

# Computer model more accurate at identifying sources of foodborne illnesses than traditional

6 November 2018



Credit: CC0 Public Domain

A new computer model that uses machine learning and de-identified and aggregated search and location data from logged-in Google users was significantly more accurate in identifying potentially unsafe restaurants when compared with existing methods of consumer complaints and routine inspections, according to new research led by Google and Harvard T.H. Chan School of Public Health. The findings indicate that the model can help identify lapses in food safety in near real time.

"Foodborne illnesses are common, costly, and land thousands of Americans in emergency rooms every year. This new technique, developed by Google, can help restaurants and local [health departments](#) find problems more quickly, before they become bigger public health problems," said corresponding author Ashish Jha, K.T. Li Professor of Global Health at Harvard Chan School and director of the Harvard Global Health Institute.

The study will be published online November 6, 2018 in *npj Digital Medicine*.

Foodborne illnesses are a persistent problem in the U.S. and current methods by restaurants and local health departments for determining an outbreak rely primarily on consumer complaints or routine inspections. These methods can be slow and cumbersome, often resulting in delayed responses and further spread of disease.

To counter these shortcomings, Google researchers developed a machine-learned [model](#) and worked with Harvard to test it in Chicago and Las Vegas. The model works by first classifying search queries that can indicate [foodborne illness](#), such as "stomach cramps" or "diarrhea." The model then uses de-identified and aggregated location history data from the smartphones of people who have opted to save it, to determine which restaurants people searching those terms had recently visited.

Health departments in each city were then given a list of restaurants that were identified by the model as being potential sources of foodborne illness. The city would then dispatch health inspectors to these restaurants, though the health inspectors did not know whether their inspection was prompted by this [new model](#) or traditional methods. During the period of the study, health departments continued to follow their usual inspection procedures as well.

In Chicago, where the model was deployed between November 2016 and March 2017, the model prompted 71 inspections. The study found that the rate of unsafe restaurants among those detected by the model was 52.1% compared with 39.4% among inspections triggered by a complaint-based system. The researchers noted that Chicago has one of the most advanced monitoring programs

in the nation and already employs social media mining techniques, yet this new model proved more precise in identifying restaurants that had food safety violations.

In Las Vegas, the model was deployed between May and August 2016. Compared with routine inspections performed by the health department, it had a higher precision rate of identifying unsafe restaurants.

When the researchers compared the model with routine inspections by health departments in Las Vegas and Chicago, they found that the overall rate across both cities of unsafe restaurants detected by the model was 52.3%, whereas the overall rate of detection of unsafe restaurants via routine inspections across the two cities was 22.7%.

Interestingly, the study showed that in 38% of all cases identified by this model, the [restaurant](#) potentially causing foodborne illness was not the most recent one visited by the person who was searching keywords related to symptoms. The authors said this is important because previous research has shown that people tend to blame the last restaurant they visited and therefore may be likely to file a complaint for the wrong restaurant. Yet clinically, foodborne illnesses can take 48 hours or even longer to become symptomatic after someone has been exposed, the authors said.

The new model outperformed complaint-based inspections and routine inspections in terms of precision, scale, and latency (the time that passed between people becoming sick and the outbreak being identified). The researchers noted that the model would be best leveraged as a supplement to existing methods used by health departments and restaurants, allowing them to better prioritize inspections and perform internal food safety evaluations. More proactive and timely responses to incidents could mean better public [health](#) outcomes. Additionally, the model could prove valuable for small and mid-size restaurants that can't afford safety operations personnel to apply advanced [food safety](#) monitoring and data analysis techniques.

"In this study, we have just scratched the surface of

what is possible in the realm of machine-learned epidemiology. I like the analogy to the work of Dr. John Snow, the father of modern epidemiology, who in 1854 had to go door to door in Central London, asking people where they took their water from to find the source of a cholera outbreak. Today, we can use online data to make epidemiological observations in near real-time, with the potential for significantly improving [public health](#) in a timely and cost-efficient manner," said Evgeniy Gabrilovich, senior staff research scientist at Google and a co-author of the study.

**More information:** Adam Sadilek et al, Machine-learned epidemiology: real-time detection of foodborne illness at scale, *npj Digital Medicine* (2018). [DOI: 10.1038/s41746-018-0045-1](https://doi.org/10.1038/s41746-018-0045-1)

Provided by Harvard T.H. Chan School of Public Health

APA citation: Computer model more accurate at identifying sources of foodborne illnesses than traditional (2018, November 6) retrieved 20 May 2019 from <https://medicalxpress.com/news/2018-11-accurate-sources-foodborne-illnesses-traditional.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.*