

Data Analytic Capabilities Assessment for Medicaid Program Integrity

**Centers for Medicare & Medicaid Services
Medicaid Integrity Institute
Medicaid Data Analytics Working Group
September 2014**



TABLE OF CONTENTS

Executive Summary	2
Definitions.....	4
Frequently Asked Questions	5
Self-Assessment: State Capability and Needs	10
Talking Points and Checklist	12
How to Achieve and Measure Success in Data Analytics	14
Fact or Myth?.....	18
Resources	19
Contributors	20

Executive Summary

This document is a reference for Program Integrity Directors and other Medicaid leadership to become better consumers of, and guide stakeholder discussion about, data analytics and predictive modeling products for program integrity purposes.

The use of data analytics and predictive modeling in the detection of fraud, waste, and abuse in healthcare programs can be a powerful tool for Medicaid program integrity administrators. Data analytics allows for detection and identification of patterns of fraudulent behavior not otherwise readily apparent. One strength of these tools is the ability to identify patterns of suspicious behavior based on historical data, thereby creating an opportunity for additional system edits to prevent future overpayments or any kind of fraud, waste, and abuse.

These tools combine powerful data modeling in diverse data sets to recognize patterns in providers and beneficiaries to focus limited investigative resources. Through an oversight of claims, suspicious patterns can be identified and scrutinized for further investigation. These systems, services, and software use state Medicaid Management Information System (MMIS) data, as well as other data sources, to build analytical products such as peer billing comparison reports, diagnosis cluster groupings, and other statistical comparisons to group like-providers. Systems and software may even be integrated within a state MMIS as a functional component serving the program integrity and surveillance utilization review business areas.

Implementing an effective program of data analytics offers several advantages, including a positive return on investment that can exceed that of traditional methods and lead to a cascade of strengthened program integrity safeguards throughout the Medicaid agency. A primary advantage of predictive modeling techniques is the potential for earlier detection of improper payments and the identification of predictors of billing anomalies. After such patterns are identified and validated as reliable predictors of improper billing, the information can be used to strengthen program integrity safeguards by implementing more effective edits, improving review of pended claims and prior authorizations, closing policy loopholes, and identifying areas of potential loss in Medicaid managed care. Similarly, the identification of predictors of improper billing can also lead to the development of new and more effective models for post-payment audits and recoveries, and to the maintenance of up-to-date provider education.

While states are currently performing some of these functions post-payment, predictive modeling tools can provide a more systematic approach to pre-payment claims. For example, by comparing providers of the same type, the system can identify long-term trending that is indicative of abusive billing behaviors, such as upcoding or high frequency use of certain codes. These trends can then be applied to future claim submissions in a pre-payment capacity.

Data analytics and predictive modeling are not stand-alone tools, but critical first steps in the investigative process. Staff investigators can use information from these tools to focus their efforts and resources to areas of the greatest risk and return, thereby more efficiently managing the Medicaid program, leading to greater recoveries, and discouraging future abuse.

While these systems and services can be expensive, both the Centers for Medicare & Medicaid Services (CMS) using predictive analytics pre-pay in Medicare and states using post-payment predictive modeling have seen positive returns through recoveries, savings, and prevention that may far exceed the cost of funding the services and hiring the technical staff to successfully implement them. Enhanced funding for such systems may be available in certain circumstances from CMS.

The Small Business Jobs Act of 2010 requires CMS to submit a report on the feasibility and cost-effectiveness of expanding the use of predictive analytics technologies to Medicaid and the Children's Health Insurance Program (CHIP), an analysis of the effect on states and territories, and recommendations regarding any technical assistance that may be necessary.

In Medicaid, many states are just beginning to leverage predictive modeling. Currently, only a few states have explored using these types of models. At the federal level, CMS is working with states to improve access to data, to identify opportunities that augment state efforts, and to leverage the initial success of the Medicare Fraud Prevention System.

