Within moments of the explosions at the finish line of the 2013 Boston marathon, Twitter messages began appearing describing the incident and casualties—some preceded the alerts from Massachusetts public health and emergency agencies sent to emergency departments (EDs) in the area (Cassa CA et al. PLoS Curr. Published online July 2, 2013).

The prospect of using such real-time data to provide early public health warnings—possibly within minutes—and timely alerts to hospitals about oncoming surges or emerging public health concerns has led to a new wave of studies probing ways to use Twitter in medicine and public health.


Now, scientists are finding that Twitter data—especially when combined with other real-time data streams like environmental sensors or data from fitness apps—also have the potential to provide early warnings about chronic disease, emergencies, adverse drug reactions, or even safety problems like prescription drug misuse.

Finding the Signal
Enormous amounts of real-time data are generated each day as people use mobile devices and apps like Twitter. In fact, the average US resident now spends about 3 hours each day using a mobile device, according to analysts at Yahoo's Flurry Insights (http://bit.ly/1NzuROH).

“We live in the era of big data,” said Sudha Ram, PhD, professor of management information systems and computer science at the University of Arizona, Tucson, whose research focuses on predicting asthma surges. “We have tons of data pouring out from everywhere. Why not look at [asthma surges] in a very different way and bring other data to bear?”

Ram’s research focuses on extracting health signals from this fire hose of data using data mining techniques. In Twitter, Ram has found a particularly rich vein of information.

“A lot of people dismiss Twitter as frivolous, but my work has shown there...
are a lot of nice signals you can get out of Twitter if you know how to extract them," Ram said.

She and her colleagues discovered many tweets about trouble breathing, forgetting asthma medications, receiving calls from teachers about a child's asthma attack, or cutting exercise short because of asthma symptoms (Ram S et al. IEEE J Biomed Health Inform. 2015;19[4]:1216-1223).

Many people share symptoms or information about health-related behaviors on Twitter long before they ever see a doctor, explained John Brownstein, PhD, chief innovation officer at Boston Children's Hospital and associate professor at Harvard.

"It could be an early indicator," Brownstein said. "Twitter represents a place where people are disclosing symptoms in the millions."

These firsthand reports are much more timely than the weekly reports from the US Centers for Disease Control and Prevention (CDC) that many hospitals rely on to predict asthma-related visits. Often, the CDC data, which are based on reports from state public health departments, are weeks old by the time hospitals see them, Ram explained.

But are Twitter posts a reliable indicator of asthma exacerbations? To figure that out, Ram and her colleagues used semi-automated techniques to sift through more than 1 million asthma-related tweets posted in the Dallas–Fort Worth area in the fall of 2013 and weeded out Twitter posts made in jest or that were clearly not relevant. For example, someone posting that their chemistry examination is giving them trouble would not be given asthma status.

"After excluding this noise, they found that the number of asthma-related tweets was associated with ED visits for asthma in the area during that fall. Asthma tweets were not correlated with ED visits for abdominal pain or constipation, which were used as a control."

Next, they looked at 2 other real-time data streams: Google search term data and environmental data streams: Google search term data and air pollution sensor, and Twitter data for more than 80 hospitals and their surrounding areas. Preliminary analyses appear to support the value of using Twitter and air sensor data to predict surges in asthma-related ED visits.

"Once these final data are in, Ram and her colleagues hope to build a tool that will help hospitals predict whether they will have a high, low, or medium number of asthma visits on a given day. This would provide a more timely population-level warning to compliment hospitals' modeling based on individual patients."

For example, Parkland Center for Clinical Innovation in Dallas, with which Ram collaborated on the study, has also developed a model to use electronic medical records to predict whether individual patients are likely to have an exacerbation, but within a longer prediction window of 3 months rather than weeks.

"The physicians are really excited [about having more timely alerts]," she said.

**Fly on the Wall**

In addition to symptoms, Twitter users also share information about their behavior or concerns about drugs and other medical products that might help alert public health authorities to emerging concerns.

"In the midst of the ongoing epidemic of prescription opioid abuse, Urmi Sarkar, MD, associate professor of medicine at the University of California, San Francisco, and her colleagues used Twitter to gauge public sentiment about such medication misuse.

"People tweet about their Pap smears—nothing is off limits," Sarkar explained. "So, we thought it would be an interesting data source."

