THE ACADEMY FOR PROFESSIONAL EXCELLENCE

SAN DIEGO STATE UNIVERSITY SCHOOL OF SOCIAL WORK

LITERATURE REVIEW:
PREDICTIVE ANALYTICS IN HUMAN SERVICES

PREPARED BY: DR. THOMAS PACKARD, PROFESSOR

FEBRUARY 2016

SACHS
Southern Area Consortium of Human Services

INSPIRING INNOVATIVE SOLUTIONS IN HEALTH AND HUMAN SERVICES

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY

Predictive analytics (PA) and predictive risk modeling (PRM) are receiving increased attention in human service agencies. PA involves “the practice of extracting information from data sets to determine patterns and predict outcomes and trends” (Casey Family Programs, 2015, p. 18). PRM is a specific type of PA focused on using data patterns to identify predictors of risk and assign risk categories based on these patterns to individuals or families. Several PA and PRM models will be reviewed, including those of Deloitte Consulting, the Public Consulting Group, the Case Commons Casebook model, and cutting edge work in New Zealand.

Sources of knowledge which can be used by child welfare agencies include clinical knowledge, actuarial tools, data generated by the agency, new data from PA and PRM, and the use of outside research. An agency can consult the existing literature, collaborate with university researchers, and/or do its own primary research, for example, by using existing administrative data to predict risk and then assessing clients for identified risk factors. Other outside consultants with PA expertise and occasionally their own software systems may also be useful. Issues relating to the use of such consultants will be discussed in this report.

Several existing best practices, including those in Allegheny County, Pennsylvania; Tennessee; Florida; and New Zealand will be reviewed. Brief mention will be made of applications in Los Angeles County’s Department of Children and Family Services, other states, and the field of child support enforcement.

Regarding implementation, any agency considering initiating PA or PRM systems will need to thoughtfully consider issues including the goals of the system, the agency’s existing capacities, and other available resources and how to engage them. Some available consulting firms will be mentioned.

Because implementation of such a system should also be recognized as a major organizational culture change, guidelines for change management will be reviewed. Another major issue in PA and PRM is data sharing and interoperability of different systems. Some challenges and guidelines in this area will be discussed. The report will end with a review of ethical issues and suggestions on how to address them.

II. BACKGROUND AND NEED

The subject of "Big Data" is receiving increasing attention in many fields. Common examples include Amazon and Netflix, which use past data on customers to predict future choices: products which the customer may be interested in. "Harnessing big data" has been declared a Grand Challenge for social work (Coulton, Goerge, & Putnam-Horsntein, 2015). These authors have asserted that:

> there is a growing movement to build and maintain multiagency integrated data systems (IDS) as a permanent utility for the social sector. In most IDS, administrative records from many agencies are retrieved on an ongoing basis, linked at the individual level, cleaned and organized, and made available for analysis. Though these systems are under development, they have great potential to deliver high quality big data with almost unlimited possibilities to yield vital information to transform social policy and practice. (pp. 8-9)

They also note the importance of going beyond standard quantitative data:

> To get the most out of big data, it is necessary to move beyond the reliance on structured data fields and standard statistical models, an area in which the social sector has made some progress. In particular, there is a great deal that can be learned from detailed case notes, assessment reports, or other digital text that are part of agency records. Text mining methods are evolving rapidly and are now being applied to unstructured notes to gain a more complete picture of behaviors. (p. 10)

According to consultants with Deloitte Consulting, a major firm in the predictive analytics field,

> while human services agencies have always collected, stored, and reported a glut of data, the information rarely was readily available for problem-solving or managing day-to-day work. With today's nimble and relatively inexpensive tools for data management and manipulation, however, information and insights that once might have taken a roomful of analysts weeks to understand can be put in front of workers and clients in near-real time (Walker and Fishman, 2015).

While data in human services agencies are now commonly automated, Walker and Fishman (2015) suggest that programs and their data systems often need to be redesigned to meet specific customer needs.

At a recent summit (About Data Analytics Summit II, 2015), it was noted that in the emerging field of Big Data Analytics, "about 70-80% of the data used in organizations is typically unstructured."

III. DEFINITIONS

A subfield of Big Data, *predictive analytics* (PA), including a subarea known as *predictive risk modeling* (PRM), has been applied in many areas of the human services, including child support, health care, Medicaid, juvenile justice, adult behavioral health, homelessness prevention, and child welfare. PA has a variety of definitions. A definition with reference to child welfare is:

> The practice of extracting information from data sets to determine patterns and predict outcomes and trends. Predictive models typically analyze current and historical data to produce easily understood metrics (quantifiable measure that are used to track and assess such as rates of child protection reports or substantiated cases). For example, these scores rank individuals or families by likely future performance, actions or risk. (Casey Family Programs, 2015, p. 18)

Casey further defines PRM as a "specific type of PA focused on using data patterns to identify predictors of risk and assign risk categories based on these patterns to individuals or families" (p. 18).

Davenport and Harris (2007, as cited by APHSA [2014]) provided context to this definition by adding that "analytics involves the collection, synthesis, and analysis of field-specific data that can lead to improved decision-making as a result of understanding underlying patterns and trends" (p. 6). Descriptive reporting of this data includes:

- Standard Reports—What just happened and why?
- Ad Hoc Reports—How many, how often, who, and where?
- Drill Down—Exact root cause; identify the problem
- Alerts—What actions are needed?

Advanced analytics (including PA) "goes beyond the collection and sorting of data to turn the information into data capable of providing future options and predictive capabilities" and includes:

- Statistical Analysis—Why is this happening?
- Forecasting Scenarios—What if trends
- Predictive Analytics—What happens next?
- Optimization—Predict and prescribe the best that can happen (APHSA, 2014, p. 6-7)

Providing some implementation detail, Casey Programs (2015) noted that,

> many [statistical] methods fall under the predictive analytics umbrella, including those typically taught in graduate statistics classes (regression, hierarchical linear modeling)… PA looks at how a combination of predictors impacts an outcome. …. Child welfare researchers, for example, have used PA to examine child welfare outcomes such as legal permanency or high school completion. **The question being examined is what variables predict these outcomes?** Researchers collect information including demographics, risk factors, placement information, familial information, etc. and seek to determine which of these variables, when examined as a collective, tell us the most about achieving legal permanency. (p. 4)

PA is different in some respects from the more commonly used actuarial risk assessments. According to Russell (2015a):

> Predictive analytics is an approach to how we learn from the past. The past is recorded in some way, in data, in text, in information, in case files. Predictive analytics most often uses a computer algorithm to search through all that information, in millions of iterations, to look for patterns, and interactions, and signals. In most cases, what is learned through predictive analytics is formed as one of three tools: a checklist, a decision tree, or a black box. Actuarial risk assessment is an approach that takes what can be learned about a case from other similar cases, and forms that into a weighted checklist. It uses a list of known useful predictor items and says that when a case has more than an average number of those items, that case can be classified as high risk. When a case has fewer than an average number of those items, it can be classified as low risk. Actuarial risk assessments are tested for their accuracy by comparing what outcome rates really result for each risk level.

In a comprehensive discussion of PRM, Vaithianathan, et al. (2012) also made these distinctions:

> While a plethora of "operator driven" risk assessment tools exist (sometimes referred to as "actuarial" risk tools in the literature), these are inadequate for a number of reasons. One concern is that operator driven tools rely on the social worker or frontline agency correctly applying the model. Compliance is dependent upon an agent who is sufficiently trained and motivated to apply the model, and to then respond to the estimated risk. A second concern is that operator driven risk assessment tools are infrequently validated for the population being risk rated. (p. 6)

In summary:

Actuarial Risk Assessment tools-

1. Are "Operator driven" and require frontline staff (using checklists) to enter the variables that are used to predict risk;
2. Provide a coarse classification of risk;
3. Are validated on other populations, often in other jurisdictions.

PRM tools-

1. Use routinely collected administrative data to exploit historical correlations and patterns;
2. Assign a precise risk score, enabling early detection of high risk.
   (Vaithianathan, et al., 2012, p. 37)

Based on their literature review, Vaithianathan, et al. (2012) also found that:

- Actuarial risk models are common in frontline social work and have been increasingly popular because it is believed that they reduce the "cognitive biases" of frontline social workers.
- Critics claim these tools undermine professionalism and could be used to reduce the accountability of frontline staff.
- Interviews with frontline staff reveal that frontline workers often do not adhere to the tools.

- Ensuring fidelity to risk tools requires a consultative approach with frontline staff so that the tools are seen to be complementary to professional judgment and helpful to the work of staff. (p. 28)

While the great potential PRM's use historical data, which may go as far back as birth records, is clear, de Haan and Connolly (2014) noted a concern that PRM tools

> also identify families who may well benefit from support but are not on a maltreatment trajectory - the so called 'false positives' who would not be among those families later identified as mistreating their children. Whilst early identification of families through the use of PRM has the potential to offer opportunities to provide supportive services that could ameliorate future harm to children, it is clear that it also has the potential to mistakenly target and label families as potential child abusers. (p. 86)

Caution has also been suggested an international authority on PA, Rema Vaithianathan, who noted that "we're still in research mode. I worry sometimes that policy makers and practitioners, because they are under so much pressure to do something, just end up adopting things. We need to go slowly" (Hamovitch, 2015, p. 11).

According to another top researcher in the field, Emily Putnam-Hornstein, "One strategy to ensure that the tool is not misused is to restrict access to risk scores, perhaps only allowing the hotline operator and a supervisor to view the results of the model" (Hamovitch, 2015, p. 12). That would prevent caseworkers in the field from being overly concerned due to a high score or ignoring red flags during a family visit because of a low score. Dr. Putnam-Hornstein suggested that "if a hotline model allows the county to identify the top 10 percent riskiest referrals, perhaps the protocol is simply that those referrals cannot be screened out without an investigation and they are assigned to a more experienced worker" (Hamovitch, 2015, p. 12).