To determine whether a state is ready to implement predictive modeling for program integrity purposes, the state should assess and evaluate:

- The flexibility of the state's MMIS to incorporate predictive modeling,
- The readiness of state staff to provide support to data analytics,
- Available resources or gaps in resources,
- The state Medicaid agency's priorities, and
- System-edit capabilities.

Definitions

“Predictive analytics” is a term of art and is often used to describe data analytics and predictive modeling. However, this document will use the broader term “data analytics” when addressing both post-payment and pre-payment analytics generally and the term “predictive modeling” when specifically discussing the use of techniques designed to extrapolate future events and/or behavior. Both are defined below:

Data analytics is a broad category of data assessment approaches and tools. Data analytics includes predictive modeling but also includes other approaches, e.g., link analysis, machine learning, graph pattern analysis, scoring, trend analysis, spike analysis, and cluster analysis. Some of these techniques are more appropriately applied to post-payment as compared to pre-payment analysis.

Predictive modeling describes a range of analyst-driven statistical and analytical techniques that use historical data to predict future events or behavior. Predictive modeling requires human intervention and evaluation to make a final decision regarding the payment status of a healthcare claim.

Frequently Asked Questions

1. What is predictive analytics?

“Predictive analytics” is a term of art often used to describe data analytics and predictive modeling. This document will use the more precise terms “data analytics” and “predictive modeling.” Please refer to page 4 of this document for detailed definitions of these terms.

2. What do I want or need? A case management system? A data warehouse? A list of providers who scored high risk?

It is important to fully contemplate your state’s needs to implement a "data analytics" system for program integrity purposes. Before reaching out to vendors or responding to their requests, take time to reflect upon what information will be actionable for your organization. The answer is critically important, and will inform all of your decisions going forward.

Initiating a new approach to data analysis is an opportunity to think about where your data analytics should be in five or ten years, not just where it could be in a year. Recognize that analytics are a significant investment, and that a mistake is costly to correct.

Data analytics require more than just a program that can assess data or draw linkages. Often states need the machinery to make all that work—servers, upgraded computers, data storage, and a case management system to track new cases and mine them for additional data. A simple system such as a risk scoring approach still requires a significant amount of support that can be expensive. Think about what your specific needs and goals are, then work back from that to determine what you have that will serve in the new approach, what you need to replace, and what you will need to purchase.

3. Do I want a pre-pay or post-pay system?

Your decision here will be governed in part by whether your MMIS system will accept the type of data analysis you want. If it does, you can consider a pre-pay system. If not, it is a post-pay solution. One of the factors to consider when deciding between a pre or post-pay system is to consider the impact to current state contractors such as Managed Care Organizations (MCOs) and fiscal agents.

Consider your prompt pay laws: follow-up associated with investigating results from data analytics often can make the claim payment late. In either pre- or post-pay scenarios, consider what staff you will need to implement the project and whether the vendor you select will provide

any resources. Bear in mind that a post-pay system also has pre-pay implications: the lessons learned from post-pay analytics can inform pre-pay decisions that may in turn help to reduce or eliminate potential losses.

4. What types of analytics do I want? Risk Scoring? Graph Pattern Analysis? A claims-based approach or a pattern approach?

Vendors often claim their product does what you want—whatever that may be. States must decide for themselves—first—what technology they need and can support.

A risk scoring methodology is the simplest approach, and provides a statistical analysis of individual claims to show which claims are most likely to have payment issues based upon what claims have had problems in the past. Risk scoring can be exceptionally informative but requires a good deal of state investigative resources to follow up on leads. Risk scoring cannot take into consideration state peculiarities, such as great size (which could affect geospatial calculations and provide false negatives), but is the least expensive system to implement.

A link analysis approach is more complicated and assesses both statistical probabilities and connections that exist in the data (such as two providers using the same address). Link analysis is significantly more informative than risk scoring and provides better and more refined leads, but requires more infrastructure and support.

