Proposal 14: Smarter, Better, Faster: The Potential for Predictive Analytics and Rapid-Cycle Evaluation to Improve Program Development and Outcomes

Scott Cody
Mathematica Policy Research

Andrew Asher
Mathematica Policy Research

Introduction

Public administrators have always been interested in identifying cost-effective strategies for managing their programs. As government agencies invest in data warehouses and business intelligence capabilities, it becomes feasible to employ analytic techniques used more-commonly in the private sector. Predictive analytics and rapid-cycle evaluation are analytical approaches that are used to do more than describe the current status of programs: in both the public and private sectors, these approaches provide decision makers with guidance on what to do next.

Predictive analytics refers to a broad range of methods used to anticipate an outcome. For many types of government programs, predictive analytics can be used to anticipate how individuals will respond to interventions, including new services, targeted prompts to participants, and even automated actions by transactional systems. With information from predictive analytics, administrators can identify who is likely to benefit from an intervention and find ways to formulate better interventions. Predictive analytics can also be embedded in agency operational systems to guide real-time decision making. For instance, predictive analytics could be embedded in intake and eligibility determination systems, prompting frontline workers to review suspect client applications more-closely to determine whether income or assets may be understated or deductions underclaimed.

Rapid-cycle evaluation, another decision-support approach, uses evaluation research methods to quickly determine whether an intervention is effective, and enables program administrators to continuously improve their programs by experimenting with different interventions. Like predictive analytics, rapid-cycle evaluation leverages the data available in administrative records. It can be used to assess large program changes, such as providing clients with a new set of services, as well as small program changes, such as rewording letters that encourage clients to take some action. This type of formative evaluation can be contrasted with the summative program evaluations familiar to many in the policy community. Summative program evaluations often assess whether a program has an impact by comparing program participants with nonparticipants. Rapid-cycle evaluation uses similar techniques, but does not examine the overall impact of the program. Instead, it assesses the impacts of changes to the program by comparing some program participants (with the change) to other program participants (without the change).[1] For example, rapid-cycle evaluation can determine whether an employment training program can use text message prompts to encourage more clients to successfully complete program activities. In this way, rapid-cycle evaluation can identify incremental changes that make the program more effective for its clients, increasing the likelihood that a subsequent summative evaluation would identify large impacts relative to individuals not in the program.

We believe that these techniques can be used to help government programs—including social service programs serving low-income individuals—to improve program services while efficiently allocating limited resources. We believe that the use of predictive modeling and rapid-cycle evaluation—both individually and together—holds significant promise to improve programs in an increasingly fast-paced policy and political environment.

We propose that social service agencies take two actions. First, agency departments with planning and oversight responsibilities should encourage the staff of individual programs to conduct a thorough needs assessment. This assessment should identify where predictive analytics and rapid-cycle evaluation can be used to improve service delivery and program management. The assessment should also evaluate whether the benefits of adopting these tools outweigh the costs, resulting in a recommendation of whether and how these tools should be deployed. Second, federal agencies should take broad steps to promote the use of predictive analytics and rapid-cycle evaluation across multiple programs. These steps include investments in data quality and data linkage, as well as measures to support and promote innovation among agency staff.

The Challenge

Our proposal is based on the simple assumption that government programs could do better. This seems self-evident: despite decades of antipoverty efforts, the reality is that unemployment and underemployment, low food security, high poverty rates, and related problems persist. Rigorous evaluations of federal social programs show that many programs have little or even no impact on program participants.

In fact, even those programs held up as examples of proven, evidence-based programs demonstrate that government programs could do better. For example, the Coalition for Evidence-Based Policy identifies top-tier social programs with rigorous evidence of effectiveness, such as the Nurse-Family Partnership, Nevada's Reemployment and Eligibility Assessment Program, the Transitional Care Model, and other programs (Coalition of Evidence-Based Policy 2012). Multiple randomized controlled trials on each of these programs show positive impacts on client outcomes. But even this positive evidence suggests these programs could be more effective. A systematic review of research on the Nurse Family Partnership program concludes that there is evidence of a positive impact on only seven of the twenty-five measures of child maltreatment, and on only five of the fifty-nine measures of child development and school readiness (U.S. Department of Health and Human Services n.d.). The Nevada Reemployment and Eligibility Assessment Program increased employment among participants, but only modestly: 52 percent of program participants were employed, which is higher—but not substantially higher—than the rate in the control group, in which 48 percent of participants were employed (Michaelides et al. 2012). In short, even programs highlighted as success stories have room for improvement. They could benefit more clients and they could have a larger impact on the clients they benefit.