This literature review will focus most specifically upon applications in child welfare but will also draw upon experiences in other fields which could be applied in human service organizations.

IV. MODELS AND MODEL DEVELOPMENT

New Zealand Model

Work on PRM in New Zealand (e.g. Panattoni, et al., 2011) has been particularly valuable to those in American child welfare agencies. In introducing their model Vaithianathan, et al. (2012) notes the principal requirements for the utilization of a PRM include:

1. A sufficiently wide net of the target population captured in the systems from which data are harvested;
2. Comprehensive and timely data on risk factors;
3. Risk scores that can be generated immediately; and

   Outcomes that can be predicted with sufficient accuracy. In the case of child maltreatment, it is particularly important that the protocols followed once the risk score is generated are ethical. (p. 6)

Their model begins with the development of a "risk scoring tool," or algorithm that begins when a child enters the child welfare system. They actually recommend that data sources for the algorithm begin before the child is two years old, so that these data can be used to predict system involvement by age five. Their model includes separate algorithms which can "predict each child's risk of having a substantiated finding of neglect, emotional abuse and physical/sexual abuse by age five and behavioural problems by age seven" (p. 8). In their initial analysis, 224 *predictor*, or independent variables were tested, and 132 were selected for inclusion (Vaithianathan, et al., 2013). The *outcome,* or dependent, variable was substantiated maltreatment. Nearly 45% of predictor variables related to the demographics, SES, and histories of the primary caregiver, whereas 37% related to the primary caregiver's partner (present in 28.9% of the times in which a child entered or re-entered the public benefit system). In this model, predictor variables are given different weights that generate a *probability score* for system involvement, using historical data and statistical methods. The algorithm is adjusted each time a child has contact with the system. Findings are used to identify children and families at the greatest risk, so that they may receive targeted services.

They developed their model using a linked data set from the New Zealand Ministry of Social Development. The data, with no individual identifiers, came from the work and benefits and well as child and family health and welfare systems. Family demographics as well as service data were used. Examples of predictor variables are shown in Table 1.

TABLE 1: Predictor Variable Examples (Vaithainathan, et al., 2013, p. 356)

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

Based on the use of this model, Vaithianathan, et al. (2013) concluded that:

- Although a statistical model cannot replace more comprehensive clinical assessments of a child's risk, automated predictive risk models could be cost effectively implemented within a broader array of assessment tools employed at varying points in the trajectory of children's engagement with service sectors.
- The application of an automated predictive risk model has the potential to not only support an upstream shift toward maltreatment prevention activities but also to do so in a cost-effective and targeted manner.
- Research indicates that early intervention programs often yield greater benefits when offered to mothers and families at higher risk compared to those at low risk.
- Risk stratification has the potential to maximize the impact of programs that may vary in effectiveness across populations. (p. 358)

Deloitte Model

According to Deloitte Consulting, "The objective of child welfare predictive modeling is to use advanced analytics to help the agency caseworkers better identify warning signs and signals to help improve the timeliness and stability of re-unifications" (Deloitte Development, 2015). Their model design is essentially a linear scoring engine that assigns lower scores to those child placements where timely and stable reunification is less likely. The predictive, or in the New Zealand model, predictor variables, include all the factors in the case which may impact child safety or well-being, both risk factors and strengths. The target variable measures the desired outcome, such as, for example, the overall success of a child removal (i.e., what the model is designed to predict): whether the child will be reunified with family within 365 days after removal and not be removed again within 365 days after reunification - whether the reunification is both timely and stable.

With reference to the predictive variables, the Deloitte model makes the important distinction between *uncontrollable* and *controllable* variables, with the controllable variables being extremely important, since these are the ones that a worker can, in fact, influence:

> Examples of *uncontrollable variables* found in child welfare predictive models include:
>
> - the parent's age
> - the size of a household
> - income level
> - the number of times the child has been reported to the agency.
>
> Examples of *controllable variables* found in child welfare predictive models include:
>
> - the number of case workers who have been assigned to the child over the lifecycle of the case
> - the number of contacts between the agency and the family after taking the child into foster care
> - the number of completed medical checkups
>
> For things that we can see, but cannot control, there isn't much a child welfare caseworker can do. For example, a caseworker cannot change the fact that a parent only makes $20,000 a year, his age, or where he lives. On the other hand, controllable variables allow caseworkers to affect outcomes by reducing the number of caseworker

> assignments, increasing agency contacts and family member visits after removal, or helping to ensure that medical checkups are scheduled and happen on a more regular basis. (Deloitte Development, 2015)

In the original model, over 200 predictive variables were used, including Allegation, Allergies, Caretaker Strength and Needs Assessment, Child Strength and Need Assessment, Client, Client Relationship, Home Removal, Income, Investigation, Medication, Medical Appointment, Placement Episode, Provider, Psychological Assessment, Reunification Assessment, Risk Assessment, Safety Assessment, and Visits (Deloitte Development, 2015).

Their overall model includes problem identification, data selection, target variable identification, and data analysis. The predictive modeling in the analysis uses the scoring engine mentioned above to conduct a statistical analysis to create a numerical score. The model begins with a base model which uses only the uncontrollable variables that cannot be impacted by case worker efforts. Then, the final model adds variables that case workers or the agency can impact. Controllable variables are extremely important; first, of course, because these are things that can be impacted by the agency, and second, because their inclusion improves "model performance." The model not only identifies the cases which have a lower success rate, but also gives caseworkers, supervisors, and management staff insights into ways to improve the success rate through actions taken on the controllable variables in the model. For example, the model may show that reducing caseload size, increasing contacts with the family, or medical checkups on schedule lead to better outcomes.

Their research found "a caseworker, acting on factors they can control, can positively identify and impact an additional 20% of families and children on both ends of the spectrum. Caseworkers, with access to the kinds of foresight offered by including controllable variables inside predictive models can make more targeted decisions and expect to have a greater impact on outcomes" (Deloitte Development, 2015).

Behavioral Health Outcomes Management System

A similar process was used developed a model to predict resiliency for youth in the child welfare system (Toche-Manley, et al., 2013). Using a population of youth in out of home placement, youth self-assessments and reports from caseworkers and caregivers, and a clinician assessment were used to identify variables that may be related to resilience. With several clinical assessment instruments available, their original item pool included 243 self-report items and eight summary scales.

> Items included demographics, youth strengths, family and interpersonal relationships, past/current stressors, PTSD, ADHD, Depression/Anxiety and other conditions. Univariate analyses and scatterplots showed relationships between each item and change in resiliency. Items with statistically significant relationships were considered first, with others added later. Items having the strongest correlations with the dependent variable were retained for further analysis. This process reduced the number of variables to less than 100. Multiple regression analyses were used to find the best set of predictors. … In each analysis, days since intake and the initial Resiliency Scale score were entered in a single block. A second block, including up to 12 additional variables from the pool, were evaluated using stepwise regression. Variables from the second block that entered the equation were retained for further testing. This process continued in an iterative manner until the final set of 11 predictors was identified.

These 11 predictors were:

- Intake Resiliency Scale
- Days since intake
- Trying to change life (AVA)
- Does some things well (Resiliency)
- Others concerned about drug/alcohol use (Substance Screen)
- Concerned about own drug/alcohol use (Substance Screen)
- Uses drugs/alcohol to change mood (Risk Factor)
- Feels has caused trouble to parents (Parent-Child Relationship)
- Feels wouldn't be liked if known (AVA)
- Feels on edge, and (PTSD)
- Feels cranky or grumpy (PTSD)

> After the item pool was reduced, a standard multiple regression analysis was performed between change in resiliency as the dependent variable and all eleven predictors… The resulting prediction line shows the expected change in resiliency given the current "mix" of services offered at that site. This information can be used to identify treatment goals and to target appropriate additional services as necessary (p. 8-9).

This process can be seen as an example of using existing data to assess (through later scores) the effectiveness of interventions and assist caseworkers in determining services to be delivered.

Public Consulting Group Model

The Public Consulting Group (Hussey & Shutt, N.D.) has a process with the following steps:

A. *Focusing questions to understand the problem:*

- What is the problem you are trying to solve; or the outcomes you are trying to achieve?
- Know the business.
- Focusing questions give you a starting point to help you identify the data you need to analyze.
- Clarifying questions are generated from your initial data analysis and may require additional data.
- This is achieved through structured surveys, interviews, and focus groups.

B. *Gathering data:*

- Inventory relevant data sources
    - Data locations
    - Data ownership
    - Policy/privacy issues
- Identify specific data elements/sets
    - Data architecture
    - Data models
    - Data dictionaries

C. Analyze the data:

- Data management maturity
  - Do you have good quality data
  - Has the data been validated/cleansed
- Create crosswalks or relationship maps to determine usage scenarios
- Statistical analysis
- Use of tools
  - Automate formulas and calculations
  - Merge data

D. Action:

- Apply that meaning when we create data displays and presentations that answer the questions
- Make informed decisions that will lead to improved outcomes

Case Commons Casebook Model

The Case Commons approach, the Casebook Model (Case Commons, N.D.) recommends as an early step:

> *To develop a clear understanding of your business objectives.* This involves first identifying some of your primary objectives for a case management system. Examples can include:

- Workers being able to efficiently conduct day-to-day tasks
- Meeting federal reporting requirements
- Receiving appropriate federal and state reimbursement
- Integrating more than one human service delivery system, such as child welfare, mental health, juvenile justice, and others
- Collecting longitudinal data for better policymaking and practice
- Ensuring data is exchanged between systems in a way that doesn't hinder users

The Case Commons "Designing an Engagement and Implementation Process" includes developing a budget and timeline and "assembling a decision-making team from your agency that includes caseworkers, supervisors, administrators, policymakers and technologists" (p. 9). Finally, "Mapping Solutions and Configuring Your System" can occur (Case Commons, N.D.).