Graph Pattern analysis is a further evolution, providing statistical and link analysis but also providing the capability to ingest vast quantities of unrelated data (such as private commercial records, property tax records, exclusions, banking records, and voter records) and process the data to identify patterns and relations among apparently unrelated data points. Graph Pattern analysis can handle both structured and unstructured queries and is designed to work through large quantities of data rapidly. It is designed to root out bad behavior, not just bad claims, and is not a good selection if your goal is simply to identify improper claims. Graph Pattern analysis requires the greatest amount of support, both in terms of staffing and data capacity, but provides substantial, detailed results requiring less investigative time than other approaches.

5. What data are available to analyze, and who is responsible for obtaining the data—the state or the vendor?

The data you provide for analysis will be the basis for all your results. Better, cleaner, and sufficient data yield better outcomes. Although you have control over your Medicaid data, you will likely want to expand your database to include other useful data sources. Third party data sources—such as but not limited to Medicare data, driver's license information, birth/death records, voter records, banking records, property tax records, professional licenses, corporate records, private commercial records, and exclusions by the Office of Inspector General (OIG)—add depth and meaning to health care data but can be expensive.

6. What is the vendor's overall approach to implementing a data analytic system?

Who hosts the system is important. If the system is integrated, where does the tool reside? Does the solution have a management reporting capability? How are claims routed to end users? Can thresholds be predetermined to flag high risk claims? Make certain the vendor clearly explains its ability to incorporate external data sources into the tool. What user support and training will you need, and what does the vendor provide on an ongoing basis? How does the vendor manage false positives?

Ensure the vendor describes its overall approach without guidance from you. You know what you need; determine whether the vendor's approach meets your identified needs without telling the vendor what you are looking for. Questions to ask are the type of analytic approach the vendor uses, ease of deployment, time to deploy, previous deployments and results (good and bad), whether the state or the vendor hosts the technology and software, who will use the software, and its ease of use.

7. What results can I expect? How much and how often will information be communicated and what format will be provided to the end user?

Make sure the vendor provides you with specifics about the information that the system and services provide. Lack of previous deployments should not necessarily rule a vendor out, but previous successful deployments are a good indicator of future success. Have the vendor identify specific successes.

The content, format, and frequency of output that a vendor's system provides are essentially what analytics is all about. The output of an analytics system can be as simple as a monthly list of potentially bad claims or as involved as a full investigation of a fraud scheme. You should control the type of information you receive and how you receive it, because the ability to integrate results into an existing system is important.

8. How much will the tool, licenses, training, and support cost? Will there be a need for special hardware to purchase server(s)? How will the state meet these costs?

Data manipulation is expensive. Project costs will vary widely, based on basic requirements (the type of analytics you select, whether you need a case management system, the type of output you want), infrastructure needs, staffing needs, and the amount of data you have and want to obtain. Case management systems, more extensive data, and more complicated analytics platforms will multiply that cost substantially. Smaller programs need to consider whether the increased results from more sophisticated systems justify the additional costs, which can be significant. Federal support may be available in certain instances through CMS to help states upgrade or implement a data analytics program. Federal match rates range from 50/50 to 90/10, depending on the nature of the cost.

9. What is the timeline for the project from startup, design, development, and testing to deployment, stabilization, and maintenance? What state resources are required for each phase of the project? What vendor support will be provided during each phase?

Data projects always require state resources, from explaining Medicaid rules to resolving information technology concerns. Ask whether the vendor will completely staff the project after deployment, whether state employees will manage it, or whether there will be shared responsibility. You should plan for significant staff time to familiarize the provider with basic business operations and needs.

10. How is data transferred to and from the vendor? Are the vendor's processes compliant with the Health Insurance Portability and Accountability Act (HIPAA)?

Have the vendor explain the process and their capabilities for transferring the data from state systems and returning results back to the state. This process will provide insight into the vendor's overall approach and will help you to be certain the vendor complies with federal and state privacy laws.

11. What is the most important factor in vendor selection?

One of the most important predictors of vendor success is prior success. Once you have determined the capability you want to acquire and you have identified a panel of vendors, it is critically important to ascertain the prior experience of the vendors. This includes experience with the tool they are proposing, as well as the vendor's experience with large data sets and with healthcare claims analytics. Focus should be given to the vendor's understanding of your business requirements in combination with the vendor's proposed technology. You should focus

on the references the vendors can provide and whether those references are able to attest to both positive experience and a strong return on investment. Do those references identify the vendor as a source of analysis expertise, or is their expertise limited only to statistics and technical modeling, without an understanding of the business patterns being modeled and the significance of the outcomes?

