BY SARAH WILLIAMS

CAN BIG DATA STOP OPIOID ABUSE?

Armed with electronic health records and insurance claim information, data scientists are trying to predict who's going to become addicted to opioids—and stop them before it's too late.

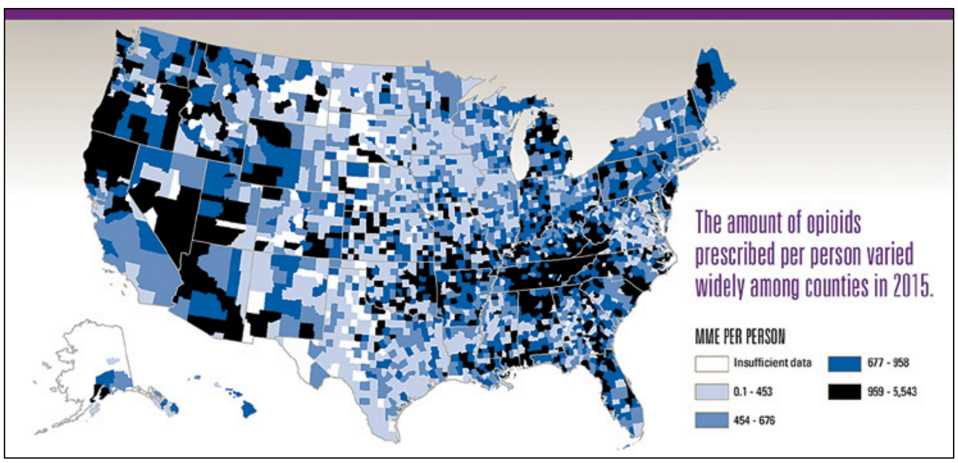
A person who is overdosing on opioids exhibits telltale signs: a limp body, slowed breathing and heart rate, and blueish or purplish fingernails and lips. But millions more who live with an opioid problem are harder to spot. They may be male or female, old or young, employed or homeless, lonely or part of a large supportive family. It's often an invisible problem until it's too late.

To identify those who are abusing opioids as well as those who are at risk, researchers are turning to datasets—including electronic health records (EHRs), insurance claims, statewide pharmacy databases, and emergency medical service calls.

"The more knowledge we have and the more we can use automated tools to fight this, the better," says **Caleb Alexander, MD**, co-director of the Johns Hopkins University Center for Drug Safety and Effectiveness.

Indeed, data science could become an important weapon in the battle against the skyrocketing rates of opioid abuse in the United States. Though some opioids are street drugs, such as heroin, many people become addicted after being prescribed pain relievers that are considered safe in small doses and for a short period of time. But these drugs also cause a sense of euphoria that can lead people to crave them. As the cravings increase, prescription users may end up taking opioids without a prescription or in a larger dose than prescribed. The epidemic has been linked not only to increasing deaths by overdose (according to the Centers for Disease Control, in the US roughly 100 people die of an opioid overdose each day), but also to staggering medical costs, a rise in the number of children entering foster care, and even a shrinking labor force.

It can be difficult for clinicians to identify addicts or other individuals at high risk of abusing opioids. Even armed with a thorough physical and mental health history, the most skilled doctor may miss the signs; and that's where hard data comes in. Researchers are finding that information on when and where patients fill prescriptions, what doctors they visit, and where they live, among other factors, may reveal an addiction or risk of overdose. Using data to flag these individuals may be one part of a much broader effort to slow the opioid epidemic.



Higher opioid prescribing puts patients at risk for addiction and overdose. The wide variation among counties' MME (morphine milligram equlvalents) per person (as shown on this map) suggests a lack of consistency among providers when prescribing opioids. SOURCE: Centers for Disease Control, Vital Signs, *July 2017.*

Calculating Risk from Claims and EHR Data

Thomas Ciesielski, MD, a doctor of internal medicine at Washington University School of Medicine, recently collaborated with Express Scripts, a pharmacy benefit manager, to determine whether they could predict opioid abuse or dependence using pharmacy and health insurance claims information. He and his colleagues analyzed de-identified data on nearly 700,000 Express Scripts users and discovered 12 patient characteristics that increased the odds of opioid abuse or dependence. The strongest predictors: chronic use of opioids, mental illness, alcohol and non-opioid substance abuse, younger relative age, and male gender. The work was published in the *American Journal of Medicine* in 2016.

"My hope in adding to this body of literature was to give clinicians a better understanding of the risk factors that a patient sitting in front of them has," says Ciesielski. "If they have this information before they write a prescription, they can have more informed discussions with their patients."

Joseph Boscarino, PhD, an epidemiologist and social psychologist at Geisinger Health in Danville, Pennsylvania, says that pinpointing a patient who may be addicted to opioids—based on EHR data—can be useful even after initial prescriptions are written. Boscarino and his colleagues recently discovered that a patient's healthcare costs often spike just before an overdose.

"We found the signal by accident," he says, pointing out that a spike in costs after an overdose—when a patient needs life-saving medical care and mental health support—would be more expected. When Boscarino and his team looked closer at their data—which included the electronic health records of more than 2,000 patients admitted to the hospital for an overdose between 2005 and 2015—they found other risk factors associated with overdoses: being unmarried, unemployed, and taking other prescriptions along with opioids.