Recommended Steps for Developing and Implementing a PA Model

After a PA model is chosen, Walker (2013) suggested that those beginning to design a PA system ask the following questions:

- Is our agency driving data and the resulting insights down to everyone who can do something to affect the outcomes of the case –frontline workers, providers, biological and foster parents, and even children, when appropriate?
- Do we fully understand what data are available to us and what that data are saying? Are we exploring what certain patterns and relationships can mean?
- Are we driving resources to what we have learned and moving resources away from what we have always done?
- Are we mining data in good times as well as bad times?

- Are we using data as an ongoing source of information to both challenge and inspire the workforce – to help remove the ambiguity that often surrounds decision-making at the frontline? Are we willing to create a forum where data sit at the center of an honest and straightforward dialogue between leadership, supervisors, and the frontline?
- Are we willing to use data in ways that can hold us more accountable for results at the level of individual children and not simply track trends and aggregate statistics? (pp. 24, 33)

Regarding ways to organize data, Russell (2015b) offers three common model output formats:

- *Decision trees* can be a visually appealing way to portray a model. For each tree, every case starts at the root and at the first branch; some cases split off according to some criteria, and at the next branch others are split again. Branches split into even smaller branches, with fewer and fewer cases going down smaller branches. Tree models are especially good at accounting for complex relationships, interdependencies, and nonlinear combinations….
- *Checklists* assign different weights to particular factors or variables to predict an outcome. Checklists are the form most actuarial models take. For example, factors that can help predict an auto insurance claim might be owning a hot rod car, a history of past accidents, and living in an area with poorly maintained roads. The model could predict that someone with two or three of these factors would be more likely to have a future insurance claim than someone with none or one of these factors. Checklist models are especially intuitive and useful when clarity is a priority.
- A third common predictive analytics model form is a *black box*. A model is called a black box when the computer algorithm making the prediction is so complex it cannot easily be interpreted. This type of modeling can be successful at accurately assigning likelihood scores, but the reasoning behind how it works may not be clear.

American Public Human Services Association (APHSA) (2014) has suggested these points to bear in mind when designing a system:

1. Regardless of where or how data are used, the cleanliness of the data requires constant vigilance.
2. Data definitions can vary, so ensuring data transparency is critically important.
3. Inconsistencies in data can be due to various factors and, to the extent possible, should be explained to the user of the data.

Finally, in planning to implement a PA system, Casey Family Programs (2015) suggested considering the following:

- *Clearly communicate complicated analyses*.
    - Not all of the analyses that fall under the PA umbrella are complicated, but some are. Regardless, when findings from analyses are presented (i.e. reports, journals or presentations), they must be made accessible to disparate audiences. That is not to say audiences must understand the complicated mathematics involved in algorithms or advances statistics, but a general overview of what is being done must be provided in an understandable format.

- *Implement adaptive analyses.*
  - It is important that as more information becomes available (e.g., variables and/or new participants are added), that analyses are adaptive—in other words, they learn from new information. Basically, analyses need to be re-run as new information becomes available.
- *Have a plan.*
  - Jurisdictions must be prepared to act on findings. When engaging in PA, jurisdictions should have the end in mind. In other words, jurisdictions must be prepared to take action, otherwise they are engaging in a strictly academic exercise. The idea is not to conduct PA, rather the idea is to use PA to serve youth more effectively. For example, Georgia's Cold Case Project[1] didn't just run analyses, managers created teams to act on findings. PA can be used to tailor services, support decision making, and for a variety of other uses. (pp. 5-6)

V. SOURCES OF KNOWLEDGE

PA will be an increasingly important source of knowledge to inform child welfare practice. To assist child welfare workers, their supervisors, and their research and program design staffs, clinical knowledge will continue to be essential, as will other methods including Structured Decision Making and other actuarial methods for risk assessment. These use weighting of items in the tool, and are eventually interpreted by the worker. Beyond those sources, practitioners may use knowledge from the literature, including journal articles, conference presentations, and papers by professional and capacity building organizations such as Chapin Hall and Casey Family Programs. For example, studies, some summarized here, have noted the importance of factors ranging from birth records and data from behavioral health data systems to the agency's own system which can indicate higher risk.

A child welfare agency can consult the existing literature, collaborate with university researchers, and/or do its own primary research, for example using administrative data such as CWS/CMS (California's version of SACWIS) in assessing clients for risk factors. Other outside consultants with PA expertise and occasionally their own software systems may also be useful. Issues relating to the use of such consultants will be discussed later in this report.

This report will now review some of the current research and best practice models that may offer guidelines for agencies wanting to design PA or PRM systems, including information on variables for agencies to consider. Then, in case an agency may want to use outside resources, some of the firms offering PA expertise in child welfare will be mentioned.

---

[1] A good example comes from the Georgia Cold Case Project where a set of variables was used to predict whether youth would exit care without a permanent family. The cold case project put together teams, which included specially trained lawyers that helped overcome barriers preventing youth from having a permanent family. As the project experienced success, the population changed. As analyses were re-run, a different set of variables became important for predicting whether a youth exited care without a permanent family. These new variables were addressed by the cold case team. Had they relied on the old information, the cold case teams would have been addressing the wrong issues for these youth. As the population changed, the analyses changed, which dictated a change in strategy for serving youth (Casey Family Programs, 2015, pp. 5-6).

VI. EXAMPLES OF CURRENT APPLICATIONS

Allegheny County, Pennsylvania

A major national example of the use of PA is in Allegheny County, Pennsylvania. In response to major administrative, workforce, and service delivery challenges, crystallized by a child death in 1994, Allegheny County embarked upon major change initiatives their Department of Children and Youth Services (Smith, N.D.). Their basic approach was to "use data already collected about each child to evaluate the possibility of future adverse outcomes in the child welfare system" (Vaithianathan, et al., 2015). In addition to other major organizational changes, they created a Data Warehouse of more than 17 internal and 10 non-DHS data sources, including data regarding child welfare, behavioral health and intellectual disability, AOD, aging services, community services (e.g., Head Start, homeless assistance), public schools, corrections, and the Department of Public Welfare. There have been challenges. For example, criminal justice data were available by month, not year, so data needed to imported and grouped by year. It has also been difficult identifying the outcomes to use (e.g., ending up in foster care, near-fatal abuse) (Emily Putnam-Hornstein, personal communication, Nov. 9, 2015).

After the system harvests the data, a Risk Score is produced for each child. A score from 1-20 indicates how likely the child is to have a placement in the 365 days following the index call. At birth, a score could be used to predict a case opening within 3 years (Vaithianathan, et al., 2015).

If the score is at a high risk level, an in-person risk assessment can be done. Starting in February, when a call comes into their child abuse hotline

> call screeners will enter the alleged perpetrator's name, address, or Social Security number into a computer. The computer will then instantly search dozens of records for that person and others in the household. The person will then be assigned a higher 'risk score' if they or others in the home have a criminal record, have been in drug or alcohol rehab, have received mental health counseling, have been in prison, or were in the child welfare system when they were young. The system can also access school records, so frequent absences will raise a red flag… Allegheny County could, for example, identify young mothers who have used drugs and are former foster children, factors that make their children statistically more likely to suffer from abuse. Officials could then recommend that they receive in-home nurses' visits. Some say officials could intervene even earlier. Because research has shown that a woman's age, lack of prenatal care, and level of education are among the predictors that her child will be abused by age 5, 'prenatal risk assessments could be used to identify children at risk of maltreatment while still in the womb.' [Putnam-Hornstein & Nedell, 2011, p. 2406] (Levinson, 2015)

Florida

Another influential national PA model began in a county in Florida. Florida's Rapid Safety Feedback (RSF) Tool was developed in Tampa's Hillsborough County after a spate of nine child deaths from 2009 through 2012. The Hillsboro model, developed by Eckerd, a private child welfare services provider, looked retrospectively at child abuse cases to determine risk factors for abuse (Heimpel, 2015).

> Prior to assuming case responsibility in Hillsborough County, Eckerd organized, funded and completed a multidisciplinary quality and safety improvement review of all open cases in the county. From this analysis, two distinct sets of criteria emerged. The first was

> a profile of those cases with the highest probability of serious injury or death occurring. These cases had multiple factors in common, including child under the age of three; a paramour in the home; substance abuse/domestic violence history; and a parent previously in the foster care system. The research and analysis identified nine child welfare practice skills that were critical to ensuring that children in the target cases remained safe. Among these were quality safety planning, quality supervisory reviews and the quality and frequency of home visits. (Eckerd Kids)

One of the most important things Florida did was build a query system into their SACWIS system. Now they can search their database for cases with these targeted characteristics and respond to them differently. They conducted a review of every open dependency case in the Tampa Bay area— 1,500 cases impacting more than 3,000 children. They discovered that parents were not involved in the development of safety plans, safety plans were not individually tailored, and changes within families were poorly monitored (e.g. new boyfriends in the home). Even when workers and supervisors identified what needed to be done, there was no follow through.

The first stage of the project was working with Mindshare to develop a system within SACWIS to search by criteria selected, mine thousands of cases, and identify those that met the criteria. They then reviewed the cases utilizing a safety-focused review tool with nine core questions, to hone in on the real issues facing families.

Mindshare produced a system for caseworkers to make daily decisions on their active cases. With access to data from the state system, certain school boards, and the Department of Juvenile Justice, Mindshare produced real-time dashboards that identify children who are at high risk for re-entering care, being re-abused, leaving with no diploma, and aging out.

