

# Chance for advance flu warning in search-based tracking method

November 10 2015, by Alvin Powell

---



CDC data are considered something of a gold standard in disease surveillance, but it takes one to three weeks for reports to be compiled, making it hard for public health officials to stay ahead of the flu, said Samuel Kou (center), a professor of statistics. Kou worked with Shihao Yang (left) and Mauricio Santillana to deliver results in real time. Credit: Kris Snibbe/Harvard Staff Photographer

Harvard statisticians have devised a new method to track the flu via Internet search data, potentially providing public health officials and consumers alike with advance warning this flu season.

The system adopts an approach—tracking searches for key words and phrases such as "flu," "[flu symptoms](#)," "treating the flu"—that has been tried before with varying levels of success, but combines it with additional data to improve accuracy.

The result, according to Samuel Kou, a professor of statistics, is the most precise method yet.

"If you see a spike in search volume, it probably indicates that something is going on," Kou said.

The history of tracking the flu using search data goes back several years. The most notable effort, Kou said, was that of Google Flu Trends, released in 2009 and discontinued last summer. Though Google Flu Trends used sophisticated algorithms and was revised repeatedly, its predictions often missed the retrospective analysis of the Centers for Disease Control and Prevention, which tracks actual reports of people with flu-like symptoms seeking medical attention across the country.

CDC data are considered something of a gold standard in disease surveillance, Kou said, but it takes one to three weeks for reports to be compiled, making it hard for [public health officials](#) to stay ahead of the flu.

Though most sufferers recover relatively quickly, the disease can be deadly, killing 500,000 annually around the world. In the United States, it kills an estimated 3,000 to 5,000 people each year.

Kou worked with Ph.D. student Shihao Yang and Mauricio Santillana, a

lecturer in applied mathematics, to deliver results in real time.

The approach, called ARGO, for AutoRegression with Google search data, combines Google data with historical records from the CDC and information on seasonality of the [flu](#). It also accounts for changes in the inner workings of Google's [search engine](#) and shifts in search behavior. People learn as they search for information, Kou said, changing their queries and becoming better searchers.

"If I want to search for something, I do it better now than I did two years ago. Besides, Google's search engine evolves, and so does the interaction between people and the engine," Kou said.

The project began about a year ago after a casual conversation on the topic between Kou and Santillana, who also has an appointment at Harvard-affiliated Boston Children's Hospital. The team is currently working to make ARGO open source and widely available.

**More information:** S. Yang et al. Accurate estimation of influenza epidemics using Google search data via ARGO, *Proceedings of the National Academy of Sciences* (2015). [DOI: 10.1073/pnas.1515373112](https://doi.org/10.1073/pnas.1515373112)

*This story is published courtesy of the [Harvard Gazette](#), Harvard University's official newspaper. For additional university news, visit [Harvard.edu](#).*

Provided by Harvard University

Citation: Chance for advance flu warning in search-based tracking method (2015, November 10) retrieved 20 April 2024 from <https://medicalxpress.com/news/2015-11-chance-advance-flu-search-based-tracking.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.