The public's no-holds-barred approach to tweets extended to prescription drug misuse. Ram and her colleagues quickly discovered. She found that 70% of the messages about personal experience using opioids referred to taking the medication to "get high," or other forms of misuse (Chan B, Lopez A, Sarkar U. PLoS One. 2015:10[8]: e0135072). Many posts described using opioid medications to sleep or in dangerous combinations with other drugs.

"I was pretty shocked by the problematic behaviors people were disclosing," Sarkar said. "They were using [opioids] in a way that is really not safe."

The behaviors were so problematic, in fact, that Sarkar said she would make a referral for addiction treatment or discontinue the medication if she heard similar comments from a patient. The results of the small study, which looked at a hand-sorted set of 540 messages that referenced opioid medication names or related slang, suggest that analyzing Twitter data may be a way to quickly assess attitudes toward opioid abuse.

"Twitter would likely supplement rather than supplant more traditional techniques for measuring health behaviors," Sarkar explained. She noted that currently, surveys are the mainstay of such research. But reduced timeliness is often a trade-off when trying to boost survey quality. Results are often reported a year or even longer after the survey was administered. These delays stymie public health officials' efforts to use the data to respond to an ongoing epidemic.

Another advantage of Twitter data is that the data allow researchers to be the proverbial fly on the wall. With illegal or potentially stigmatizing behaviors, people may not be as frank in answering survey questions as they are when talking with friends on social media.

"Their survey responses are affected by the fact they know they are being questioned about their behavior," Sarkar explained.

"Twitter also has emerged as a promising source for reports of adverse experiences. Brownstein explained that it is diffi-
cult to get people to report drug adverse events through traditional channels, such as the US Food and Drug Administration’s (FDA’s) Adverse Events Reporting System (FAERS) or through the drug’s manufacturer. But people do discuss drug adverse events on Twitter.

With funding and support from the FDA, Brownstein and his colleagues used a combination of automation and hand sorting to cull 6.9 million tweets mentioning 23 medical products to identify 4400 describing probable adverse events (Freifeld CC et al. Drug Saf. 2014;37:343-350). The tweets provided much of the same information collected by FAERS, but a larger sample would be needed to detect a safety signal, according to Brownstein and his colleagues.

“A tweet will not lead to a black box warning, but it can lead to a nudge to look at other data,” Brownstein said. “There is still value in looking across many data sources.”

Moving Toward Personalization

While the research so far suggests great potential for public health warning systems built off of Twitter data, the field is still very young. The high-volume data analytics techniques it relies on are still developing and, like any field, it has its limitations.

Brownstein, who is currently working with officials in the European Union to develop a warning system about emerging crises that uses Twitter data, explained that especially in emergency situations it is important for machine learning tools to be as robust as possible to avoid false alarms.

“Lots of work is [still] needed on the methods and how to interpret the data,” he said.

Twitter data may also not work for every population.

“It is very clear it works better in places with good Internet connectivity and a lot of people,” Brownstein said. “Otherwise, the noise can outweigh the signals.”

Sarkar also acknowledged that social media users aren’t necessarily a representative sample. According to the Pew Research Center reports never going online (http://pewrsr.ch/1LPDFNJ). These nonusers are disproportionately older, less well educated, less affluent, and more likely to live in rural areas.

“It’s important to remember who is being left behind,” Sarkar said.

Still, three-quarters of adults who go online report using social media sites, according to the Pew Research Center (http://pewrsr.ch/1doWsil), with one-quarter reporting using Twitter. Use of social media is highest among younger adults, with 90% of 18- to 29-year-olds using these online networking forums compared with 78% of those aged 30 to 49 years, 65% of those aged 50-64, and 46% of those aged 65 years or older. But social media use has grown across all age groups, according to the Pew data.

“There are millions of Twitter users,” Sarkar said. “It’s okay if this doesn’t extrapolate to non-Twitter users. If we can reach [Twitter users] and improve their health behavior that would be a substantial step.”

The possibility of translating Twitter-based research into more personalized outreach is something Brownstein and Ram endorse. Ram said the asthma data she collected might also be used to develop apps that warn individual patients when they are at elevated risk of an asthma exacerbation. Brownstein, who also has researched chronic conditions on Twitter, said it could be a powerful tool to recruit individuals for clinical trials.

“Online networks are people’s real worlds now,” said Sarkar said. “We have to learn how to engage with people where they are and where they are is online.”