OnPoint: Issue Brief

A Bi-Monthly Publication by the Massachusetts Association of Health Plans

Volume VI, December 2016

MAHP
MASSACHUSETTS ASSOCIATION *of* HEALTH PLANS

Written by Elizabeth A. Leahy, Esq.

Can Analytic Tools Help at Risk Children?
Harnessing Big Data across Multiple Systems.

An Overview of Tools in the Market Today

Child Abuse and Neglect in Massachusetts

In a recent report from the U.S. Department of Health and Human Services, Massachusetts was identified as the state with the highest reported rate of child abuse and neglect during fiscal year 2014.[1] State officials indicated that the number of cases in the Commonwealth rose in part because of public awareness raised by several high-profile cases that ended in tragedy. Like many other stakeholders, Massachusetts health plans looked for opportunities to assist the state in efforts to protect this vulnerable population. The Massachusetts Association of Health Plans put together a workgroup to explore the use of analytic tools, including examining the potential use of claims data and related privacy concerns, as a means to help identify children in need of assistance. While plan data alone is not sufficient for analysis, MAHP plans reviewed several analytic tools currently in use in other states and have convened a forum for plans and policymakers to hear from two experts in the field. MAHP's latest *OnPoint* examines the field of predictive analytics and is a companion to the December 9, 2016 Policy Forum: ***Can Analytic Tools Identify Children at Risk for Abuse?*** *A roundtable conversation with risk analytics experts, health plans, and state stakeholders*.

Big Data and Predictive Analytics

As Governor Baker and the Administration embark on a multipronged and multiyear reform of the state's Department of Children and Families, policymakers and stakeholders have begun to examine a number of tools to help identify children at risk for abuse and neglect. One such tool, predictive analytics, is receiving increased attention in human services agencies across the country. Predictive analytics, or "the practice of extracting information from data sets to determine patterns and predict outcomes and trends,"[2] is not a new concept; in fact, it has been in use by insurers for quite some time. Many health care payers mine claims, pharmacy, and demographic data to identify sources of risk and create targeted interventions. As the amount of available data increases across multiple systems, the opportunities for linking data sources to analyze outcomes increases exponentially, and health plan data sets can be a useful resource. Taken alone, claims and encounter data are not appropriate sources for identifying at risk children; however, the addition of unique information from health plans may help close gaps in data and paint a broader picture for the state. Whether through the state's All Payer Claims Database (APCD), or in partnership with participating health plans, opportunities exist for the Commonwealth to combine health plan data with human services data to build a more robust data set.

Like analytics, the big data movement seeks to harvest intelligence from data and translate that into useable information. According to the term's originator, Doug Laney, "Big data is high-volume, -velocity and –variety information assets that demand cost-effective, innovative forms of information process for enhanced insight and decision-making."[3] The amount of available data is vast, some with the potential for real-time collection, and many organizations struggle to dig through the data for valuable insight.

