the fourth step). In the short term, an agency needs to do rapid *assessments* to ensure that the interventions are being implemented properly and that there are no immediately visible problems. The longer-term criminal response is measured through changes in the collected data, which, in turn, drives additional analysis and modified operations, and the cycle repeats.

Predictive Policing Myths and Pitfalls

Many types of analysis that inform predictive policing have been widely used in law enforcement and other fields, just under different names. The lessons from these prior applications can highlight many well-known pitfalls that can lead practitioners astray and can provide recommendations to enhance the effectiveness of predictive policing efforts.

Predictive Policing Myths

"Predictive policing" has received a substantial amount of attention in the media and the research literature. However, some myths about these techniques have also propagated. This is partly a problem of unrealistic expectations: Predictive policing has been so hyped that the reality cannot live up to the hyperbole. There is an underlying, erroneous assumption that advanced mathematical and computational power is both necessary and sufficient to reduce crime. Here, we dispel four of the most common myths about predictive policing:

- *Myth 1: The computer actually knows the future.* Some descriptions of predictive policing make it sound as if the computer can foretell the future. Although much news coverage promotes the meme that predictive policing is a crystal ball, these algorithms predict the risk of future events, not the events themselves. The computer, as a tool, can dramatically simplify the search for patterns, but all these techniques are extrapolations from the past in one way or another. In addition, predictions are only as good as the underlying data used to make them.
- *Myth 2: The computer will do everything for you.* Although it is common to promote software packages as end-to-end solutions for predictive policing, humans remain—by far—the most important elements in the predictive policing process. Even with the most complete software suites, humans must find and collect relevant data, preprocess the data so they are suitable for analysis, design and conduct analyses in response to ever-changing crime conditions, review and interpret the results of these analyses and exclude erroneous findings, analyze the integrated findings and make recommendations about how to act on them, and take action to exploit the findings and assess the impact of those actions.
- *Myth 3: You need a high-powered (and expensive) model.* Most police departments do not need the most expensive software packages or computers to launch a predictive policing program. Functionalities built into standard workplace software (e.g., Microsoft Office) and geographic information systems (e.g., ArcGIS) can support many predictive methods. Although there is usually a correlation between the complexity of a model and its predictive power, increases in predictive power have tended to show diminishing returns. Simple heuristics have been found to be nearly as good as analytic software in performing some tasks. This finding is

especially important for small departments, which often have insufficient data to support large, sophisticated models.

- *Myth 4: Accurate predictions automatically lead to major crime reductions.* Predictive policing analysis is frequently marketed as the path to the end of crime. The focus on the analyses and software can obscure the fact that predictions, on their own, are just that—predictions. Actual decreases in crime require taking action based on those predictions. Thus, we emphasize again that predictive policing is not about making predictions but about the end-to-end process.

Predictive Policing Pitfalls

To be of use to law enforcement, predictive policing methods must be applied as part of a comprehensive crime prevention strategy. And to ensure that predictive methods make a significant contribution, certain pitfalls need to be avoided:

- *Pitfall 1: Focusing on prediction accuracy instead of tactical utility.* Suppose an analyst is asked to provide predictions of robberies that are as "accurate" as possible (i.e., to design an analysis in which as many future crimes as possible fall inside areas predicted to be high-risk, thus confirming that these areas are high-risk). One way to accomplish this is to designate the entire city a giant "risk area." However, such a designation has almost no tactical utility. Identifying a hot spot that is the size of a city may be accurate, but it does not provide any information that police officers do not already have. To ensure that predicted hot spots are small enough to be actionable, we must accept some limits on "accuracy" as measured by the proportion of crimes occurring in the hot spots.
- *Pitfall 2: Relying on poor-quality data.* There are three typical deficiencies that can affect data quality: data censoring, systematic bias, and relevance. Data censoring involves omitting data for incidents of interest in particular places (and at particular times). If the data are censored, it will appear that there is no crime in a given areas. Systematic bias can result from how the data are collected. For example, if especially heavy burglary activity is reported between 7:00 and 8:00 a.m., it may not be immediately clear whether a large number of burglaries actually occurred during that hour or whether that was when property owners and managers discovered and reported burglaries that took place overnight. Finally, relevance refers to the usefulness of the data. For some crime clusters, it can be very useful to have data going back many months or years. Conversely, if there is a spree of very similar robberies likely committed by the same criminal, several months of data will not be of much use.
- *Pitfall 3: Misunderstanding the factors behind the prediction.* Observers—especially practitioners tasked with making hot spots go away—may reasonably ask, "For a given hot spot, what *factors* are driving risk?" "The computer said so" is far from

an adequate answer. In general, predictive tools are designed in a way that makes it difficult, if not impossible, to highlight the risk factors present in specific areas. There has been some improvement, however. When applying techniques, such as regression or any of the data mining variants, using common sense to vet the factors incorporated into the model will help avoid spurious relationships.