Self-Assessment: State Capability and Needs

Prior to spending money and resources on new technology, it is important to have a good understanding of your state's current situation and capabilities. An in-depth understanding of current capabilities and future goals is essential to making wise investments in new analytical systems. Any new technology will either need to integrate with your current system or may require a more comprehensive plan for a system upgrade. It is critical that you understand the needs of the end-user.

Program Characteristics

CMS requires states to participate in periodic state program integrity reviews. The program integrity review report from CMS provides a useful outside perspective on your state's current situation and capabilities.

Program assessment is contained within your state program integrity review report. Information such as the following will also be helpful as you plan for additional technology:

- Number of providers—enrolled in both fee-for-service and MCO networks,
- Number of MCOs,
- Number of Medicaid enrollees—fee-for-service, managed care, and other types of newly evolving payment methodologies,
- Current Medicaid Expenditures, and
- Pre-Pay Edit Capabilities.

Where am I now?

What are my current resources?

- Audit capabilities – consider state staff, contract staff, Third Party Liability vendor, Recovery Audit Contractor, and Medicaid Integrity Contractor.
- Investigative capabilities – consider state staff and contract staff.
- Analytic capabilities – assess data mining and analysis staff, and analytic tools.
- Analytic and business knowledge – this knowledge and understanding will help in implementing predictive analytics that meet the needs of the end-user.
- What data are available?
- What is the current state of the data?
- How are the data currently accessed?

- Can the data be migrated?
- Can the existing system handle new technology?
- Do I have an effective and efficient case management system? Is the case management system integrated with analytics?

Where do I want to go?

Should I hire staff, or contract for additional staff?

- Caps on Full Time Equivalent staffing
- Staff required for deployment
- Staff required for ongoing data analysis
- Staff required for increased audits and investigations

What are funding possibilities?

- CMS through Advance Planning Document
- Legislatures (General Revenue)
- Medicaid program dollars

What do I expect?

- Increase of inappropriate payments identified
- Increase of fraud cases and recoveries
- Identification of policy lapses

Should we procure an analytics system?

- Do my program needs justify a new analytic system and if so, what kind?
- What are my next steps?

Reviewing these frequently asked questions will assist you in completing the self-assessment process.

Talking Points and Checklist

Now that you have decided to implement a data analytics system:

Internal Discussion

- Do you want pre-pay analytics, post-pay analytics, or both?
- How will the data analytics system fit with the current system?
- Do we need any hardware upgrades or additions?
- Do we have existing staff capability to deploy and implement an analytics system?
- What data sources are we including?
- What claim types or provider types will the analytics assess?
- Do we have or will we need to develop a governance process to prioritize program integrity vulnerabilities?
- Will the system track financial recovery information?
- Will investing in data analytics further my agency's capacity and capabilities to prevent and reduce fraud, waste, and abuse? How? What are the measures of effectiveness?

Vendor Discussion

- Has the vendor worked with other states/entities?
- Was the vendor successful? How does the vendor describe and demonstrate success?
- How does the vendor plan to model or demonstrate the system?
- What rules or edits will be used to detect aberrant claims and/or providers?
- Do rules/edits come ready-made in the system? Does the state agency define and build the rules and edits, and then integrate them into the system? Or is this work done by the vendor?
- What is the time commitment for state agency analysts/staff?
- What is the frequency and cost of vendor support and training at implementation?
Ongoing?
- Will the proposed solution incorporate social media or geospatial mapping?
- Will analysts have access to proprietary systems logic so that results can be verified?