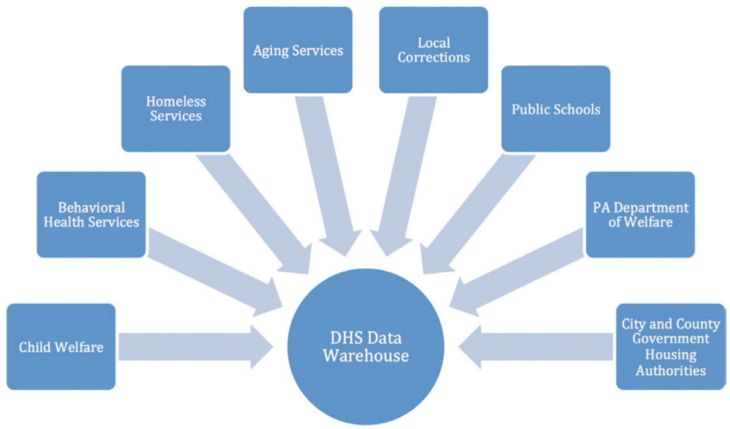
Human services agencies are no exception. While they have always collected, stored, and reported a vast amount of data, the information was rarely readily available for problem-solving.[4] With data collection occurring in silos across multiple agencies, the ability to combine data sets to identify risk factors was traditionally limited. However, with the application of predictive analytics and predictive risk modeling, human services agencies have begun to harness data across multiple systems. These predictive models typically analyze current and historical data to produce easily understood metrics, results that go beyond merely collecting and sorting data by turning the information into data capable of providing future opinions and predictive capabilities.[5] Predictive analytics uses a computer algorithm to search through multiple data sets to look for patterns, interactions, and signals,[6] while predictive risk modeling uses those patterns to identify predictors of risk and assign risk categories to individuals or families.[7] In human services agencies, quantifiable measures used to track and assess, such as rates of child protection reports or substantiated cases, would be just some of the metrics used to examine and predict child welfare outcomes. The use of predictive analytics and predictive risk modeling is fairly new to human services agencies, but there are several models currently in development and in use in agencies in the U.S. and abroad.

Models in Development

Pennsylvania – Allegheny County Data Warehouse

In 1999, Allegheny County, Pennsylvania's Department of Human Services became one of the first agencies in the U.S. to harness data collected across systems for use in evaluating the possibility of future adverse outcomes in the child welfare system. The county developed a Data Warehouse of more than 17 internal and 10 non-DHS data sources, including data regarding child welfare, behavioral health and intellectual disability, alcohol and drug use, aging services, community services, public schools, corrections, and the Department of Public Welfare. The system harvests the data and produces a Risk Score for each child, indicating how likely the child is to be at risk for a case opening. Call screeners at the county's child abuse hotline are provided with a risk score when a call comes in, and high risk level callers can be assessed in person.[8] Allegheny County is monitoring the effectiveness of the program through training and the build out of an independent evaluation tool.

Figure 1 Internal DHS Data Sources (not inclusive) / External Data Sources (not inclusive)



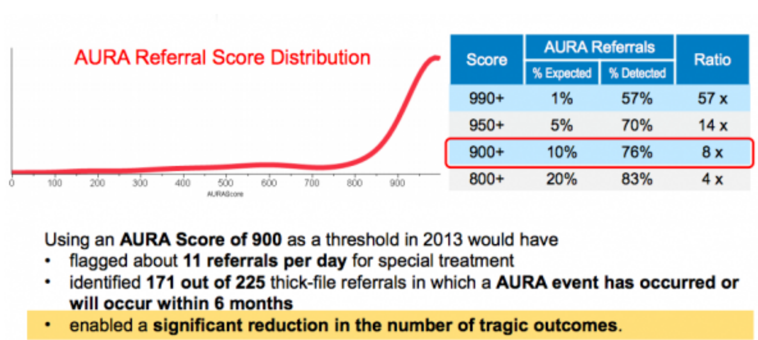
Florida – Eckerd's Rapid Safety Feedback

Launched in 2013 following several child deaths from maltreatment in Hillsborough County, Florida, the Rapid Safety Feedback Tool (RSF) is a predictive analytics tool built into the state's automated child welfare information system (SACWIS). After review of both the open child welfare cases in the county and the child death cases, Eckerd, the Community Based Care Lead Agency for Hillsborough County identified common risk factors for review and developed a profile of cases with the highest probability of serious child injury or death. Eckerd then contracted with Mindshare to provide system overlay software that allows cases to be mined in real time for the common risk factors identified with cases that have a high risk of child tragedy or death. Each high probability case is reviewed by Eckerd quality assurance staff utilizing the Eckerd RSF tool; if any safety concerns are identified during the review, the quality assurance staff meets with the case manager and supervisor to develop a plan to mitigate concerns and provide coaching and support for case management staff. Eckerd is currently working to apply its predictive analytics model to other states, including Connecticut and Maine.[9]