The administrators of these and other programs are constantly seeking ways to improve outcomes. Some administrators seek to match clients with the right services. But without the right analytic tools, these administrators cannot determine if their services are targeted as effectively as possible. Other administrators seek to test new procedures aimed at improving program services. But again, without the right analytic tools, these administrators may get biased results, leading them to implement ineffective changes or to dismiss effective ones. In the end, progress toward program improvement is slow, and programs end up spending resources inefficiently and leaving participants underserved.

A New Approach

Because predictive analytics and rapid-cycle evaluation have the potential to improve program effectiveness, we believe that social service agencies should conduct thorough needs assessments to identify, program by program, where these tools can be used. The needs assessments should examine the quality of existing program data to determine whether they are robust enough for use in predictive analytics and rapid-cycle evaluation. The assessments should also examine whether and how programs can deploy predicted outcomes operationally in a way that improves program performance. Furthermore, they should assess whether and to what extent experiments can be conducted to test changes in program operations. In addition to conducting program-level needs assessments, agencies should also take steps to promote the use of these tools broadly across multiple programs. These steps could include investments to improve data systems, improve data governance, and promote a willingness among program staff to test program innovations.

To inform the needs assessment, this section begins with an explanation of how predictive analytics and rapid-cycle evaluation can be deployed in the administration of public programs. These tools are not commonly used at this time. Where possible, we provide real-world examples of the application of these tools. We supplement these examples with a discussion of potential applications. Agencies should consider these real-world and potential applications when conducting their needs assessments.

PREDICTIVE ANALYTICS

At the individual level, predictive analytics leverages the fact that key outcomes and outputs for program clients are often correlated with the client's prior behaviors, circumstances, and characteristics, as well as those of the client's family, associates, service providers, and surroundings. By examining these correlations, predictive analytics methods can be used to rank program clients based on the likelihood that an outcome, whether positive or negative, will occur.

For example, an analysis predicting which participants of a job training program are likely to find employment might leverage existing information about the clients' education levels and their attendance at job training sessions. The model might tap these factors and other information to rank participants on the likelihood that they will find employment. Using these rankings, program administrators could decide, based on their goals and resource constraints, the exact sub-population that they want to target with their additional services. Depending on their program's objectives, administrators might focus on individuals most likely to find employment, or might target additional services to individuals less likely to find employment.[2]

Below we describe two key uses of predictive analytics for policymakers: (1) identifying program participants at risk of an adverse event and (2) predicting the optimal service path for an individual. We then discuss deploying predictive analytics to impact decision making.

IDENTIFYING PROGRAM PARTICIPANTS AT RISK OF AN ADVERSE EVENT

Program administrators can use predictive analytics to identify clients who are at risk of an adverse outcome such as unemployment, fraud, unnecessary hospitalization, mortality, or recidivism. Knowing which participants are most likely to experience an adverse outcome, program staff can provide targeted interventions to reduce the likelihood that such outcomes will occur.

Reducing readmission rates for certain patients discharged from the hospital provides an example of how predictive analytics can be used effectively. Reasons for unplanned readmissions can include clinical and social factors, such as patients' timely access to quality primary health-care services, their underlying conditions, whether they are homeless, and whether they lack social support and other factors that affect their ability to recuperate at home without incident (Peikes et al. 2012–13). If Medicaid programs could anticipate which patients are likely to be readmitted, they could intervene to address some of the factors contributing to the higher likelihood of a repeat visit. This would enable the patients to avoid another hospitalization while the Medicaid program would avoid paying for expensive hospital care.

Researchers at New York University have developed such a predictive model to identify a combination of characteristics and circumstances that indicate an elevated risk that a New York Medicaid beneficiary discharged from a hospital will return within one year (Raven 2009; Raven et al. 2009). New York City Health and Hospitals Corporation is using this model within its operational systems to screen admitted patients and identify interventions for those most likely to be readmitted for a preventable reason (Evans 2011).[3]

A similar approach could be used to prevent recipients of public assistance benefits from letting their eligibility lapse. Assistance programs such as the federal Supplemental Nutrition Assistance Program (SNAP), formerly known as the Food Stamp Program, require beneficiaries to demonstrate eligibility through a periodic recertification process. If clients do not complete the recertification process, their benefits are terminated. Clients often do not reapply for the program until they realize their benefits have been terminated. This creates two problems. First, clients who are eligible for assistance forgo benefits for one or two months until they reapply. Second, the program must bear the costs of processing a new application—which is more expensive than recertification. State agencies that administer the federal SNAP program could use predictive analytics to identify clients at risk of such churning. What would be required, beyond the tested and validated analytics themselves, is that the models be built directly into the case maintenance systems. Identifying these at-risk clients prior to the redetermination would enable program administrators to direct targeted, intensive communication efforts to these clients to prevent churning and help the clients maintain benefits while saving program funds.