- Will information produced by the analytics system be understandable to end users? (graphs, charts, comparisons, lists scores, description of associated risks [trigger reason])
- How will system defects and problems be tracked and resolved?
- What level of customization within the program is possible? Does that affect cost?
- Additional considerations to keep in mind as you move forward with successfully implementing predictive analytics technology are the needs to:
 - Develop a robust process for incorporating field intelligence, policy knowledge, and clinical expertise into the development of predictive or other sophisticated algorithms to ensure that the results of the technology are actionable.
 - Develop a method for tracking, measuring, and evaluating the actions taken based on the information produced by the technology. The technology is a tool to provide more accurate leads more quickly; in order to achieve savings, the information must then be used appropriately to take action.
 - Develop an analytic environment for data exploration that includes historic information necessary for predictive modeling and an operational environment that quickly displays results and visualization (graphics, maps) that assists the end user in taking action.

How to Achieve and Measure Success in Data Analytics

Better analytic tools, including data analytics, may allow State Medicaid program integrity groups to be more efficient, more productive, and more effective. However, regardless of the analytics, richer and more reliable data will produce more detailed and accurate analyses and better overall decision-making processes. State Medicaid agencies should review the CMS reports to Congress on implementation of the Center for Program Integrity's Fraud Prevention System listed in the Resources section. In considering the incorporation of Data Analytics into state decision-making, it is important to consider several key factors:

1. Direct Fiscal Impact

Potential Savings from:

- Denial of claims
 - Based on pre-payment reviews
 - Based on auto-denial edits
- Costs avoided by more thorough screening of providers and revocation of billing privileges
 - It is strongly recommended that cost savings is incorporated into the return on investment methodology to ensure that the expenditures that are prevented are part of the savings. Since there is not an audit trail for prevented payments, as there is for recoveries, consider engaging actuarial expertise in the methodology development.
- Payment suspensions
- Recoveries of overpayments
 - From audits
 - From Credible Allegations of Fraud investigations
- Sentinel effects including self-modification of behavior by providers and beneficiaries
- MCO rate reviews which compare expected treatments and outcomes with actual treatments and outcomes to validate rates

Costs:

- Contractor costs
 - Overseeing analytics system

- Conducting investigations
- Conducting audits
- State costs
 - Overseeing contractors
 - Conducting audits
 - Other direct activities
 - Assess ability to fund and maintain the system after initial investment and implementation

2. Improvements in Fraud, Waste, and Abuse Prevention, Detection, and Deterrence

- Greater ability to identify new patterns and trends related to
 - Systemic vulnerabilities
 - Suspicious provider activities
 - Suspicious beneficiary activities
- Increase in referrals to Medicaid Fraud Control Units (MFCUs)
- Greater ability to validate and compare claims across different provider types
- Optimization of existing staff resources through better and more focused data retrieval
- Inform policy changes to close vulnerability gaps to prevent future risks

Better prepayment predictive models are used to flag suspicious claims for additional review by staff before payment is made. Post-payment review identifies patterns and trends, and leads to recoveries, edits, and policy changes.

3. What are the different uses for predictive modeling?

Some different uses are:

- Screening providers and conducting post-enrollment provider analysis,
- Identifying possible provider and recipient relationships,
- Verifying home visits, and
- Providing claims analysis prior to payment.

4. Is predictive modeling applicable to all Medicaid claims payments?

Medicaid personnel should scrutinize vendors' representations that their systems will apply to the entire Medicaid claims population. Pre-payment analytics may not be suited to managed care payments, waiver programs, point of sale systems, or Medicaid/Medicare dual-eligibles, depending on the state payment system.

5. How do predictive and other data analytics apply to managed care?

Post-payment analytics assist the states with accomplishing their responsibilities to perform surveillance and other oversight of MCOs. This oversight helps ensure that future capitation rate setting is accurate, that provider payments are valid, and that important policy changes are considered.

6. Why are human resources essential to this process?

The analytics are the initial tool used to identify possible issues in billing claims. Findings need to be verified and vetted for reasonability through human review of the data to ensure accuracy.