Mindshare’s role is to look at the operational data in the SACWIS system and use that data to make a prediction about each case. At that point, they turn the process over to Eckerd, who applies Rapid Safety Feedback. The model can then be re-run based on the services provided to determine whether the services are having an impact on the original risk factors. Once they have total access to the data, Mindshare can start producing output in about six weeks. The models are easily replicated; they just need to be populated. The cost has been between 5 and 8 cents per child, per day. Right now the project is using just two data sources, but they originally identified 24. Ideally, they would get MOUs with all of them and add them into the system to get a better picture.

> Predictive analytics can be applied in states without a SACWIS system, as well. .. The state is not dependent on contractors; Rapid Safety Feedback is an internal process, although they do have ongoing technical assistance as they make the shift. The state’s title IV-E waiver has been critical in providing Florida with the flexibility to tailor programs by community. Secretary Carroll argued for even more flexibility, across all states. (Commission to Eliminate Child Abuse and Neglect Fatalities, 2014)

Since RSF was launched in 2013, Hillsborough County has been spared many child deaths. While the addition of a sizeable number of child abuse investigators in 2014 and the anomalous nature of child deaths makes the absence of child death impossible to attribute to RSF alone, the tool has caught the attention of other jurisdictions. Eckerd is working to apply its PA model to Connecticut, Alaska, Oklahoma, Nevada and Maine (Heimpel, 2015).

New Zealand

The research discussed above in the Model section (Vaithianathan, Maloney, et al. 2013) was from a study to "explore the potential use of administrative data for targeting prevention and early intervention services to children and families" (p. 354). The authors believed that "determining a child's risk of maltreatment at or shortly after birth provides an opportunity for the delivery of targeted prevention services" (p. 354). They used a data set of "integrated public benefit and child protection records for children born in New Zealand between January 1, 2003, and June 1, 2006…to develop a risk algorithm using stepwise probit modeling" (p. 354).

Their final model included 132 variables. Nearly 48% of the children who were in the top 10% of those predicted to be at risk were substantiated for maltreatment by five years of age. They also found that "of all children substantiated for maltreatment by five years of age, 83% had been enrolled in the public benefit system before two years of age" (p. 354). While their analysis showed the great potential of developing risk scores, they added that "although a PRM cannot replace more-comprehensive clinical assessments of abuse and neglect risk, this approach provides a simple and cost-effective method of targeting early prevention services" (Vaithianathan, Maloney, et al., 2013, p. 354).

According to Emily Putnam-Hornstein, who has collaborated with the New Zealand researchers

> New Zealand's predictive risk modeling is based on incredibly rich, integrated data. Models were developed based on hundreds of variables spanning data collected over time and across systems. And these models were then used to divide or stratify children into 'risk-deciles,' (i.e. ten groups of equal frequency) based on the likelihood the child would be substantiated for abuse or neglect during the first few years of life. And the findings are promising. (Casey Family Programs, 2015, p. 11)

Los Angeles County Department of Children and Family Services

In 2014, the Los Angeles County Department of Children and Family Services contracted with SAS, the world's largest private software firm, to test out risk modeling. The experiment, dubbed AURA, or Approach to Understanding Risk Assessment, tracked child deaths, near fatalities and "critical incidents" in 2011 and 2012. The firm called these very rare and tragic happenings "AURA events" and looked six months back in the histories of those children and families to find reports of abuse, which they called "AURA referrals." Using a mix of data including, but not limited to: prior child abuse referrals, involvement with law enforcement, as well as mental health records and alcohol and substance abuse history, SAS statisticians created a risk score from one to 1,000, wherein high numbers demark high risk. The next phase involved applying those risk scores to DCFS referrals in 2013 to gauge if AURA was any good at identifying which kids were most likely to be victims of severe and even deadly abuse. The Project AURA Final Report, a PowerPoint presentation created by SAS and dated Oct. 14, 2014, stated that if the department had used the tool in 2013, it would have "enabled a significant reduction in the number of tragic outcomes" (Heimpl, 2015).

Indiana

In 2014, Indiana launched an initiative using PA along with other IT resources including laptops for caseworkers (Goldsmith, 2015). In 2012, the state implemented a child welfare case management system called Management Gateway for Indiana's Kids (MaGIK) using Casebook software, a spinoff of the Annie E Casey Foundation. This initiative, Case Commons, is a nonprofit corporation dedicated to improving outcomes for vulnerable children and families (Feely, Ebendorf, & Hollen, 2015). In July 2012, Casebook went live as a core element of the Indiana Department of Child Services' (DCS) MaGIK system, replacing their legacy child welfare system. A survey of DCS users in 2013 found that the majority of survey participants agreed that they were better able to serve children and could make better decisions helping make children safer. While this is not strictly a PA system, it appears to be comprehensive and useful in monitoring clients and service provision.

Rhode Island

In Rhode Island's Department of Children Youth and Families (Northeast and Caribbean Implementation Center, 2010), "a monthly Data Analytic Group meets and analyzes SACWIS data as well as data that service providers submit. This consists of primarily survival and predictive analyses examining the potential of returning to the system. If a negative trend is identified, the head of the QA office, who is a member of the executive team, can bring the issue to the weekly executive team meeting for discussion and possible interventions." (p. 5)

Other States

"Following success with reduced child fatalities in Florida, Eckerd is donating its new predictive tool to Connecticut's Department of Children and Families for one year. States such as Alaska and Maine are also working with Eckerd to improve their decision-making for children under investigation" (Woods, 2015).

The Nebraska Foster Care Review Office (2014) has recommended that the state acquire PA tools.

Wisconsin has developed a predictive model to identify children at highest risk of re-entering substitute care (Fuller, N.D.). Massachusetts is also considering this approach (Levinson, 2015).

Child Support Enforcement

Deloitte Consulting and others have applied PA in the child support enforcement field. In one case (Deloitte Development, 2015) to improve child support collection, a payment score calculator identified non-custodial parents who are at the highest risk of failing to pay 80% or more of their child support obligations over the next three months. This allows case workers to calculate a score for each case based on a diverse mix of predictive variables, including demographics (e.g., cases, acknowledgment of paternity, unemployment, incarceration, and employment history), payment history, financial data (e.g., total arrears owed, monthly support obligation), enforcement data (e.g., wage withholding, driver's license suspension, passport denial), and case data (e.g., case type, medical support order) (Richard, et al., 2014).

Case management includes a Performance Improvement Module (PIM) which helps county caseworkers sort cases based on the model score, and view the top three reasons for a particular case's score at the time a support order is established. Caseworker actions to help prevent a case from becoming delinquent include rapid follow-up conferencing, phone call reminders, and mandatory job searches. This has increased quantity and frequency of collections against child support orders, improved relationships with defendants through more effective meetings and new methods of outreach, and improved operational efficiency and process improvement through more strategic case assignment (Deloitte Development, 2015).

VII. OTHER RELEVANT RESEARCH OF APPLICATIONS IN CHILD WELFARE

Putnam-Hornstein, Webster, et al. (2011) framed the thinking on broadening perspectives on data use beyond child welfare sources by considering a public health perspective. They reported that:

> Historically, data concerning children reported for abuse or neglect in the US have been compiled by child protective service agencies and analyzed independently from other sources of information. Yet these data suffer from the 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) (p. 256).

Putnam-Hornstein, et al. (2013) showed the value of aligning records across systems when they linked birth records, administrative CPS records, and death records, and found that children with a previous allegation of physical abuse sustained fatal injuries at 1.7 times the rate of children referred for neglect.

In a related study, Putnam-Hornstein and Needell (2011) examined variables including birth weight, prenatal care, birth abnormality, maternal education, total number of children born to the mother, and birth payment method and found significant differences in the unadjusted rates of reported maltreatment for nearly all of the variables.

Another study discussed the use of official child maltreatment data, emergency department and hospitalization data, death certificates, and data from child death review teams, and how integrating this information can advance efforts to protect children (Putnam-Hornstein, Wood, et al., 2013).

Individual, family, and community-level factors have been suggested as explanations of foster care entry rates and average lengths of time that children remain in foster care. They do not, however, provide a sufficient explanation of the substantial geographical variation in entry rates and average lengths of stay across the United States. In a study by Russell and Macgill (2015), a data set of 104 state-level variables was constructed to help answer the question of what accounts for geographic differences in foster care entry rates and average lengths of stay in foster care. A predictive analytics approach (classification and regression trees) was used to sort through all the potential explanatory variables, their interactions, and combinations. The results show that state cultural orientations and socioeconomic facts together best explain foster care entry rates. In contrast, child welfare policy and practice differences together best explain average lengths of stay in foster care. Thus, interventions aimed at goals relating to who goes into foster care and how many children go into foster care might be most effective if they focus on culture and socioeconomic facts. Interventions aimed to change lengths of time in care, on the other hand, might be most effective if targeted at state child welfare policies and practices.

While using data from sources such as birth records and behavioral health has been very relevant and useful, Nguyen (2015) has noted that attention can also be paid to larger economic factors in a community as they can impact child well-being:

> The conventional logic supported by research and statistics suggests that there will be more child maltreatment as the economy becomes worse and less child maltreatment as the economy becomes better. However, in some local jurisdictions in California, statistics indicate the opposite. A closer examination of one county, San Mateo, suggests that this may be due to the fact that the County has a very high Self-Sufficiency Standard in which people get jobs with incomes that do not exceed the Standard, but in fact disqualifies

> them from the safety net of Federal benefits. Further, children born around the time of the last recession have a higher chance of adverse mental health issues and are now entering schools with issues that may reflect child abuse and neglect. (p. 1543)

VIII. OTHER RESOURCES

An agency wanting to embark upon a PA or PRM initiative will need to address many of the questions mentioned prior, develop a model, and assess internal capacity to implement.