Other potential areas for using predictive analytics include enforcement and fraud detection applications. For example, some child support enforcement agencies are developing predictive models to identify noncustodial parents who will not make their child support payments. This information can be used to triage enforcement efforts, making sure fewer resources are devoted to collection efforts against those who will ultimately pay without enforcement and identifying those who are likely to pay in response to more-aggressive efforts.

In addition, predictive analytics can be used to identify provider, client, vendor, and billing entity fraud patterns in health-care and social service programs. In SNAP, for instance, geographic patterns of electronic benefits transfer redemption and historical investigative data can be used to predict which program clients and retailers may be engaged in benefit trafficking (exchanging SNAP benefits for cash at a discount).

PREDICTING OPTIMAL SERVICE PATHS

Many government programs have different approaches to working with clients to achieve the same outcome. For example, there are multiple approaches to preventing recidivism among juvenile offenders, encouraging preventative health care, and boosting the parenting skills of new mothers. These paths may differ in the services involved or the time at which the services are offered. Under the right circumstances, predictive analytics can be used to determine which approaches are most likely to benefit which clients. Administrators can then identify the optimal service path for a client among the available options.

Consider a caseworker trying to find the right jobs program for a nineteen-year-old unemployed man with no high school diploma. This caseworker can enroll the individual in a low-intensity résumé support and job search program, a more-intensive program teaching specific manufacturing skills, or even a very intensive apprenticeship program. Each path has a different cost, and possibly a different outcome, for this individual. The caseworker's job is to match the program to the individual's background and interests. Combining this information—which is readily known at intake—with a prediction, based on which programs are associated with success for similar clients, could yield a better match between client and services, increasing the likelihood that the client will find employment and reducing the likelihood of wasting funds on ineffective training. Many agencies are interested in developing optimal service path predictions, yet in practice few exist. We believe there is an opportunity for optimal service path modeling to benefit the clients of public programs.

RAPID-CYCLE EVALUATION

Rapid-cycle evaluation, another tool that supports decision-making, is increasingly used in public programs with readily available administrative data and the ability to analyze those data in a rapid, cost-effective manner. This type of evaluation uses rigorous experimentation to test changes in agency operations.[4] To determine any impacts from the changes, administrators can compare client outputs and outcomes with those for other clients who are included in the evaluation but continue to receive regular services. The evidence from these tests can be more reliable than other sources, such as feedback from staff, complaints from selected clients, or anecdotes from other agencies.

To better understand how rapid-cycle evaluations can be used to test changes, it is useful to consider the three defining terms:

1. **Rapid.** The "rapid" means that the impact of the intervention will be identified quickly. To facilitate rapid identification of results, the outcomes of interest should be observable in administrative data. This eliminates the time-consuming process of collecting new data. Additionally, any impacts of the intervention should be observable within a short time frame. For example, it would not be possible to rapidly assess whether an intervention delivered to ninth-grade students leads more of those students to graduate from high school.

2. **Cycle.** The "cycle" refers to the iterative nature of the tests. Rapid-cycle evaluations can support a formative, continuous improvement model in which an intervention is tested, the results are examined, the intervention is modified if needed, and the modified intervention is tested again or a new intervention is tested.

3. **Evaluation.** The "evaluation" refers to the use of rigorous research techniques that generate confidence that observed changes in outcomes are due to the intervention and not to other factors (such as differences between the group that received the intervention and the group that did not).

This approach has been used by businesses for years to continuously improve the match between customers and services. For example, Capital One claims it runs more than 30,000 experiments each year to help identify the techniques that cause customers to sign up for new credit cards as well as techniques that encourage customers to pay Capital One back (Davenport and Harris 2007). The company experiments with changes in interest rates, promotional incentives, and even the color of the envelopes used in customer mailings.

Rapid experiments are used in the public sector as well to test a variety of program interventions, including changes in staff procedures, the services provided to clients or customers, and when and where those services are provided. Rapid-cycle evaluations of experiments can assess whether the interventions meet goals such as improving (1) the agency's ability to serve more clients, (2) the quality of information agencies get from clients, (3) client outcomes, and (4) agency efficiency. It is sometimes possible to test numerous variations of program services simultaneously. Box 14-1 shows how experimentation and rapid-cycle evaluation can fit into overall program operations by presenting applications used by New York City Human Resources Administration.

