

# IBM researchers use grocery scanner data to speed investigations during early foodborne illness outbreaks

August 15 2016

---

Foodborne illnesses, like salmonella, E. Coli and norovirus infections, are a major public health concern affecting more than one out of six Americans each year, according to the Centers for Disease Control and Prevention (CDC). During a foodborne illness outbreak, rapidly identifying the contaminated food source is vital to minimizing illness, loss and impact on society.

IBM announced its scientists have discovered that analyzing retail-scanner data from grocery stores against maps of confirmed cases of foodborne illness can speed early investigations. In the study, researchers demonstrated that with as few as 10 medical-examination reports of foodborne illness they can narrow down the investigation to 12 suspected [food products](#) in just a few hours.

In the study, researchers created a data-analytics methodology to review spatio-temporal data, including [geographic location](#) and possible time of consumption, for hundreds of grocery product categories. Researchers also analyzed each product for its shelf life, geographic location of consumption and likelihood of harboring a particular pathogen – then mapped the information to the known location of illness outbreaks. The system then ranked all grocery products by likelihood of contamination in a list from which public [health officials](#) could test the top 12 suspected foods for contamination and alert the public accordingly.

A traditional investigation can take from weeks to months and the timing can significantly influence the economic and health impact of a disease outbreak. The typical process employs interviews and questionnaires to trace the contamination source. In 2011, an outbreak of E. coli in Europe took more than 60 days to identify the source, imported fenugreek seeds. By the time the investigation was completed, all the sprouts produced from the seeds had been consumed. Nearly 4,000 people became ill in 16 countries and more than 50 people died before public health officials could pinpoint the source, according to the European Food Safety Authority.

"When there's an outbreak of [foodborne illness](#), the biggest challenge facing public health officials is the speed at which they can identify the contaminated [food source](#) and alert the public," said Kun Hu, public health research scientist, IBM Research – Almaden in San Jose, Calif. "While traditional methods like interviews and surveys are still necessary, analyzing big data from retail grocery scanners can significantly narrow down the list of contaminants in hours for further lab testing. Our study shows that Big Data and analytics can profoundly reduce investigation time and human error and have a huge impact on [public health](#)."

Already, the method in this study has been applied to an actual E. coli illness outbreak in Norway. With just 17 confirmed cases of infection, [public health officials](#) were able to use this methodology to analyze grocery-scanner data related to more than 2,600 possible food products and create a short-list of 10 possible contaminants. Further lab analysis pinpointed the source of contamination down to the batch and lot numbers of the specific product – sausage.

The study, "[From Farm to Fork: How Spatial-Temporal Data can Accelerate Foodborne Illness Investigation in a Global Food Supply Chain](#)," was published in the Association for Computing Machinery's

*Sigspatial Journal.*

Provided by IBM

Citation: IBM researchers use grocery scanner data to speed investigations during early foodborne illness outbreaks (2016, August 15) retrieved 25 April 2024 from <https://phys.org/news/2016-08-ibm-grocery-scanner-early-foodborne.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.