APHSA (2014) has developed a capability assessment model which can help an agency decide what type of system is needed, from Basic reports to Predictive Modeling and Optimization. These types are further delineated as Basic, Advanced, and Leading levels. The uses being considered for child welfare in PRM are likely to be at the Advanced level, which is when "basic multi-variable models are developed, monitored, and modified as necessary using regression and extrapolation techniques, and the most current data available are used to operate the model". Beyond this, is the Leading level in PRM which is when "a data scientist reviews the data available, internally and externally, structured and unstructured, then assesses what is meant by each piece of data. Data are enhanced using univariant and bivariant investigations with subject matter experts. Collectively, questions are developed, the precise answers to which are extremely important to the organization. Then, through a series of iterative regression analysis using many variables, an algorithm is developed and validated, and when applied, the best predictor of the question's answer is known. The algorithm is periodically reviewed, retested, and updated" (p. 9-10).

Agencies who are considering the use of consultants should be sure to be educated and clear on what they are buying. Ideally, the agency should own the tool, so that it is updated as needed, to change practice or change weights of risk factors (Emily Putnam-Hornstein, personal communication, Nov. 9, 2015).

Counties are likely to have some internal capacity for developing and using PA systems. <u>To the extent that external consultants may be needed, APHSA (2014) has suggested these steps:</u>

1. *Decide upfront how much analytical power you really need.* The Capability Assessment Model for Analytics is a great place to start. It lets you analyze where your agency or program stands today and where you want it to go.
2. *Collaboration, cooperation, and standardization of requirements across departments are important to ensure that all participants know what data they have and what they need to get.* If your agency plans to seek outside assistance through a Request for Proposals (RFP), make sure there are no overlaps, redundancies, or conflicts in your requirement specifications. Obtaining input from staff on this will enable the organization to write a clear and concise description of its needs.
3. *Decide in advance what a vendor proposal should address*: Is it only the analytic capabilities provided by a specific application, or a soup-to-nuts approach, including data preparation, training, installation, configuration, and maintenance? While a specific application can address particular functions, a solution can remove the burden to figure out what needs to be done to get your organization to an analytics-ready state.
4. *Try to avoid describing in legacy thinking what your organization wants to achieve*. If you use the lexicon of older systems and/or technology, vendors may be hesitant to

offer a more current solution or, worse, not understand what is being asked for. One way to address this problem is to make use of the Capability Assessment Model described earlier. By putting your agency needs in the larger context of your overall objectives, the vendor will have a clearer understanding of not only your organization's immediate concerns, but where their solution should fit within the broader picture.

5. *Include a use case in your RFP that makes sense to your community of stakeholders as an evidence-based example.* By so doing, a prospective bidder is provided with a sense of the current state and may be able to address it in very specific terms, perhaps as a demonstration pilot, thus ensuring that both sides clearly understand what is involved.
6. *Describe your organization's data sources so as to provide bidders with a realistic frame of reference.* Analytics consumers of all kinds have repeatedly said that getting data ready to process with analytics is the hardest part of the journey. Each program within and across the organization collects a variety of data—many times this means that the data elements are defined differently and are from a variety of places. Lack of standard data definitions, identification of such sources, and the processes by which they are collected can hamper enterprise-wide efforts to further its analytic capability. Setting realistic expectations with the user community is a critical first step. Being forthcoming about the quality and status of the available data will confirm to the anticipated user base that they will be able to trust the data as complete, accurate, and current once these issues have been addressed. Without such trust, the results of the most sophisticated analytical application will fall victim to the old adage, "Garbage in, garbage out."
7. *While identifying possible data sources does not necessarily translate into the data being ready to use, prospective bidders may be able to suggest ways in their responses on how to turn them into a useable state.* (p. 11-12)

One capacity that an organization will need to assess is its staff capabilities in statistics. Correlations and regressions are commonly used methods. Software systems with these capabilities include SAS Analytics, PASW Statistics 18/SPSS, and SAP Predictive Analytics (http://www.predictiveanalyticstoday.com/top-predictive-analytics-software/). As noted earlier, if an agency does not have these internal capacities, a partnership with a local university may be mutually beneficial: the agency could get directly relevant statistical and analytical expertise, and university researchers could have opportunities to advance their own research agendas through publications and conference presentations.

The APHSA document also has extensive guidelines on drafting an RFP for consulting services.

For an agency considering outside consultation, several firms doing this work are mentioned below.

- *SAS*
    - The common engine behind the PA work in Florida, Connecticut, and Los Angeles is the analytics software called SAS, developed by the SAS Institute. The SAS Institute is not a traditional child protective services partner. SAS began in 1976 as researchers at the North Carolina State University sought to improve agriculture crop yields…. Business analytics is a $14.4 billion a year business

globally, dominated by market leaders like SAP, Oracle, and IBM. SAS faces strong competition from these companies. There is even a free, open source software package similar to SAS called R, which draws away paying business. (Woods, 2015)

- *Deloitte*
    - As noted prior, Deloitte Consulting is very active in the child welfare field with its Advanced Analytics and Modeling (AAM) Practice.

- *Mindshare*
    - Also as noted prior, Mindshare (http://www.mindshare-technology.com/child_welfare.php) has worked with a number of agencies in Florida including Eckerd Kids. This system uses SAS, and brings together data from:
        - county school board systems
        - accounting and financial systems
        - prevention services data sources
        - 2-1-1 services and referrals
        - state and federal medication listings
        - medication black box warnings
        - court documents and county sexual offender predator data sources

      regardless of location or format. TACF correlates the disparate data to produce a previous, current and predicted synopsis of individual case and child profiles. (Mindshare)

- *IBM*
    - IBM Cúram Solution for Child Welfare offers an "outcome management function" which provides caseworkers with the information they need to analyze the challenges that face children and families. A dynamic assessment and decision-making framework offers the flexibility to link to outside assessments, build in existing assessments or integrate add-on packages." (IBM Corporation (2015, p. 2).

      The core Structured Decision Making assessments have been built into this package.

- *The National Council on Crime and Delinquency*
    - NCCD has done extensive work in this area:
        - NCCD's analytics services are used by agencies committed to using data to drive and describe decision making at the organizational level. We combine data analysis expertise with extensive content knowledge to produce insights that will be impactful for your organization and your clients. (NCCD)

- *Case Commons*
    - Case Commons, also mentioned earlier, is a nonprofit corporation dedicated to improving outcomes for vulnerable children and families. Its Casebook model has been applied in Indiana as discussed above.

- *Public Consulting Group*
    - The Public Consulting Group works with software company NTELX to implement a holistic approach to complex decision analytics that:
        - links together both Medicaid and human services agency utilization data in order to produce a full picture of all of the health and human services;
        - monitors, analyzes and manages key performance indicators; and
        - plans, analyzes and researches platforms to support client agency programs, policies and initiatives to improve the quality of organizational performance and work processes. (Hussey and Shutt, N.D.)

IX. IMPLEMENTATION ISSUES: CHALLENGES AND GUIDELINES

An issue of the publication *Child Welfare 360°* covered many aspects of the use of information technology in child welfare, including using data for child welfare system improvement, and decision support technology (Center for the Advanced Studies in Social Welfare, 2011). Particular attention was given to the issue of organizational culture change to value the use of data in decision making (Webster, Putnam-Hornstein, and Needell, 2011).

Beyond the content of the actual PA or PRM system which is being developed, great attention should be paid to the *process* of implementation. Such a change will, for many agencies, be a major change in organizational culture. There is growing literature on the implementation of evidence-based practices and implementation science; and implementation of PA or PRM systems would be within this arena. Only a few sets of principles for change management will be included here. These are all based specifically upon PA or PRM implementation.

Deloitte consultants (Bingham, et al., 2014) have noted these concerns regarding the implementation of advanced analytics:

- *Executive Ownership:* Without buy-in from senior leadership and a clear corporate strategy for integrating predictive models, advanced analytics efforts can end up stalled at model development.
- *IT Involvement:* Failure to involve IT from the very beginning of the analytics journey can lead to significant issues down the road if technology gaps and limitations aren't understood up front.
- *Available Production Data vs. Cleansed Modeling Data:* Access to historical data for model development is very different from access to real-time data in production, and a strong model is only as good as its ability to be practically implemented within the technology infrastructure. Real life limitations may restrict the data that's available for historical modeling. Sometimes a proxy variable can be used for modeling until the data is available. Analytics initiatives often risk being stymied by the belief that data for modeling must be perfectly clean and organized.

- *Project Management Office (PMO)*: Lack of clear ownership of the end-to-end journey is a common stumbling block for organizations that have struggled (and failed) in implementing their predictive models. Without the right project management structure in place, a clear cadence of project milestones, and the ownership of deliverables by pre-identified business owners, the project could be doomed before it starts. Most importantly, the PMO must be able to connect with all interested parties and adopt an agile approach.
- *End User Involvement and Buy-In:* Lack of end user involvement in the planning, design and ultimate roll out of the predictive models can be detrimental to the efforts. … End users also have more insight into the business process and may be able to better identify potential gaps or roadblocks to successfully incorporate models in day-to-day operations. If the end users feel as if they have a stake in the predictive model roll out, then the company may be more likely to realize the potential financial benefits. If done correctly, some of the early doubters can eventually become analytics advocates.
- *Change Management:* Organizations often fail to understand how predictive models change the current business and technology operations — policies, procedures, standards, management metrics, compliance guidelines and the like. Without the proper design, development and roll out of training materials to address the impacted audiences in the field and home office, the analytics journey can come to an abrupt end. Educating end users and other related stakeholders on how the model will be used on a day-to-day basis, and how their life may change, is important. A communication plan should be developed to answer frequently asked questions (FAQs), address common concerns, and help end users appreciate the strategic vision of the organization. Change management doesn't start and end with training; it begins on day one and lasts well beyond the roll out of the models.
- *Explainability vs. the "Perfect Lift":* It is important to balance building a precise statistical model with the ability to explain the model and how it produces results. What good is using a non-linear model or complicated machine learning method if the end user has no way to translate the drivers of the score and reason codes into actionable business results? Experience shows that a less complex statistical model development method yields results similar to those from more complex approaches, and a small sacrifice of predictive power can result in marked improvement in the explainability of technical model recommendations for end users.