In some cases it may not be feasible to collect the necessary outcome data. For example, target outcomes may occur too far in the future to be examined in a rapid experiment (e.g., the eventual graduation of ninth graders). It may still be feasible, however, to employ rapid-cycle evaluation by looking at impacts on intermediate outcomes (such as class

BOX 14-1.

New York City Human Resources Administration

Agencies such as the New York City Human Resources Administration (HRA) have deployed rapid-cycle evaluation to improve program services. HRA recently tested new administrative procedures to increase the establishment of child support for children receiving cash assistance (Dinan 2013). Since court appearances are assumed to be a deterrent to establishing child support, the agency tested new procedures that would avoid a court appearance for the noncustodial parent. The agency's goal was to increase the percentage of these children with established child support orders, reduce the time needed to establish orders, and increase the proportion of noncustodial parents that comply with their established orders. HRA staff developed a simple random assignment process for determining which cases were eligible for the new procedures, and trained frontline workers to administer the pilot. The analysis showed the new procedures were unsuccessful. The rate of child support order establishment for the treatment group (57.3 percent) was essentially the same as the rate for the control group (56.5 percent). Moreover, it took longer to establish those child support orders established through the new procedures.

HRA also used experimentation to test streamlined procedures for investigating Medicaid eligibility fraud (Weinberg 2013). These streamlined procedures were designed to reduce the number of steps needed to investigate potentially fraudulent Medicaid enrollees. They used a four-month random assignment experiment to evaluate the impact of these new procedures. Although the new procedures reduced the time spent investigating cases by 12 percent, fraud investigations conducted with the streamlined procedures were less likely to be successful. Fraud was established for 44 percent of cases investigated through streamlined procedures, compared with 61 percent of those investigated through status quo procedures.

In both of these experiments HRA's new procedures proved unable to achieve the desired impact. But in each case HRA clearly and quickly established that the procedures were unsuccessful without having to implement these new procedures program-wide.

attendance and grades), as well as program outputs (such as the amount and quality of services provided). Such rapid experiments and rapid-cycle evaluations can often still help improve program services.

Rapid-cycle evaluation also could be used to measure real responses to potential policy changes. For example, programs like SNAP and Temporary Assistance for Needy Families (TANF) have numerous eligibility criteria and other regulations that are often debated by policymakers. These include deduction amounts, certification period lengths, benefit formulas, reporting thresholds for income changes, and even the required number of hours for participation in work programs. When the changes to these regulations are discussed, policymakers debate whether these changes will lead to higher or lower participation rates, and whether they will lead to longer program dependence or encourage employment. Rapid-cycle evaluation has the potential to generate rigorous, reliable information that can take the guesswork out of these policy debates. Regulatory changes can be tested to identify—and quantify—clients' behavioral response to these changes. This information can ensure that regulatory changes better meet policymakers' goals.

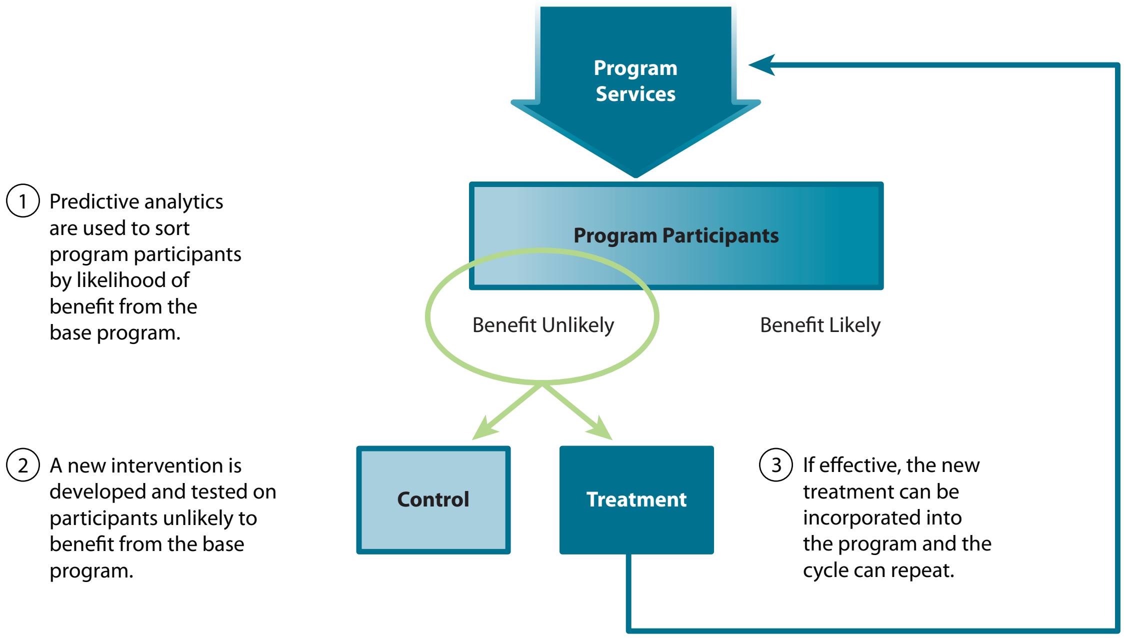
The greatest benefit of rapid-cycle evaluations to the agencies is the rigorous nature of the evaluation, which can replace other, nonexperimental techniques for assessing programmatic changes. For example, programs may pilot new procedures with all staff in a single location. In such cases, it is often not possible to know whether differences in outcomes are caused by the new procedures or simply by the unique circumstances of that location. This can lead program administrators to the false conclusion that a new procedure has promise, only to learn there is no benefit once it is implemented agency wide. Alternatively, it can lead them to reject a procedure that actually has promise.