- *Pitfall 4: Underemphasizing assessment and evaluation.* During our interviews with practitioners, very few said that they had evaluated the effectiveness of the predictions they produced or the interventions developed in response to their predictions. The effectiveness of any analysis and interventions should be assessed as part of the overall effort to keep the data current. Measurements are key to identifying areas for improvement, modifying interventions, and distributing resources.
- *Pitfall 5: Overlooking civil and privacy rights.* The very act of labeling areas and people as worthy of further law enforcement attention inherently raises concerns about civil liberties and privacy rights . Labeling areas as "at-risk" appears to pose fewer problems because, in that case, individuals are not being directly targeted. The U.S. Supreme Court has ruled that standards for what constitutes reasonable suspicion are relaxed in "high-crime areas" (e.g., hot spots). However, what formally constitutes a "high-crime" area, and what measures may be taken in such areas under "relaxed" reasonable-suspicion rules, is an open question.

Recommendations

Our conclusions center on advice to three communities: police departments (the buyer), vendors and developers, and crime fighters. Our advice centers on the role of predictive policing in the larger context of law enforcement operations.

Advice for Buyers (Law Enforcement Agencies)

All departments can benefit from predictive policing methods and tools; the distinction is in how sophisticated (and expensive) these tools need to be. In thinking about these needs, it is important to remember that the value of predictive policing tools is in their ability to provide *situational awareness* of crime risks and the information needed to act on those risks and preempt crime. The question, then, is which set of tools can best provide the situational awareness a department needs?

Small agencies with relatively few crimes per year and with reasonably understandable distributions of crime (e.g., a jurisdiction with a few shopping areas that are the persistent hot spots) are unlikely to need much more than core statistical and display capabilities. These tools are available for free or at low cost and include built-in capabilities in Microsoft Office, basic geographic information tools, base statistics packages, and perhaps some advanced visualization tools, such as the National Institute of Justice–sponsored CrimeStat series.

Larger agencies with large volumes of incident and intelligence data that need to be analyzed and shared will want to consider more sophisticated and, therefore, more costly systems. It is helpful to think of these as enterprise information technology systems that make sense of large data sets to provide situational awareness across a department (extending, in many cases, to the public). These systems should help agencies understand the where, when, and who of crime and identify the specific problems that drive crime in order to take action against them. Key considerations include interoperability with the department's records management, computer-aided dispatch, and other systems; the ability to incorporate "intelligence" tips from officers (e.g., via field interviews) and the public; the types of displays ("dashboards") and supporting information the system can provide; and, of course, the types of analyses and predictions the system can support and under what conditions.

Advice for Vendors and Developers

The list of questions for purchasers doubles as guidance on desired capabilities for those who develop predictive tools. Looking ahead, it could be useful to move beyond predictions to offer explicit decision support for resource allocation and other decisions.

We emphasize that predictive policing tools and methods are very useful, but they are not crystal balls. Media reports and advertisements can give an impression that one merely needs to ask a computer where and when to go to catch criminals in the act. We ask that vendors be accurate in describing their systems as identifying crime risks, not foretelling the future.

Finally, developers must be aware of the major financial limitations that law enforcement agencies face in procuring and maintaining new systems. Licensing fees of into the millions of dollars are simply not affordable for most departments. We suggest that vendors consider business models that can make predictive policing systems more affordable for smaller agencies, such as regional cost sharing.

Advice for Crime Fighters

Generating predictions is just half of the predictive policing business process; taking actions to interdict crimes is the other half. The specific interventions will vary by objective and situation. (A number of examples are described in Chapters Three and Four of this guide; core resources on interventions are the Office of Justice Programs' CrimeSolutions.gov and the Center on Problem Oriented-Policing.) However, we have identified some promising features of successful intervention efforts:

- There is substantial top-level support for the effort.
- Resources are dedicated to the task.
- The personnel involved are interested and enthusiastic.
- Efforts are made to ensure good working relationships between analysts and officers.

- The predictive policing systems and other department resources provide the shared situational awareness needed to make decisions about where and how to take action.
- Synchronized support is provided when needed.
- Responsible officers have the freedom to carry out interventions and accountability for solving crime problems.
- The interventions are based on building good relationships with the community and good information ("intelligence").

Designing intervention programs that take these attributes into account, in combination with solid predictive analytics, can go a long way toward ensuring that predicted crime risks do not become real crimes.