BUSINESS

Professors: Google site shows flu waning

Data from Web searches used to quickly track flu's spread, and now its decline

By Julie Wernau
TRIBUNE REPORTER

The flu in the United States is on track to bottom out by the end of the year, returning to levels seen nine months ago, before the first case of H1N1 hit the U.S. That forecast is from three University of Chicago professors who have created a new method for predicting the spread of the disease.

The professors, Nicholas Polson and Hedibert Lopes at the university's Booth School of Business and Vanja Dukic in the Department of Health Studies, are using data from Google Flu Trends, currently the fastest way to track the flu, and applying a mathematical model that aims to do the one thing Google can't yet do: predict what will happen next.

The same modeling approaches, with minor alterations, could be applied to other infectious disease epidemics such as avian flu, SARS or measles. Google created its Flu Trends tool after it discovered that the searches people conduct when they are sick with the flu can be mathematically tied to the actual spread of the flu.

And although the professors are using the resource for surveillance data, the same model, they say, could be applied to any data that tracks infection over time, including CDC and European surveillance data.

A few weeks ago the Google graph had them worried.

Its steep curve, laid over the flatter data from previous years, had many of the symptoms (mathematically speaking) that predicted an epidemic along the lines of the 1918 Spanish flu. That outbreak, caused by a deadly H1N1 virus strain, quickly spread around the world and killed about 50 million people, including an estimated 500,000 in the U.S.

But now, the fact the Google searches are subsiding indicates the disease might be ebbing. Dukic said the flu seems to have hit its peak nationwide in mid- to late October and has been trending downward since. There might be a slight increase in January, she said, due to more travel over the holidays. Illinois, she said, is expected to trend with the rest of the nation.

"The flu has now reached a lot of people. People are immune," Dukic said. "The number of people who can get the flu has decreased."

She also explained how mathematics plays a role in their predictions.

First, the data — in this case, from Google — is plugged into a model called SEIR (susceptible–exposed–
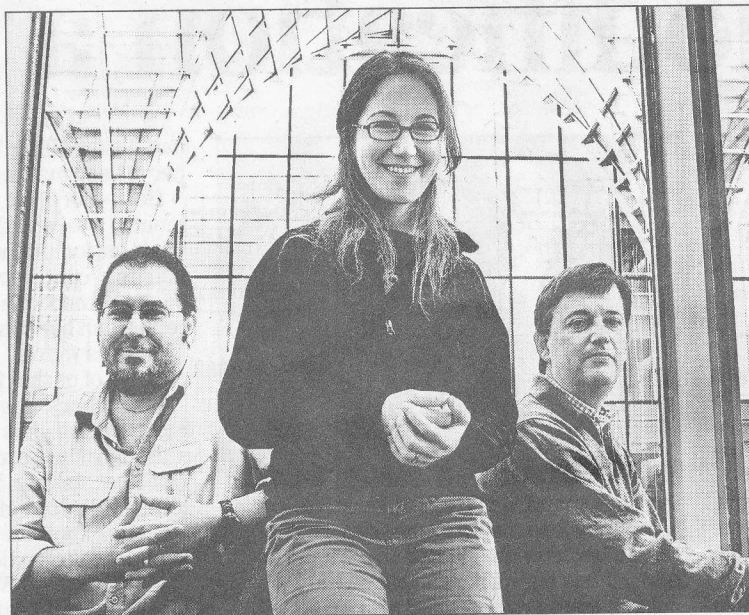
University of Chicago professors Hedibert Lopes, from left, Vanja Dukic and Nicholas Polson are using Google Flu Trends data to predict the spread of flu cases. **DAVID PIERINI/TRIBUNE PHOTO**

infected–recovered), she said, which accounts for the number of possible contacts between a susceptible person and a healthy person given the current flu data.

"Our simple model tracks the true number of people with the flu through time, and this number reflects the impact of some people getting vaccinated, subject to vaccine availability," Dukic said.

Fewer susceptible people, from immunity or vaccination, is good news for preventing epidemics, she said.

Polson said a Bayesian statistical model, a kind of math that allows for continuous updating based on new data, then takes the past historical data and plots out the many probable scenarios to figure out the most likely trend the flu could take in the future.

"It's just a very powerful set of techniques and tools to predict what's going to happen next," Polson said.

The professors chose to use Goo-

gle Flu Trends data, which is available at **google.org/flutrends**, because it closely mirrors Centers for Disease Control and Prevention data for actual flu cases and is unique in that it makes it possible to track the flu with a lag time of just one day, versus the week it takes to gather and release CDC data.

The model can also estimate flu activity on a state-by-state basis, data the CDC does not publicly release, and provides experimental estimates for flu activity in regions around the world that have no CDC-like agency tracking influenza.

"They have a stream of the data of the searches you've done throughout the day. And imagine that multiplied by millions and millions of users," Lopes said.

According to Google, the Flu Trends searches are tied to IP addresses, which can be tracked to a region of the country but not to a particular individual.

They also are not specific to a particular strain of the flu, such as H1N1.

"Given that the data is publicly available on the Web site, we don't know exactly how many scholars and scientists are using it to track the flu," said Google spokesman Jamie Yood. "We have heard that some public health officials are using our daily updates to give them a preview of flu activity in various states."

jwernau@tribune.com