COMBINING PREDICTIVE ANALYTICS AND RAPID-CYCLE EVALUATIONS

Predictive analytics and rapid-cycle evaluations can be combined to help program administrators build better interventions. Predictive analytics allow administrators to anticipate which individuals are most (and least) likely to benefit from a program. These predictions can help program administrators guide the formulation and scope of the interventions, and determine the group or subgroups to which they would apply. By creating targeted experiments, program

FIGURE 14-1.

Combining Predictive Analytics and Rapid-Cycle Evaluation: A Simplified Example



administrators can identify a series of effective, tailored interventions to maximize their ability to make an impact.[5]

Consider a program administrator seeking to test new approaches for reaching hard-to-serve clients. Initial predictive models could identify which of the program's current clients are least likely to benefit from the program, and could separate the clients into treatment and control groups. New interventions (or potentially multiple variations of the same intervention) could be tested rapidly and, if effective, could be incorporated into service delivery. After the new interventions operate for sufficient time, the entire cycle could be repeated (see figure 14-1) or applied to a different subgroup of program participants.

The new procedures can also be tested on individuals who are likely to benefit from the program. Such tests can help administrators determine whether new approaches would yield even greater improvements for individuals positioned to benefit the most from program services. Some administrators may view targeting those most likely to benefit as the most effective way to achieve gains for participants and improve the program's overall success.

For illustration, consider the hospital readmission prediction model mentioned earlier. A predictive model could be used to identify at-risk patients who are most likely to return to a hospital within one year. If program administrators want to test two different interventions for these at-risk patients, they could randomly assign the at-risk patients to one of three groups—one for each of the two interventions plus a control group—that receive the hospital's normal discharge planning and other services. The team would then monitor hospital admission rates for three months and assess whether the new interventions cause a significantly lower readmission rate. Any successful intervention could be integrated into program operations; the unsuccessful ones could be discarded.

If multiple interventions prove successful, program administrators could implement all of them or choose one based on cost and potential sustainability. The predictive model could be rerun and follow-up analysis could suggest new, tailored interventions for the remaining at-risk population. These interventions could be formulated and tested as in the previous cycle, evaluated, and either discarded or included in program operations.[6]

THE POLICY PROPOSAL

We propose that federal social service agencies take two actions. First, agency departments with planning and oversight responsibilities should encourage the staff of individual programs to conduct thorough needs assessment. This assessment should identify where predictive analytics and rapid-cycle evaluation can be used to improve service delivery and program management.

For predictive analytics, program administrators should assess:

- Whether predictions about specific client and program outcomes could be employed to target program services;
- Whether the program's current administrative data contain accurate, valid, and reliable measures of those outcomes—as well as valid and reliable measures of information that could predict those outcomes—to support predictive modeling; and
- The magnitude of systems enhancement efforts required to enable frontline workers to use the results of predictive models in real-time when they interact with clients.

For rapid-cycle evaluation, program administrators should assess:

- Whether program changes under consideration would benefit from precise, causally-valid impact estimates generated through rapid-cycle evaluation. The assessment can rely not only on program staff, but also on funders and outside experts to identify program features that they believe would be beneficial to test but were not sure should be implemented permanently without assessment;
- Whether program operations can be modified to facilitate experimentation of these program changes;
- What types of investments in data and systems would be required to deploy predictive analytics and rapid-cycle evaluation together as an integrated strategy;
- What types of programmatic waivers and other policy changes would be needed to facilitate predictive analytics and rapid-cycle evaluation; and
- Whether it would be beneficial to use predictive analytics to subset the program population, and to test program changes on different types of individuals (e.g., those most likely to benefit from current services).