Russell (2015c) has also noted the importance of forming a data leadership group: "This is a culture change – a move from gut level to using data. Have champions to move the data-driven culture forward."

X. INTEROPERABILITY AND DATA SHARING

Challenges and frustrations regarding the inabilities to share data across systems in the human services have been long-standing. In PA and PRM systems, there are technical issues regarding interoperability of systems and data sharing issues which involve multiple agencies, sometimes with different cultures and missions, working together to develop workable processes and MOUs.

Benefits of sharing across are obvious regarding, for example, child welfare and behavioral health. In addition, Nguyen (2014) has noted another connection: food insecurity may impact a child or family in terms of both behavioral health and well-being. For example, "staff from different systems could utilize interoperability, based on predictive analytics, to identify families who may be eligible for SNAP and have higher risk factors for child maltreatment. They could then work with those responsible for administering SNAP at the local level to use public health approaches to maximize SNAP participation" (Nguyen, 2014, p. 4). Putnam-Hornstein, et al. (2013) have made the same case regarding health records: "surveillance and prevention efforts must be broader than one system and should more effectively incorporate health care systems" (p. 65). Russell (2015d) made a similar point about juvenile justice records: "data in other systems (e.g., court data, juvenile probation data, or public health) may be useful for the child protection agency to have at the time of applying a predictive model" (p. 186).

In 2013, the United States Government Accountability Office (GAO) released a report which examined:

1. How selected states or localities have shared data across programs to improve the administration of human services
2. Challenges state and local human services agencies face in balancing privacy protections with greater data sharing
3. Actions that the federal government could take to help address these challenges (GAO, 2013, p. 1)

Human services agencies in Michigan; Utah; Allegheny County, PA; and New York City (Health and Hospitals Corporation) all were found to use central systems to share client data. Success factors included strong leadership (the most commonly cited factor), "having an organizational structure that houses multiple human services programs under one agency" (GAO, 2013, p. 16) and funding to initially develop the system (e.g. in Allegheny County, the local foundation community provided funds). Outside legal review regarding privacy requirements was also important.

Challenges identified among these human services agencies included:

> Confusion or misperceptions around what agencies are allowed to share, as well as a tendency to be risk averse and overly cautious in their interpretation of federal privacy requirements. … Stakeholders also reported that potential inconsistencies in federal privacy requirements that apply to data sharing across multiple programs are a challenge. In particular child welfare workers have difficulty meeting a federal obligation to monitor and support foster care children's educational stability and performance because of the federal law limiting access to education records without parental consent. An amendment enacted on January 14, 2013, includes provisions to address this issue (GAO, 2013, p. 1).

Other specific challenges included a lack of training, outdated IT systems, a past culture suggesting that agencies should not share data, cumbersome data sharing agreements, and concerns that other agencies will not protect data.

To address identified challenges, stakeholders suggested that federal agencies:

> Clarify federal privacy requirements and consider harmonizing requirements….develop model data sharing agreements and informed consent language that comply with federal

privacy requirements, or providing existing examples… [and] reexamine requirements to ensure more consistent privacy rules for data sharing across human services programs and agencies. The report noted that Federal agencies have some related efforts under way" (GAO, 2013, p. i).

According to Barth (2014), states currently have the opportunity to share birth data with CWS at the discretion of the Secretary of their Health Departments, and Maryland, Minnesota, and Michigan all have a birth match protocol in place. Several Federal initiatives and one by the National League of Cities have also been developed to address these issues. Brief summaries and web links are provided below.

Administration for Children and Families Confidentiality Toolkit

The Administration for Children and Families (2014) released a Confidentiality Toolkit as a product of its Interoperability Initiative. It covers Child Welfare, Child Care, TANF, Child Support, Low Income Home Energy Assistance Program (LIHEAP), and SNAP. Also covered are legal and regulatory requirements, including reference to major Child Welfare programs and related requirements including the Family Educational Rights and Privacy Act (FERPA). Recommendations include:

- form two working groups: a program group and a legal group.
- consider including representatives from the provider community on these working groups and throughout the planning process
- accumulate all of the state child welfare laws and general state privacy laws to determine if there are additional state requirements that must be met to share case information between systems working with the same person

This toolkit is available for download at: https://www.acf.hhs.gov/sites/default/files/assets/acf_confidentiality_toolkit_final_08_12_2014.pdf

National Information Exchange Model (NIEM)

The National Information Exchange Model (NIEM) Human Services (HS) Domain Workgroup was established in early 2015 to provide a combination of programmatic, policy, business, and technical expertise in creating standardized data exchanges for the HS domain by adopting the NIEM. The collaborative workgroup, overseen and coordinated by the Administration for Children and Families, consists of federal, state, local and non-profit organization representatives.

The NIEM is a community-driven, standards-based approach to exchanging information. Diverse communities can collectively leverage NIEM to increase efficiencies and improve decision making. It was started by a handful of organizations supporting state and local government to overcome the challenges of exchanging information across state and city government boundaries…. All 50 states and the majority of federal agencies are using (at varying levels of maturity) or considering using NIEM (NIEM, N.D.)

For more information on NIEM visit: https://www.niem.gov/aboutniem/Pages/niem.aspx
NIEM 3.1 is available for download at: https://www.niem.gov/news/Pages/NIEM-version-3-1-is-Now-Available.aspx

National Human Services Interoperability Architecture (NHSIA)

The National Human Services Interoperability Architecture (NHSIA) proposes "a framework to facilitate information sharing, improve service delivery, prevent fraud, and provide better outcomes for children and families….NHSIA offers a foundation for common understanding, interoperability, standards, and reuse. Objectives include establishing a common vocabulary, providing a business and technical framework, promoting sharing and reuse, encouraging data exchange standards development, developing standard data structures, and improving operational efficiency and effectiveness" (Administration for Children and Families).
For more information on (NHSIA) visit: http://www.acf.hhs.gov/nhsia-definition

Center for Substance Abuse Treatment
A guide developed by SAMHSA "describes the primary data-reporting systems used in the child welfare, alcohol and other drug services, and court systems. The document describes 15 data-reporting systems, including 8 child welfare systems, 5 alcohol and other drug service systems, 2 initiatives to implement a national data reporting system in the courts, and 1 enterprise health information system for data on American Indian and Alaska Native families" (Center for Substance Abuse Treatment, 2011, p. 1).
This SAMHSA Guide is available for download at:
http://search.usa.gov/search/docs?affiliate=samhsa-store&dc=1415&query=11-4630

National League of Cities and Stewards of Change
The National League of Cities and Stewards of Change (Stewards of Change, 2014) prepared a toolkit on data sharing for cities and other organizations. The document covered details in fields including education, health, mental health and drug and alcohol treatment, criminal justice, and human services. They noted several key tasks for agency leaders who engage in data sharing initiatives:

- Taking the lead to get all parties to agree on "why" the systems involved should share data. Designating a team of staff representing the systems from which data will be shared to create the "what," or list of minimally necessary information that needs to be shared for the legitimate governmental purpose to succeed and "who" needs to receive such information.
- Forming a team, including the privacy officials and the information technology staff, to determine "how" to share the information and how to use it once it has been shared. This group also will develop policies and procedures regarding the privacy security and safeguards of the shared information. The result will be enforced by the privacy officials from the affected agencies.
- Formalizing these agreements in a Memorandum of Understanding (MOU) or Memorandum of Agreement (MOA).
- Arranging for extensive training of all members of the workforce on the policies and procedures of the information sharing project once it is initiated and fully implemented.
- Putting in place a system to monitor the implementation and impact of data sharing agreements to determine if they are having the intended positive impact and if not, to make necessary adjustments. (p 12-13)

They also noted the importance of articulating a shared vision as part of the "why".
This report is available for download at:
http://www.nlc.org/Documents/Find%20City%20Solutions/IYEF/Data%20Sharing%20for%20Better%20Results.pdf

XI. ETHICS ISSUES

Ethics issues regarding PRM have been raised in both the professional literature and in media reports.

Christian (2015) raised several ethics questions regarding PRM:

- ***Does PRM violate people's right to privacy?*** PRM involves the use of personal information without the consent of the individual in order to generate an individually identifiable risk score. That information may be shared within the child welfare agency, private service providers and others for the purpose of intervening with the family to prevent child maltreatment that may or may not occur, given that PRM also returns false positives.
- ***When risk is identified, what are the ethical obligations of the child welfare agency?*** What are the obligations of the agency to provide services or to the family if services are refused? From a practical standpoint, will families be likely to refuse services if and when they find out how they were identified and does PRM thus interfere with family engagement?
- ***Does PRM raise issues of due process and fairness to families?*** Should families have the right to contest a risk score in the same way that they have the right to appeal a substantiation of maltreatment?
- ***How do agencies balance the need for transparency in the PRM process against the likelihood that data may be misinterpreted?*** The use of certain demographic and economic predictor variables may reinforce stereotypes and prejudices regarding race, poverty and ethnicity.
- ***Is PRM likely to lead to a more risk-averse, coercive and deficit-based child welfare system?*** Would reliance on PRM make risk the central organizing principle of child welfare, undermining the best practice paradigm that emphasizes family strengths and resiliency?

