

Can an Algorithm Tell When Kids Are in Danger?

Child protective agencies are haunted when they fail to save kids. Pittsburgh officials believe a new data analysis program is helping them make better judgment calls.

By DAN HURLEY JAN. 2, 2018

The call to Pittsburgh's hotline for child abuse and neglect came in at 3:50 p.m. on the Wednesday after Thanksgiving 2016. Sitting in one of 12 cubicles, in a former factory now occupied by the Allegheny County Police Department and the back offices of the department of Children, Youth and Families, the call screener, Timothy Byrne, listened as a preschool teacher described what a 3-year-old child had told him. The little girl had said that a man, a friend of her mother's, had been in her home when he "hurt their head and was bleeding and shaking on the floor and the bathtub." The teacher said he had seen on the news that the mother's boyfriend had overdosed and died in the home.

According to the case records, Byrne searched the department's computer database for the family, finding allegations dating back to 2008: parental substance abuse, inadequate hygiene, domestic violence, inadequate provision of food and physical care, medical neglect and sexual abuse by an uncle involving one of the girl's two older siblings. But none of those allegations had been substantiated. And while the current claim, of a man dying of an overdose in the child's home, was shocking, it fell short of the minimal legal requirement for sending out a caseworker to knock on the family's door and open an investigation.

Over the course of an 18-month investigation, officials in the county's Office of Children, Youth and Families (C.Y.F.) offered me extraordinary access to their files and procedures, on the condition that I not identify the families involved. Exactly what in this family's background led the screening tool to score it in the top 5 percent of risk for future abuse and neglect cannot be known for certain. But a close inspection of the files revealed that the mother was attending a drug-treatment center for addiction to opiates; that she had a history of arrest and jail on drug-possession charges; that the three fathers of the little girl and her two older siblings had significant drug or criminal histories, including allegations of violence; that one of the older siblings had a lifelong physical disability; and that the two younger children had received diagnoses of developmental or mental-health issues.

Finding all that information about the mother, her three children and their three fathers in the county's maze of databases would have taken Byrne hours he did not have; call screeners are expected to render a decision on whether or not to open an investigation within an hour at most, and usually in half that time. Even then, he would have had no way of knowing which factors, or combinations of factors, are most predictive of future bad outcomes. The algorithm, however, searched the files and rendered its score in seconds. And so now, despite Byrne's initial skepticism, the high score prompted him and his supervisor to screen the case in, marking it for further investigation. Within 24 hours, a C.Y.F. caseworker would have to "put eyes on" the children, meet the mother and see what a score of 19 looks like in flesh and blood.

For decades, debates over how to protect children from abuse and neglect have centered on which remedies work best: Is it better to provide services to parents to help them cope or should the kids be whisked out of the home as soon as possible? If they are removed, should they be placed with relatives or with foster parents? Beginning in 2012, though, two pioneering social scientists working on opposite sides of the globe — Emily Putnam-Hornstein, of the University of Southern California, and Rhema Vaithianathan, now a professor at the Auckland University of Technology in New Zealand — began asking a different question: Which families are most at risk and in need of help? "People like me are saying, 'You know what, the quality of the services you provide might be just fine — it could be that you are providing them to the wrong families,'" Vaithianathan told me.

She and Putnam-Hornstein linked many dozens of data points — just about everything known to the county about each family before an allegation arrived — to predict how the children would fare afterward. What they found was startling and disturbing: 48 percent of the lowest-risk families were being screened in, while 27 percent of the highest-risk families were being screened out. Of the 18 calls to C.Y.F. between 2010 and 2014 in which a child was later killed or gravely injured as a result of parental maltreatment, eight cases, or 44 percent, had been screened out as not worth investigation.

According to Rachel Berger, a pediatrician who directs the child-abuse research center at Children’s Hospital of Pittsburgh and who led research for the federal Commission to Eliminate Child Abuse and Neglect Fatalities, the problem is not one of finding a needle in a haystack but of finding the right needle in a pile of needles. “All of these children are living in chaos,” she told me. “How does C.Y.F. pick out which ones are most in danger when they all have risk factors? You can’t believe the amount of subjectivity that goes into child-protection decisions. That’s why I love predictive analytics. It’s finally bringing some objectivity and science to decisions that can be so unbelievably life-changing.”

The morning after the algorithm prompted C.Y.F. to investigate the family of the 3-year-old who witnessed a fatal drug overdose, a caseworker named Emily Lankes knocked on their front door. The weathered, two-story brick building was surrounded by razed lots and boarded-up homes. No one answered, so Lankes drove to the child’s preschool. The little girl seemed fine. Lankes then called the mother’s cellphone. The woman asked repeatedly why she was being investigated, but agreed to a visit the next afternoon.

The home, Lankes found when she returned, had little furniture and no beds, though the 20-something mother insisted that she was in the process of securing those and that the children slept at relatives’ homes. All the appliances worked. There was food in the refrigerator. The mother’s disposition was hyper and erratic, but she insisted that she was clean of drugs and attending a treatment center. All three children denied having any worries about how their mother cared for them. Lankes would still need to confirm the mother’s story with her treatment center, but for the time being, it looked as though the algorithm had struck out.

commissioner of New York City's Administration for Children's Services, expressed worries about the use of predictive analytics by child-protection agencies. "It scares the hell out of me," she said — especially the potential impact on people's civil liberties. "I am concerned about widening the net under the guise that we are going to help them."