"Now our health system is aware that we might be able to catch overdoses by looking at this cost signal, especially if we know a patient is on all these other drugs," says Boscarino.

Moving Beyond the EHR

But while researchers like Ciesielski and Boscarino can find trends in their datasets, they also admit their limitations; the EHR or pharmacy claims datasets weren't created with the goal of diagnosing opioid misuse.

Ciesielski, for instance, suspected that the distance between a patient's home and where they had a prescription filled might be predictive of opioid misuse. "Patients driving further to get opioids might be doing it because they're no longer able to get opioids in their own area," he says. But the location data included in the Express Scripts dataset he studied only included zip codes—not the most precise way of calculating distance—and he found no clear association.

"Lots of information is just not in the EHR," points out Boscarino. Because many EHR systems are not connected to one another, "I don't know from the EHR whether a patient is also driving to another hospital system to get drugs," he says.

And studies are also plagued by the challenge of identifying opioid abusers in the first place—information that's needed in order to train a machine-learning algorithm. Today, most studies rely on International Classification of Diseases (ICD) diagnostic codes, which can be found in medical charts and claims data. But doctors, even if they suspect a patient may be misusing opioids, don't always put these codes for opioid abuse or dependence in the patient's chart. Instead, they may simply stop treating the patient or order a lower dose of opioids without documenting why.

Rather than relying on the ICD codes, **Brandon Cosley, PhD**, a researcher on the predictive analytics team at BlueCross BlueShield (BCBS) of Tennessee, says that they set a threshold for what is considered opioid abuse by looking directly at raw claims data in their possession.

"Our definition usually contains how much opioid has been prescribed, how many doctors someone has gotten prescriptions from, and how many pharmacies they've filled the prescriptions at," says Cosley. He parses patient claims data and demographics to find what other factors predict the elements of the definition. For several different subpopulations of BCBS Tennessee's members, Cosley says he has developed models that predict opioid abuse with 80 percent accuracy.

"We're talking about hundreds of different risk factors for any given individual, and they combine in many ways," he says. "The relative contribution of any one risk factor may be small."

Cosley has access to a more complete dataset than researchers studying single EHR systems, yet he says more data would be useful.

"I think one of our most exciting initiatives is to really incorporate some of the free-form text data that we get from our membership," Cosley says. "Things like nurses' and doctors' notes and customer service calls." He says BCBS Tennessee has technology that allows them to start analyzing this kind of data and is working on ways to use it effectively.

Other researchers are also eyeing whether genetics data might help predict opioid abuse. A recent research effort led by researchers at Proove Biosciences, a precision medicine company based in Irvine, California, that is dedicated to optimizing the treatment of chronic pain, studied an algorithm—dubbed the Proove Opioid Risk (POR)—which used small variations in the genome, called single nucleotide polymorphisms or SNPs, combined with clinical risk factors to predict opioid abuse. Patients in the highest category of POR scoring had 16 times greater odds of opioid use disorder. The work was published in 2017 in *Pharmacogenomics and Personalized Medicine*.

"I think in the future we're going to have to do a better job linking EHRs to other assets including genetics," Boscarino says. "Addiction is really complicated."

Data-driven Interventions

Some communities are also crunching data to identify doctors or pharmacies that are part of the problem, or neighborhoods that should be the focus of community-based interventions. Massachusetts, for instance, has used predictive analytics to allocate resources to neighborhoods with the biggest overdose rates. And Alleghany County, Pennsylvania, tracked overdose deaths in the county over a six-year timespan to pinpoint who was overdosing and when.

If communities, hospital systems and insurance companies can start recognizing the patterns of opioid abuse in their computers, can they stop the opioid epidemic? Probably not, but they might help slow it, says Alexander, who in July 2017 published a review in the *Journal of the American Medical Informatics Association* evaluating 15 algorithms that have been used to identify non-medical opioid use in EHR data.

"I don't think anyone is naive enough to believe that automated tools alone will suffice," he says. "But these tools can, in some cases, be used to simply raise awareness and promote information sharing across clinical teams when a patient is at elevated risk for injury or death."

The Geisinger pharmacy team, for instance, has started flagging opioid abuse risk factors in a patient's EHR, based on Boscarino's findings. This doesn't mean a patient with risk factors won't be able to get appropriate painkillers, but it could mean that a clinician or pharmacist reconsiders dosing or limits how many pills a patient gets at a time.

Alexander notes that insurance companies have a number of opportunities to improve patient care for those in pain while simultaneously reducing the overuse of prescription opioids. "Payers have a lot of tools in their toolbox, ranging from improving the coverage of non-drug treatments to designing new programs to identify and manage patients who are at highest risk of injury or overdose death," he says.

Data-science approaches to the opioid epidemic also have limitations: "They have to be used carefully and for the right purposes," Alexander says. "But they nevertheless can be quite powerful because they shine a light on potentially concerning patterns and allow for the identification of subpopulations who are at risk." □