The answers to each of these questions will vary by program. The assessment also should evaluate whether the benefits of adopting these tools outweigh the associated costs. In the end, the assessment should contain a recommendation of whether and how these tools should be deployed.

The second step agencies should take is to promote the adoption of predictive analytics and rapid-cycle evaluation more broadly across programs. We recommend that agencies take the following steps:

1. Help programs make individual-level data available for analytics. Individual-level data provide the best foundation for predictive analytics and rapid-cycle evaluation. These data can be obtained through internal operational systems maintained by the program, or through integrated data systems that combine administrative data across programs. A broad investment in data can facilitate predictive analytics and rapid-cycle evaluation and can promote the use of these tools across multiple programs.

   Federal agencies can help more programs benefit from these analytic tools by facilitating improvements to individual-level administrative data, and ensuring that those data are available for analytic purposes. Some agencies are already taking the lead in this respect. For instance, the Department of Education provided grants to promote the development of statewide longitudinal education data, and the Centers for Medicaid and Medicare Services funded data warehouses to help states manage all aspects of their Medicaid and Children's Health Insurance Programs. Agencies can also use their expertise to help programs identify the key measures to track for prediction and evaluation on an ongoing basis.

2. Improve data governance and facilitate data sharing. Although high-quality data are necessary, agencies also need strong data governance policies that establish accountability for data quality and that define the terms for how and where data are used (see Digital Services Advisory Group 2012). In addition, as part of data governance efforts, agencies should work to actively support efforts to link data across programs, which involves often-challenging technical and legal considerations. That said, linked data can provide a more comprehensive understanding of the services received and circumstances faced by clients, and provide more-accurate predictions and a more-complete understanding of the impact of rapid-cycle experiments.

3. Encourage analytic decision making. The use of predictive analytics and rapid-cycle evaluation requires an organizational commitment to testing program improvements. This means agency staff must develop program innovations—but be willing to abandon those innovations if they prove unsuccessful. For many program staff, this is a change in mindset from a focus on assessment of their program (and compliance with funder guidelines needed to properly evaluate their programs) to a focus on how to improve the programs and empower program administrators. Federal agencies can help foster innovation by providing performance-

based funding opportunities for program improvements.

Predictive analytics and rapid-cycle evaluation can be effective in part because they empower frontline staff to determine the services that best meet their clients' needs. However, the lessons learned for individual programs can be valuable to other programs serving these same populations. Agencies can further the effectiveness of these tools by ensuring successful efforts are highlighted, and their lessons broadly disseminated.

We believe that by taking these steps, federal agencies can help promote the use of these analytic tools at the federal, state, and local levels.

COSTS AND BENEFITS

Predictive analytics and rapid-cycle evaluation have a number of benefits. In particular, greater use of these tools would increase program effectiveness by reducing wasteful and inefficient spending. Even where the proposal results in an increase in direct outlays in one phase of a program's intervention, these outlays may generate net savings. Moreover, these analytical innovations would allow programs to fulfill their missions more effectively by better targeting their intended beneficiaries and helping them continually identify and implement cost-effective interventions.

That said, adopting these tools can require significant investments at a time when government budgets are under pressure. Developing the data and technology infrastructure necessary to deploy these analytical capabilities—if they are not already present—is expensive, as are, to a lesser extent, the resources needed to perform these analytics. For example, what may be considered to be the gold standard for data infrastructure—a full-featured, enterprise-wide data warehouse that integrates data across programs and is refreshed on a weekly basis—can cost several million dollars to build, and millions more annually to staff with dedicated maintenance and analytical personnel. Less-expensive data systems, such as purpose-specific analytical datamarts within existing warehouses or standalone databases focused on specific questions, may be more feasible and more appropriate in some cases.

As part of their needs assessments, agencies should assess the costs of any changes needed to deploy predictive analytics and rapid-cycle evaluation. In addition to data infrastructure costs, agencies should examine the costs associated with training staff, as well as the costs of altering program operations to incorporate predictive analytics and to implement rapid-cycle experiments and evaluations.

Agencies should compare these projected costs with the potential benefits obtained from these tools. In many cases, the benefits will include long-term savings in program administrative costs because the tools render the program more efficient. Other important benefits, however, such as improvements to the quality, availability, and access to services, should also be considered.