7. Can “Data Analytics” alone be used to identify potential fraudulent providers and deny claims?

Not typically. Systems generally suspend suspect claims: human resources are needed to investigate and determine the proper adjudication of claims. Claims identified as suspect are not necessarily always in error—a characteristic necessary for denials. Thus, staff who interpret and investigate the results must review suspended claims. The state should ensure that standard operating procedures are in place and sufficient staff resources are allocated and trained to manage the adjudication of claims suspended by the data analytics system as potential improper billings.

8. Potential Enhancements in Healthcare Delivery

- Better able to determine whether care is within accepted standards of practice
 - Allows for verification that adequate and appropriate services and/or products are being provided
- Ability to analyze treatment outcomes (expected outcomes versus actual outcomes) at various levels
 - Plan level
 - Provider level
 - Beneficiary level
- Ability to generate practitioner peer reviews
- Ability to generate provider comparisons, including comparisons among MCOs
- Better able to inform and educate providers

9. Is there a difference between predictive modeling and system edits?

Generally, data analytics that are performed post-payment inform front end system edits. The intersection of payment system edits and pre-payment predictive modeling tools depends on the

capabilities of a state's payment system. It is important to understand system edit capabilities while assessing the purchase of predictive modeling tools.

10. Are there different options for data analytic tools?

Options can include buying tools to integrate into existing systems, subscription services that include investigation as well as identification, or an in-house build of predictive modeling tools.

11. Will a vendor with “predictive analytics” experience in the credit card industry also be successful in the healthcare industry?

Not necessarily. While the predictive concepts developed by the credit card industry (over a 20 year period) may be applicable to health care transactions, the state should consider the complexity of its health care system in selecting a vendor.

Fact or Myth?

Myth: Our system will save your state millions of dollars by implementing predictive analytics.

Fact: Predictive modeling has not yet been proven in the Medicaid program. Specific savings amounts have yet to be determined, so vendor claims are simply guesses.

Myth: “Predictive analytics” systems can be easily and painlessly integrated into any state’s existing Medicaid payment system.

Fact: Using analytics in a state’s claims processing cycle depends on the state’s payment system, technology platform, and vendor.

Myth: Most states have already implemented a “predictive analytics” system and are already saving millions.

Fact: Some states have implemented predictive modeling tools for post-payment claims analysis, while even fewer have attempted to apply these techniques to pre-payment claims.

Myth: The *Small Business Jobs Act* requires states to implement “predictive analytics” into their Medicaid integrity programs.

Fact: The *Small Business Jobs Act* required CMS to implement predictive modeling and other analytics technologies in Medicare in three phases over a three-year period and then to submit a report on 1) the cost-effectiveness and feasibility of expanding the use of predictive analytics technologies to Medicaid and CHIP, 2) an analysis of the effect on states and territories, and 3) recommendations regarding any technical assistance that may be necessary. The expansion of predictive analytics technologies to Medicaid and CHIP is to be made based on the results of this report and its recommendations.

Resources

Centers for Medicare & Medicaid Services, Report to Congress Fraud Prevention System Second Implementation Year, June 2014, <http://www.stopmedicarefraud.gov/fraud-rtc06242014.pdf>

Centers for Medicare & Medicaid Services, Report to Congress Fraud Prevention System First Implementation Year, December 2012, <http://www.stopmedicarefraud.gov/fraud-rtc12142012.pdf>

If you have any questions or would like more information on this topic, please contact Thomas J. Kessler, Center for Program Integrity, at 410-786-1991 or Thomas.kessler2@cms.hhs.gov.

Contributors

Developed by:

**Medicaid Integrity Institute
State Medicaid Predictive Analytics Working Group
November 2013**

James Guy – State of Florida

Benjamin Oberholtzer – State of Pennsylvania

Kevin Ryan – State of New York

Joan Senatore – State of Massachusetts

Jack Stick – State of Texas

Lee A. Worth – State of California

Monica Cuba – State of Kansas

Bradley K. Hart – State of Illinois

Pamela Learned Vickers – State of Oregon

Steve Murray – State of North Carolina

Cathie Ott – State of Washington

Rick Zenuch – formerly of State of Florida