Casey Family Programs (2015) raised the issue of "'pushback' to the use of PA to target high risk families for voluntary services, namely potential stigmatization of clients and issues of proper response by child welfare agencies" (p. 9). Emily Putnam-Hornstein responded that "if we have services that are not funded at the universal level, I believe we have a moral and fiscal obligation to ensure that available service slots go to the children and families where data would indicate there is the greatest need – or to children and families where data suggest the impact and benefit will be greatest" (p. 10).

Data may have bias in terms of creating possible false positives: assuming that the child is at high risk when she or he is not. One response to this is that if the intervention is beneficial, there is not much risk, but removal may lead to significant consequences. Russell (2015c) has emphasized that it is important to have a good intervention and involve lots of community stakeholders.

Regarding informed consent, in Allegheny County, Pennsylvania, Vaithianathan, et al. (2015) noted that the rights of those who would normally be given an opportunity to consent must be balanced against rights of vulnerable children. This can be managed by appropriate implementation, ensuring interventions are supportive and non-punitive, and managed dissemination/distribution of information (e.g. restricting distribution to supervisors and caseworkers who have received training and thus appreciate what a high risk score does and does not show).

Critics "point out that newer computer systems, like the one about to be launched in Allegheny County, can base their predictions on vast tracts of data mined from the criminal justice, Medicaid, and drug treatment systems. Critics worry that African-American, Hispanic, and poor parents could have their children unfairly targeted for removal or monitoring merely because the computer models suggest their race, incomes, or criminal records make them more likely to commit child abuse in the future" (Levenson, 2015).

Gusovsky (2016) has reported that:

> Critics say rather than protect children and save money, adoption of PA would mean lucrative contracts for tech companies but could lead to racial profiling and predatory behavior on the part of case agents. "Kids deserve better than to be exploited by a piece of software that's providing (at least in their mind) a prediction that we think should be provided by the community itself," said Los Angeles community activist and executive director of Project Impact, Matthew Harris. His fundamental concern about big data is its top-down approach. "The problem with big data, especially in Los Angeles County, is that the pockets of success get overlooked and buried or aggregated by the big data,"

In their comprehensive report, Vaithianathan, et al. (2012) made these recommendations:

- *Consider an ethical framework for how agencies will and will not respond to predictive risk.*
- *Consider how agencies can educate stakeholders and frontline staff about the difference between a risk assessment and substantiation*. Ensure monitoring regime and governance arrangements that can ensure that risk assessment is used appropriately.
- *Ensure stringent confidentiality, transparency and governance is maintained. Ensure those subject to risk assessment are given appropriate support to understand and appeal the process.*
- *Ensure that the benefits are large enough to warrant the stigmatization and the false-positives that might be inevitable from a risk assessment.* Similar issues are dealt with in, for instance, HIV tests or genetic tests for inherited diseases, and these could provide some guidance for the appropriate administrative structures. Ensure steps are in place to mitigate the harm of stigmatization as much as possible, including clear communication and stringent confidentiality.
- *In general, mandatory or imposed interventions should not be considered in response to predictive risk.*
- *Develop an ethical evaluation of the predictor variables that ought not to be used in the algorithm.* Develop procedures to reduce the risk of misuse of data insofar as possible. Clear communication and appropriate confidentiality are again likely to be central features in such procedures.
- *Careful consideration needs to be given on the rights and responsibilities of the agencies, the children and the families.* (p. 32-34)

In a review of ethics issues in PRM, Dare (2013) referred to this broader discussion by Vaithianathan et al. (2012) and made a summary evaluation statement that "the application of predictive risk modeling to child maltreatment does raise significant ethical concerns. Many of these concerns can be significantly mitigated or ameliorated. Remaining concerns may plausibly be regarded as outweighed by the very considerable potential benefits of the Vulnerable Children PRM. In sum, the application of predictive risk modelling to child maltreatment is ethically justified provided the recommendations below are addressed" (Dare, 2013, p. 1).

Dare made the following 17 recommendations:

1. That targeted intensive preventive intervention is offered to children identified as at high risk of maltreatment. [Regarding under- and over-identification], as with any risk prediction tool, the Vulnerable Children PRM will inevitably make some errors at any threshold for referral, identifying as low risk some children who go on to experience abuse or neglect, and identifying as high risk some children who do not.
2. That the databases upon which a child maltreatment PRM draws are expanded to include as many New Zealand children as possible.
3. That current early identification referral routes including those initiated by health professionals and other frontline social service professionals should be maintained alongside the Vulnerable Children PRM.
4. That ways of reducing the consequences of mistaken identification as high-risk are explored, including:
    a. Providing an opportunity for experienced social services professionals to exercise judgment about appropriate responses to a family's identification as at risk;
    b. Ensuring that such professionals understand the potential of the Vulnerable Children PRM to miscategorize families;
    c. Providing training to guard, in so far as possible, against confirmation bias in the professional engagement with families identified as high-risk.
5. That interventions be at the minimum level necessary to achieve the benefits offered by the Vulnerable Children PRM.
6. That information produced by the Vulnerable Children PRM is disseminated as narrowly as possible, consistently with achieving the benefits of the program. In addition that only senior and experienced staff have access to such information and that they be carefully trained as to how to manage the information they possess.
7. That consideration be given to what level of detail is required to be disseminated to make effective use of the model's predictions.
8. That training and implementation emphasize that those identified as at risk have committed no wrong and that most of them will not go on to do so. Interventions must be preventive and supportive and not punitive.
9. That interaction with high-risk families is as similar as possible to that with other families, at least in the external presentation of those interactions.
10. That ways of engaging with the media over child maltreatment and the Vulnerable Children PRM are explored with an eye to minimizing stigmatization and promoting as accurate an account of the PRM as possible.
11. That engagement with high-risk families is on a voluntary basis.

12. That the Vulnerable Children PRM is used as an opportunity to deliver additional intensive intervention to high risk families and that existing universal services remain in place.
13. That child protection resources and workload are managed to ensure response to identified risks.
14. That invasions of privacy which could appear discriminatory are monitored and minimized, consistently with delivering the benefits of the Vulnerable Children PRM.
15. That staff having access to the information provided by the Vulnerable Children PRM are made subject to a specific duty of confidentiality.
16. That the Vulnerable Children PRM is not seen as a replacement for the judgment and engagement of experienced social service professionals. (See also Recommendation 12.)
17. That implementation decisions around the Vulnerable Children PRM identify staff or services that will have responsibility for monitoring the Vulnerable Children PRM and engaging with families, taking into account the ethical issues relevant to those decisions raised in this report.

Finally, those with experience with PRM note that it is only one element of the process, and that clinical judgment is essential in deciding how to use data.

REFERENCES

About Data Analytics Summit II (2015, December 14-16): Structuring the Unstructured: The missing element of analytics. Retrieved from http://analyticssummit.harrisburgu.edu/.

Administration for Children and Families (N.D.). *What is Interoperability?* Retrieved from http://www.acf.hhs.gov/about/interoperability

Administration for Children and Families (N.D.). *National Human Services Interoperability Architecture (NHSIA) Definition.* Retrieved from http://www.acf.hhs.gov/nhsia-definition

Administration for Children and Families (2014). *Confidentiality Toolkit: A resource tool from the ACF Interoperability Initiative*. Retrieved from https://www.acf.hhs.gov/sites/default/files/assets/acf_confidentiality_toolkit_final_08_12_2014.pdf

American Public Human Services Association (APHSA). (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

Barth, R. (2014). *Preventing Child Abuse Deaths Using Birth to CWS Matches.* Presented to The Commission to Eliminate Child Abuse and Neglect Fatalities, Tampa, Florida, July 10, 2014.

Bingham, K. (2014). *Improving Child Support Enforcement and Case Management.* American Public Human Services Association Policy and Practice.

Bingham, K., Lucker, J., Ward, L., & Peterson, S. (2014). The Challenges of Implementing Advanced Analytics. *Claims Magazine*, 24-27.

Case Commons (N.D.) *Making the Casebook Choice: Leading the Way in 21st Century Innovation.* New York, NY: Case Commons, Inc. Available at www.casecommons.org

Casey Family Programs (2015). Predictive Analytics: Applications for Child Welfare. *Practice Digest*, 7.

Center for Substance Abuse Treatment. (2011). *Introduction to Cross-System Data Sources in Child Welfare, Alcohol and Other Drug Services, and Courts.* HHS Publication No. (SMA) 11-4630. Rockville, MD: Substance Abuse and Mental Health Services Administration Retrieved from http://search.usa.gov/search/docs?affiliate=samhsa-store&dc=1415&query=11-4630

Center for the Advanced Studies of Social Welfare (2011). Child Welfare and Technology, *CW 360°.* University of Minnesota. Spring 2011.

Christian, S. (2015). Ethical Issues Raised by One Type of Predictive Analytics: Risk Modeling in Casey Family Programs. Predictive Analytics: Applications for Child Welfare. *Practice Digest*, 7.