Los Angeles – Approach to Understanding Risk Assessment

In 2014, the Los Angeles County of Department of Children and Family Services contracted with advanced analytics firm, SAS, to develop and test a risk assessment model that looked at child deaths, near fatalities, and critical incidents over a two year period to develop a risk score for application to DCFS referrals. SAS used state data, including: prior child abuse referrals, involvement with law enforcement, mental health records, and alcohol and substance abuse history to develop the tool, which was then tested against 2013 DCFS referrals to determine the model's success.[10]

Figure 2 Thick File Performance on 2013 Validation Data



New Zealand – Predictive Risk Modeling

In 2015, New Zealand's Ministry of Social Development began trialing the use of predictive risk modeling in relation to child maltreatment. The model, developed by Auckland University of Technology Professor of Economics, Rhema Vaithianathan, seeks to predict the risk of a newborn baby experiencing substantiated child abuse by the time the child turns five. The model uses a linked data set of information from agencies across New Zealand's Ministry of Social Development, including the Work and Income Agency and the Child, Youth, and Family Agency.

The model uses 132 variables relating to demographics, socio-economic status, primary caregiver historical data, the primary caregiver's partner's historical data, and the child's historical data. The Ministry of Social Development is currently testing whether the Predictive Risk Modeling can enhance social workers' decision-making at intake for children who are reported to Child, Youth, and Family concerning abuse or neglect.[11]

Figure 3 Examples of Predictive Risk Modeling Variables

Data categories	*n*
Benefit spells by age 2 years	103,397
Unique children	57,986
Outcome variables	**Proportion[a]**
Any maltreatment	0.150
Neglect	0.064
Emotional abuse	0.106
Physical or sexual abuse	0.019
EXAMPLES OF PREDICTOR VARIABLES	**M (range)**
Primary caregiver characteristics	
Age at birth of child (years)	26.937 (15–75)
Number of older children	0.908 (0–10)
Proportion of time on unemployment benefit during prior 2 years	0.148 (0–1)
Prior court-issued CPS reports for other children	0.007 (0–5)
Prior substantiations for behavioral problems for other children	0.008 (0–5)
Substantiated physical or sexual abuse before age 16 years	0.102 (0–10)
Partner characteristics	
Partner of primary caregiver present	0.289 (0–1)
Partner has criminal record	0.037 (0–1)
Proportion of partner time on sickness benefit during prior 2 years	0.027 (0–1)
Prior neglect substantiations for partner's other children	0.018 (0–5)
Prior police family violence reports for partner's other children	0.019 (0–5)
Youth justice referrals for partner before age 16 years	0.059 (0–30)
Child characteristics	
Number of different caregivers for child	1.366 (1–5)
Court-issued CPS reports for child	0.024 (0–15)
Family group conferences involving child	0.011 (0–5)
Prior substantiated reports of neglect of child	0.008 (0–5)
Prior substantiated reports of emotional abuse of child	0.010 (0–5)
Prior substantiated reports of physical/sexual abuse of child	0.002 (0–5)

Note: Calculations are based on merged administrative data provided by the New Zealand Ministry of Social Development. Predictor variables included here are examples of 242 covariates available in the data set. To protect confidentiality of individuals, all maximum values for nonbinary variables were rounded to the nearest interval of 5.
[a]Proportion of spells with substantiated reports by age 5 years
CPS, child protective services

Broadening Perspectives – Expanding the Data Set

The above risk models represent a range of options, from reactive to proactive models, but are not the only tools available in the market today. Consulting and analytics firms like SAS, Deloitte, Mindshare, IBM, and the Public Consulting Group are engaged in the testing, development, and implementation of predictive analytics tools for child welfare services. As these tools are refined, the importance of aligning records across multiple systems has become apparent. In a 2011 study, child welfare researcher Emily Putnam-Hornstein reported that:

"Historically, data concerning children reported for abuse or neglect in the U.S. have been compiled by child protective service agencies and analyzed independently from other sources of information. Yet these data suffer from notable limitations of being both narrow in scope (i.e., containing a limited set of variables) and narrow in coverage (i.e. capturing data for only those children who are reported)."[12]

Putnam-Hornstein and her fellow researchers assert value in combining data sets across systems, such as linking birth records, administrative CPS records, and death records. The addition of unique data from entities outside of state human services agencies could foster a more complete picture of child welfare. As the Commonwealth looks at opportunities to adopt tools to harness big data, health plan data, such as claims data reported into the APCD or collected through partnerships with health plans, may be beneficial to developing a robust data set. MAHP's December 9, 2016 Policy Forum: ***Can Analytic Tools Identify Children at Risk for Abuse?,*** will provide an opportunity for plans, state stakeholders, and analytic experts to further explore the use of predictive modeling and analytic tools, the development of robust data sets on which to conduct the analyses, and legal and privacy concerns that exist around data sharing across multiple systems.

Endnotes

1. U.S. Department of Health & Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2016). *Child maltreatment 2014*. Available from http://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment
2. Casey Family Programs (2015). Predictive Analytics: Applications for Child Welfare. *Practice Digest, 7*.
3. Doug Laney, *3D Management: Controlling Data Volume, Velocity, and Variety*. (February, 2001) Available from: https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf
4. Walker, B. & Fishman, T. (2015). *Rethinking human services delivery: Using data-driven insights for transformational outcomes*. Deloitte Consulting.
5. American Public Human Services Association. (2014). *Analytic Capability Roadmap 1.0 for Human Service Agencies: A White Paper by the APHSA National Workgroup on Integration Analytics Committee.* Retrieved from: http://www.aphsa.org/content/dam/aphsa/pdfs/NWI/FINAL_NWI%20Analytics%20Capability%20Roadmap_4.17.14.pdf
6. Russell, J. (2015). Predictive analytics and child protection: constraints and opportunities. *Child Abuse & Neglect*, 46.
7. Casey Family Programs, 7.
8. Santhanam, Laura. (March 22, 2016). *Can big data save these children?* PBS Newshour. Retrieved from: http://www.pbs.org/newshour/updates/can-big-data-save-these-children/
9. Eckerd Kids. (2016). *Innovations in Action.* Retrieved from: http://www.eckerd.org/about-eckerd-kids/what-were-doing/innovations-in-action/
10. Heimpel, D. (2015). Uncharted Waters: Data Analytics and Child Protection in Los Angeles. *Chronicle of Social Change.* Available at: https://chronicleofsocialchange.org/featured/uncharted-waters-data-analytics-and-child-protection-in-los-angeles/10867
11. Vaithianathan R, Maloney T, Putnam-Hornstein E, and Jiang N. (2013) Children in the public benefit system at risk of maltreatment: identification via predictive modeling. *American Journal of Preventive Medicine,* 45(3).
12. Putnam-Hornstein, E., Webster, D., Needell, B., Magruder, J., (2011). A Public Health Approach to Child Maltreatment Surveillance: Evidence from a Data Linkage Project in the United States. *Child Abuse Review,* 20(4).

Figure 1 Kitzmiller, E. (2014) Allegheny County's Data Warehouse: Leveraging Data to Enhance Human Services Programs and Policies. Available at: http://www.aisp.upenn.edu/wp-content/uploads/2015/08/AlleghenyCounty-_CaseStudy.pdf

Figure 2 Heimpel, D. (2015). Uncharted Waters: Data Analytics and Child Protection in Los Angeles. *Chronical of Social Change*. Available at: https://chronicleofsocialchange.org/featured/uncharted-waters-data-analytics-and-child-protection-in-los-angeles/10867

Figure 3 Vaithianathan R, Maloney T, Putnam-Hornstein E, & Jiang N. (2013) Children in the public benefit system at risk of maltreatment: identification via predictive modeling. *American Journal of Preventive Medicine*, 45(3).