But in Pittsburgh, the advocates for parents, children and civil rights whom I spoke with all applauded how carefully C.Y.F. has implemented the program. Even the A.C.L.U. of Pennsylvania offered cautious praise. "I think they're putting important checks on the process," said Sara Rose, a Pittsburgh lawyer with the organization. "They're using it only for screeners, to decide which calls to investigate, not to remove a child. Having someone come to your home to investigate is intrusive, but it's not at a level of taking a child away or forcing a family to take services."

The third criticism of using predictive analytics in child welfare is the deepest and the most unsettling. Ostensibly, the algorithms are designed to avoid the faults of human judgment. But what if the data they work with are already fundamentally biased? There is widespread agreement that much of the underlying data reflects ingrained biases against African-Americans and others. (Just last month, the New York City Council voted to study such biases in the city's use of algorithms.) And yet, remarkably, the Allegheny experience suggests that its screening tool is *less* bad at weighing biases than human screeners have been, at least when it comes to predicting which children are most at risk of serious harm.

"It's a conundrum," Dalton says. "All of the data on which the algorithm is based is biased. Black children are, relatively speaking, over-surveilled in our systems, and white children are under-surveilled. Who we investigate is not a function of who abuses. It's a function of who gets reported."

In 2015, black children accounted for 38 percent of all calls to Allegheny County's maltreatment hotline, double the rate that would be expected based on their population. Their rate of being placed outside their home because of maltreatment was even more disproportionate: eight out of every 1,000 black

more heavily. If I had a parent who was violent, I might care more about that. What predictive analytics provides is an opportunity to more uniformly and evenly look at all those variables.”

For two months following Emily Lankes’s visit to the home of the children who had witnessed an overdose death, she tried repeatedly to get back in touch with the mother to complete her investigation — calling, texting, making unannounced visits to the home. All her attempts went without success. She also called the treatment center six times in hopes of confirming the mother’s sobriety, without reaching anyone.

Finally, on the morning of Feb. 2, Lankes called a seventh time. The mother, she learned, had failed her three latest drug tests, with traces of both cocaine and opiates found in her urine. Lankes and her supervisor, Liz Reiter, then sat down with Reiter’s boss and a team of other supervisors and caseworkers.

“It is never an easy decision to remove kids from home, even when we know it is in their best interest,” Reiter told me. But, she said, “When we see that someone is using multiple substances, we need to assure the children’s safety. If we can’t get into the home, that makes us worry that things aren’t as they should be. It’s a red flag.” The team decided to request an Emergency Custody Authorization from a family-court judge. By late afternoon, with authorization in hand, they headed over to the family’s home, where a police officer met them.

The oldest child answered their knock. The mother wasn’t home, but all three children were, along with the mother’s elderly grandfather. Lankes called the mother, who answered for the first time in two months and began yelling about what she considered an unwarranted intrusion into her home. But she gave Lankes the names of family members who could take the children for the time being. Clothing was gathered, bags packed and winter jackets put on. Then it was time for the children to get in the car with Lankes, a virtual stranger empowered by the government to take them from their mother’s care.

At a hearing the next day, the presiding official ordered the mother to get clean before she could have her children returned. The drug-treatment center she had been attending advised her to enter rehab, but she refused. “We can’t get in touch

the Washington University researcher. As an author of a recent study showing that one in three United States children is the subject of a child-welfare investigation by age 18, he believes agencies must do everything possible to sharpen their focus.

Even in Illinois, where B.J. Walker, the director of the state's Department of Children and Family Services, is terminating its contract with the companies that developed Rapid Safety Feedback, predictive analytics is not dead. "I still believe it's a good tool to make better informed decisions," Walker told me in December. Walker knows Cherna and Dalton and saw the long process they went through to develop the Family Screening Tool. "They're doing a careful job," she said. "Their transparency has been laudable. And transparency isn't often your friend, because you're going to make some mistakes, you're going to stumble, you're going to make changes."

Cherna and Dalton are already overseeing a retooling of Allegheny County's algorithm. So far, they have raised the program's accuracy at predicting bad outcomes to more than 90 percent from around 78 percent. Moreover, the call screeners and their supervisors will now be given less discretion to override the tool's recommendations — to screen in the lowest-risk cases and screen out the highest-risk cases, based on their professional judgment. "It's hard to change the mind-set of the screeners," Dalton told me. "It's a very strong, dug-in culture. They want to focus on the immediate allegation, not the child's future risk a year or two down the line. They call it clinical decision-making. I call it someone's opinion. Getting them to trust that a score on a computer screen is telling them something real is a process."

Dan Hurley is a science journalist and longtime contributor to the magazine. He is at work on a book about his experiences as a foster father and scientific efforts to prevent and treat child abuse.

Sign up for our newsletter to get the best of The New York Times Magazine delivered to your inbox every week.

A version of this article appears in print on January 7, 2018, on Page MM31 of the Sunday Magazine with the headline: Predicting the Worst.
