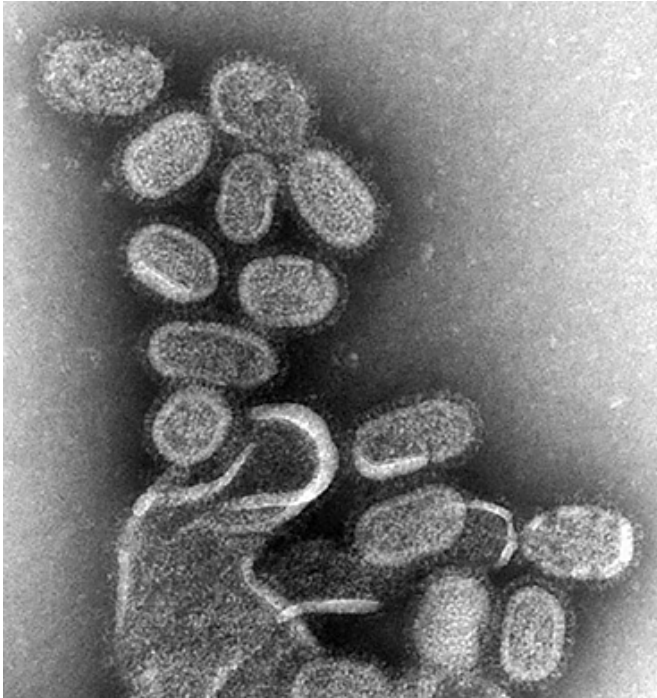


# Real-time influenza tracking with 'big data'

11 May 2016



Electron microscopy of influenza virus. Credit: CDC

Early detection and prediction of influenza outbreaks is critical to minimizing their impact. Currently, flu-like illnesses are tracked by the Centers for Disease Control and Prevention, but with a time lag of one to two weeks. Now, a team led by researchers at Boston Children's Hospital shows that cloud-based data from electronic health records (EHRs) can be used to pick up cases in real time, at least one week ahead of CDC reporting.

By combining EHR data, historical patterns of [flu activity](#) and a machine-learning algorithm to interpret the data, the researchers made accurate predictions of national and local influenza activity that matched subsequent reporting by the CDC. They reported their findings online May 11 in *Scientific Reports*, an online, open access journal from the publishers of Nature.

"Having access to near-real-time aggregated EHR information has enabled us to significantly improve our [flu](#) tracking and forecasting systems," says lead author Mauricio Santillana, PhD, faculty member at Boston Children's Computational Health Informatics Program (CHIP), who also holds a faculty appointment at Harvard Medical School and is an associate at the Harvard Institute for Applied Computational Sciences. "Real-time tracking will enable local public health officials to better prepare for unusual flu activity and potentially save lives."

The study tapped data from Athenahealth, a provider of cloud-based medical applications. The company's database encompassed more than 72,000 healthcare providers and EHRs for more than 23 million patients, mostly seen in office-based settings.

The investigators first trained the flu-prediction algorithm, called ARES, with data on weekly total visit counts, visit counts for flu and flu-like illness, visit counts for [flu vaccination](#) and other data captured from June 2009 through January 2012. They then used ARES to estimate flu activity over the next three years (through June 2015).

The team showed that ARES' estimates of national and regional flu activity had error rates 2- to 3-fold lower than earlier predictive models. ARES also correctly estimated the timing and magnitude of the national flu "peak week." It was slightly less accurate in predicting regional peak weeks, but clearly outperformed Google Flu Trends, another real-time system that tracked outbreaks by mining Internet searches. (Google Flu Trends was shut down in August, 2015.)

"Our study shows the true value of considering multiple data streams in disease surveillance," says John Brownstein, PhD, the study's senior investigator and Chief Innovation Officer at Boston Children's Hospital. "While Google data provide incredible real-time population wide information, clinical data add a more accurate and precise assessment of disease state. As EHR data become

more ubiquitously available, we will see major leaps in our ability to monitor and track disease outbreaks."

**More information:** Santillana, M. et al. Cloud-based Electronic Health Records for Real-time, Region-specific Influenza Surveillance. *Sci. Rep.* 6, 25732; [DOI: 10.1038/srep25732](https://doi.org/10.1038/srep25732) (2016).

Provided by Children's Hospital Boston

APA citation: Real-time influenza tracking with 'big data' (2016, May 11) retrieved 24 September 2020 from <https://techxplore.com/news/2016-05-real-time-influenza-tracking-big.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.*