In the end, we believe that the benefits of predictive analytics and rapid-cycle evaluation will be substantial for many programs. We believe that this potential may warrant significant investment in these tools for many programs. For virtually all programs, however, we believe that this potential warrants the costs of conducting a needs assessment.

Questions and Concerns

In this section we examine some of the factors that could affect an agency's ability to adopt predictive analytics and rapid-cycle evaluation by posing and addressing a series of key questions and concerns.

What data resources are necessary for the use of these tools?

Programs with advanced information systems that contain individual-level administrative data are better suited to deploy predictive analytics and rapid-cycle evaluation with minimal investment. For example, in many states sophisticated cross-program data warehouses have been developed to support a wide array of Medicaid and social service program monitoring needs. These systems are rapidly updated and could be easily used for both predictive analytics and rapid-cycle evaluation. For other programs, administrative data obtained from transactional systems can provide an important source of information. These programs would require additional investment to create analytics-ready data repositories through the extraction, transformation, and storage of the data.

It is important to note that predictive analytics require historic observations of key outcomes. This means that programs developing new systems may not be able to perform predictive analytics until the system has captured enough history. Similarly, programs extracting data from transactional systems would need to extract a sufficiently large volume of historical data in order for predictive analytics to be effective.

Who would implement these tools?

The program managers and staff in agencies directly responsible for program delivery would implement these tools. To be successful, the implementation of these tools requires a division of labor. Program administration staff—both program operators and those working in support of them at federal and state agencies—need to determine which interventions are worth implementing and figure out how to do so. These program experts need to be supported by analytical specialists who are charged with designing the predictive analytics and assessing the results of the experiments through the rapid-cycle evaluations. Such a partnership allows this approach to become feasible and avoid burdening those with the pressing responsibility of running programs.

Can predictive analytics be wrong?

Yes. Predictive models detect patterns, but not every individual will follow that pattern. This can lead to incorrect predictions. Administrators can take several steps to minimize problems stemming from inaccuracies in predictive models.

First, predictive models should be subjected to extensive validation. For adverse event situations, such as hospital readmissions or fraud, models should be deployed historically so that their ability to predict known outcomes can be assessed. Through repeated retrospective testing, use and learning, the models can be improved, often to the point where they can be used prospectively.

Second, it is important to ensure model predictions are followed up by human judgment. Whether it is identifying clients who should receive a caseworker visit or those who may be defrauding the government, predictive analytics should be used to prioritize cases; staff should make the final determination. Similarly, even after optimal service paths are predicted, clients should still have a say in the services they receive.

Conclusion

As integrated data repositories become common in government agencies, program administrators have become comfortable using these data to monitor their programs. Now administrators are poised to expand the use of analytics to better decide what to do next. Predictive analytics and rapid-cycle evaluation, if used individually but especially if used together, can help agencies provide services where they are needed and develop more-effective approaches for improving program outcomes.

Authors

Andrew Asher

Senior Fellow, Health Policy Division, Mathematica Policy Research

Andrew Asher is a senior fellow at Mathematica and with extensive experience in health care management and data analytics—the use of data to help agencies and health care payers strengthen the integrity of their programs, realize cost savings, and improve program outcomes. He is one of the leaders of Mathematica's data analytics initiative and his expertise includes program integrity, health policy, health program development, big data, and developing strategies to help agencies improve their performance. He is nationally recognized for developing and implementing innovative data-driven approaches to help health care purchasers and insurers reduce fraud and abuse and manage their costs effectively. He rejoined Mathematica in 2013 after serving as vice president and director of government solutions at OptumInsight from 2007 to 2013 and in various leadership roles with the state of Illinois. From 2001 to 2007, he served as the manager of the Fraud Science Team within the Illinois Department of Healthcare and Family Services' Office of Inspector General, a group whose innovative fraud detection and prevention work was featured in Government Accountability Office and Center for Medicare and Medicaid Services best practice reviews. Asher was a health researcher and analyst at Mathematica from 1989 to 1994 and holds a master's in public policy from the University of Michigan.