Commission to Eliminate Child Abuse and Neglect Fatalities. (2014). Retrieved from https://eliminatechildabusefatalities.sites.usa.gov/about-us/commissioners/

Cortez-Neavel, B. (2015, April 26). Texas Data Analytics, Prevention Efforts Could Drive Down Child Deaths. *The Chronicle of Social Change*. Retrieved from https://chronicleofsocialchange.org/featured/data-analytics-prevention-efforts-could-drive-down-child-deaths/10838

Coulton, C. J., Goerge, R., Putnam-Hornstein, E., & de Haan, B. (2015). Harnessing big data for social good: A grand challenge for social work (Grand Challenges for Social Work Initiative Working Paper No. 14). Cleveland, OH: American Academy of Social Work and Social Welfare. Retrieved from http://aaswsw.org/wp-content/uploads/2015/07/Big-Data-GC-edited-and-formatted-for-committee-review-7-17-20151.pdf

Dare, T. (2013). *Predictive risk modelling and child maltreatment: An ethical review.* Auckland, NZ: The University of Auckland, NZ. Retrieved from https://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/research/predictive-modelling/00-predicitve-risk-modelling-and-child-maltreatment-an-ethical-review.pdf

Davenport, T.H. and Harris, J.G. (2007). *Competing on Analytics: The New Business of Winning*. Boston, MA: Harvard Business School Publishing.

de Haan, I. and M. Connolly (2014). Another Pandora's box? Some pros and cons of predictive risk modeling. *Children and Youth Services Review*, *47*, Part 1, 86-91.

Deloitte Development. (2015). *Deloitte's Advanced Analytics and Modeling (AAM) Practice Driving End-to-End Analytics* [Power Point Deck]. Deloitte Consulting.

Eckerd Kids. (2014). *Innovations in Action.* Retrieved from http://www.eckerd.org/about-eckerd-kids/what-were-doing/innovations-in-action/

Feely, K., Ebendorf, B. & Hollen, A. (2015). *Transforming child welfare decision-making through modern technology and data analytics*. National Child Welfare Workforce Webinar, Nov. 4, 2015. Retrieved from https://vimeo.com/144779251

Fuller, T. (N.D.) The P.S. Program: Using predictive analytics in program implementation. Presentation to the Wisconsin Department of Children and Families. Children and Family Research Center, University of Illinois Urbana-Champaign.

Goldsmith, S. (2015, July 7). Indiana leads the way in visualization, mobility, and analytics integration. *Government Technology*. Retrieved from http://datasmart.ash.harvard.edu/

Government Accounting Office. (2013). *Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy (GAO-13-106)*. Retrieved from http://www.gao.gov/products/GAO-13-106

Gusovsky, D. (2016, January 14). Can life as a data point save America's at risk children? *CNBC.* Retrieved from http://www.cnbc.com/2016/01/14/an-80-billion-annual-tax-bill-thats-failing-our-children.html

Hamovitch, P. (2015). *Can big data help keep children safe?* A research publication of the Hamovitch Center for Science in the Human Services at the USC School of Social Work. University of Southern California, Los Angeles.

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

Hickey, K. (2014). Using analytics to reduce child abuse risk. *GCN*. Retrieved from https://gcn.com/Articles/2014/08/04/Predictive-analytics-child-welfare.aspx?Page=1

Hussey, C. & Shutt, B. (N.D.) Comprehensive Child Welfare Evaluation through Data Analytics [Power Point Presentation]. Public Consulting Group. Retrieved from http://www.publicconsultinggroup.com/humanservices/index.html

IBM Corporation (2015). *Transforming child welfare service delivery: IBM Social Programs Solution Brief.* Armonk, NY.

Institute for Child Welfare (2015). *FY 2014-2015 Annual Report*. College of Social Work, Florida State University, Tallahassee, FL.

Levinson, M. (2015, October 7). Computers may spot abuse risks: Welfare agencies look to predictive analytics to find, protect children in the most danger. *Boston Globe*, p. A.1.

Mindshare. *Predictive Analytics*. Retrieved from http://www.mindshare-technology.com/child_welfare.php

NCCD. Retrieved from http://www.nccdglobal.org/analytics/analytics-services

National League of Nebraska Foster Care Review Office. (2014, June 15). *Quarterly Update to the Legislature*.

Northeast and Caribbean Implementation Center (2010). New Jersey Department of Children and Families "Manage by Data" National Promising Practice Findings. Cutler Institute for Health and Social Policy, Muskie School of Public Service, University of Southern Maine, Portland, ME.

Nguyen, L. (2014). A public health response to data interoperability to prevent child maltreatment. *American Journal of Public Health, 104*(11), 2043-2048

Nguyen, L. (2015). The relationship between unemployment and child maltreatment: A county-level perspective in California. *Children and Youth Services Review, 35*(9), 1543-1555.

NIEM (N.D.) *About NIEM*. Retrieved from https://www.niem.gov/aboutniem/Pages/niem.aspx

Panattoni, L., Vaithianathan, R., Ashton, T., & Lewis, G. (2011). Predictive risk modelling in health: Options for new zealand and australia. *Australian Health Review, 35*(1), 45-51.

Putnam-Hornstein, E. (2014). *A population-level overview of child fatalities and child protection involvement: surveillance & risk*. Commission to Eliminate Child Abuse & Neglect Fatalities, Tampa, FL.

Putnam-Hornstein, E. (2011). Report of Maltreatment as a Risk Factor for Injury Death: A Prospective Birth Cohort Study. *Child Maltreatment, 16*(3), 163-174.

Putnam-Hornstein, E., M. A. Cleves, et al. (2013). Risk of Fatal Injury in Young Children Following Abuse Allegations: Evidence From a Prospective, Population-Based Study. *American Journal of Public Health, 103*(10), e39-e44.

Putnam-Hornstein, E. and B. Needell (2011). Predictors of child protective service contact between birth and age five: An examination of California's 2002 birth cohort. *Children and Youth Services Review, 33*(11), 2400-2407.

Putnam-Hornstein, E., Schneiderman, J., Cleves, M., Magruder, J., & Krous, H. (2014). A prospective study of sudden unexpected infant death after reported maltreatment. *The Journal of Pediatrics,164*(1), 142-148.

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), 256-273.

Putnam-Hornstein, E., Wood, J.N., Fluke, J., Yoshioka-Maxwell, A., Berger, R.P. (2013). Preventing Severe and Fatal Child Maltreatment: Making the Case for the Expanded Use and Integration of Data. *Child Welfare, 92*(2), 59-75.

Richard, D., Bean, M., Bingham, K., White, J., and Guszcza, J. (2014). Leveraging Advanced Analytics-The Evolution of Pennsylvania's Child Support Enforcement Program. *Policy & Practice*, 28-31.

Russell, J. (2015a). Is Actuarial Risk Assessment Predictive Analytics? *NCCD Blog.* Retrieved from http://www.nccdglobal.org/blog/is-actuarial-risk-assessment-predictive-analytics

Russell, J. (2015b) *Data Talk: Five Concepts about Predictive Analytics*. NCCD Blog. Retrieved from http://www.nccdglobal.org/blog/data-talk-five-concepts-about-predictive-analytics September 1, 2015.

Russell, J. (2015c). *Three Examples of Predictive Analytics in Child Protection*. Webinar available at https://vimeo.com/14182963220 15-10-08 13.05

Russell, J. (2015d). Predictive analytics and child protection: Constraints and opportunities. *Child Abuse & Neglect, 46*, 182-189.

Russell, J. & S. Macgill (2015). Demographics, policy, and foster care rates; A Predictive Analytics Approach*. Children and Youth Services Review, 58*, 118-126.

Smith, M. (N.D.). *Building an Interoperable Human Services System. Stewards of Change: How Allegheny County Transformed Systems, Services and Outcomes for Vulnerable Children and Families*. Retrieved from http://www.stewardsofchange.com/learning-center/Pages/case-studies.html.

Stewards of Change (2014). *Sharing Data for Better Results*. Washington, DC: National League of Cities. Retrieved from http://www.nlc.org/Documents/Find%20City%20Solutions/IYEF/Data%20Sharing%20for%20Better%20Results.pdf

Tampa, Florida Public Meeting Highlights—July 10, 2014. Retrieved from https://eliminatechildabusefatalities.sites.usa.gov/event/tampa-fl-public-meeting-july-10-2014/

Toche-Manley, L., L. Dietzen, et al. (2013). Revolutionizing Child Welfare with Outcomes Management. *The Journal of Behavioral Health Services & Research, 40*(3), 317-329.

Vaithianathan R, Maloney T, Putnam-Hornstein E, Jiang N, DeHaan I, Dale C, & Dare T. (2012). *Vulnerable Children: Can Administrative Data Be Used To Identify Children At Risk Of Adverse Outcomes?* The Centre for Applied Research in Economics, Department of Economics, University of Auckland. Retrieved from http://www.msd.govt.nz/about-msd-and-our-work/publicationsresources/research/vulnerable-children/index.html

Vaithianathan, R., Putnam-Hornstein, E., de Haan, I., Gambrill, E., Bitler, M., Maloney, T., Jiang, N., and Dare, T. (2015). *Using Predictive Modelling and Data Visualization to Improve Outcomes for Children in Allegheny: Options for Consideration.* Power Point. Auckland, New Zealand: Auckland University of Technology.

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), 354-359.

Walker, B. (2013). Using data to improve the welfare of children. *Policy & Practice, 71*(5), 24-33.

Walker, B. & Fisman, T. (2015). *Rethinking human services delivery: Using data-driven insights for transformational outcomes*. Deloitte Consulting.

Webster, D., Putnam-Hornstein, E., & Needell, B. (2011). Using data for child welfare system improvement: Lessons learned from the California Performance Indicators Project. *Child Welfare 360: Child Welfare and Technology*, University of Minnesota.

Woods, D. (2015, May). Who Will Seize the Child Abuse Prediction Market? *Chronicle of Social Change*. Retrieved from https://chronicleofsocialchange.org/featured/who-will-seize-the-child-abuse-prediction-market/10861

ADULT SERVICES

CHILD WELFARE

EXECUTIVE DEVELOPMENT

INDIAN CHILD WELFARE

BEHAVIORAL HEALTH

ORGANIZATIONAL SUPPORT

6505 ALVARADO ROAD, SUITE 107
SAN DIEGO, CA 92120
(619)594-3546
THEACADEMY.SDSU.EDU