Scott Cody

Vice President and Director of Human Services Research, Mathematica Policy Research

Scott Cody is a vice president and director of human services research in Mathematica's Cambridge, Massachusetts office. He is a national expert in the Supplemental Nutrition Assistance Program (SNAP). He led the first major in-depth study of state efforts to modernize enrollment procedures for programs like SNAP, Temporary Assistance to Needy Families (TANF), and Medicaid. His research also has examined the determinants of SNAP participation. He led an impact evaluation of efforts to increase program participation among the elderly, and multiple studies to identify the circumstances that lead individuals to enter and exit SNAP. Cody played a central role in adapting Mathematica's SNAP policy simulation model to reflect major changes in program policy. In addition to research on SNAP, Cody is the deputy director of the What Works Clearinghouse (WWC). The WWC is an initiative of the U.S. Department of Education's Institute of Education Sciences. The WWC assesses the quality of thousands of studies of education curricula, practices, and policies, and then summarizes the findings of well-designed studies in reports for educators, policymakers, and the general public. Cody holds a master's in public policy from the Johns Hopkins University.

Endnotes

1. We recognize that there are other definitions of rapid-cycle evaluation that will not utilize a comparison group. In this paper, we focus on the assessment of rapid experiments using comparison groups.
2. It is important to note that the performance of predictive analytics can vary. A number of considerations, including the extent to which strong predictors are available and the quality of the data, can affect performance. The strength of the underlying predictive models should be assessed before deploying predictive analytics in high-stakes situations.
3. The estimated equation generated by this model is used as part of an automated algorithm to find at-risk patients for intervention among those newly admitted to the hospital. In one early pilot, inpatient readmissions declined by 45 percent (Raven 2009; Raven et al. 2009).
4. Rigorous experimental techniques include randomized controlled trials and orthogonal research designs, and rigorous quasi-experimental designs include regression discontinuity research designs. These designs can be used to determine whether an intervention caused an outcome. Like a clinical drug trial, randomized controlled trials create randomly formed treatment and control groups, each receiving a different intervention. Orthogonal research designs use a similar approach but test variation in the components of an intervention. Regression discontinuity studies create a treatment group with individuals above (or below) a certain eligibility threshold (with individuals on the other side of the threshold forming the control group), and use analysis techniques to control for the eligibility score in the assessment of the program.
5. While we are not aware of specific, published examples of the use of these methods together, we believe the integration of these approaches is powerful and compelling, as the discussion that follows demonstrates.
6. It is noteworthy that examining multiple groups requires additional sample observations if the same level of precision is to be obtained. Generalizing beyond one hospital or program likewise requires additional sample observations.

References

Coalition of Evidence-Based Policy. 2012. "Social Programs that Work." Coalition for Evidence-Based Policy, Washington, DC.

Davenport, Thomas, and Jeanne Harris. 2007. *Competing on Analytics: The New Science of Winning*. Boston: Harvard Business School Publishing.

Digital Services Advisory Group, Federal Chief Information Officers Council, and Federal Web Managers Council. 2012. "Digital Service Governance Recommendations." http://www.whitehouse.gov/digitalgov/digital-services-governance-recommendations.

Dinan, Kinsey. 2013. "Local Agency Lessons On Implementing Random Assignment: An Example from NYC's Child Support Program." Presentation to the Association for Public Policy Analysis and Management, Washington, DC.

Evans, Melanie. 2011. "Healthcare's 'Moneyball': Predictive Modeling Being Tested in Data-Driven Effort to Strike Out Hospital Readmissions." *Modern Healthcare* 41 (41): 28–30.

Michaelides, Marios, Eileen Poe-Yamagata, Jacob Benus, and Dharmendra Tirumalasetti. 2012. "Impact of the Reemployment and Eligibility Assessment (REA) Initiative in Nevada." Impaq International LLC, Columbia, MD.

Peikes, Deborah, Rebecca Sweetland Lester, Boyd Gilman, and Randall Brown. 2012–13. "The Effects of Transitional Care Models on Re-Admissions: A Review of the Current Evidence." *GENERATIONS—Journal of the American Society on Aging* 36 (4, Winter): 44–55.

Raven, Maria C. 2009. "Improving Care and Cutting Costs for High-Risk Medicaid Patients." New York State Health Foundation, New York, NY.

Raven, Maria C., John C. Billings, Lewis R. Goldfrank, Eric D. Manheimer, and Marc N. Gourevitch. 2009. "Medicaid Patients at High Risk for Frequent Hospital Admission: Real-Time Identification and Remediable Risks." *Journal of Urban Health* 86 (2): 230–41.

U.S. Department of Health and Human Services (HHS). n.d. "Home Visiting Evidence of Effectiveness." U.S. Department of Health and Human Services, Washington, DC.

Weinberg, Stephen. 2013. "Using RCTs to Test Procedural Changes: Evidence From Medicaid Fraud Investigations." Presentation to the Association for Public Policy Analysis and Management, Washington, DC.