ISDS 2012 Conference Abstracts

Influenza Forecasting with Google Flu Trends

Andrea F. Dugas*[1], Mehdi Jalalpour[1], Yulia Gel[1, 2], Scott Levin[1], Fred Torcaso[1], Takeru Igusa[1] and Richard Rothman[1]

[1]Johns Hopkins University, Baltimore, MD, USA; [2]University of Waterloo, Waterloo, ON, Canada

Objective

We sought to develop a practical influenza forecast model, based on real-time, geographically focused, and easy to access data, to provide individual medical centers with advanced warning of the number of influenza cases, thus allowing sufficient time to implement an intervention. Secondly, we evaluated how the addition of a real-time influenza surveillance system, Google Flu Trends, would impact the forecasting capabilities of this model.

Introduction

Each year, influenza results in increased Emergency Department crowding which can be mitigated through early detection linked to an appropriate response. Although current surveillance systems, such as Google Flu Trends, yield near real-time influenza surveillance, few demonstrate ability to forecast impending influenza cases.

Methods

Forecasting models designed to predict one week in advance were developed from weekly counts of confirmed influenza cases over seven seasons (2004 – 2011) divided into training and out-of-sample verification sets. Forecasting procedures using classical Box-Jenkins, generalized linear, and autoregressive methods were employed to develop the final model and assess the relative contribution of external variables such as, Google Flu Trends, meteorological data, and temporal information. Models were developed and evaluated through statistical measures of global deviance and log-likelihood ratio tests. An additional measure of forecast confidence, defined as the percentage of forecast values, during an influenza peak, that are within 7 influenza cases of the actual data, was examined to demonstrate practical utility of the model.

Results

A generalized autoregressive Poisson (GARMA) forecast model integrating previous influenza cases with Google Flu Trends information provided the most accurate influenza case predictions. Google Flu Trend data was the only source of external information providing significant forecast improvements ($p = 0.00002$). The final model, a GARMA intercept model with the addition of Google Flu Trends, predicted weekly influenza cases during 4 out-of-sample outbreaks within 7 cases for 80% of estimates (Figure 1).

Conclusions

Integer-valued autoregression of influenza cases provides a strong base forecast model, which is enhanced by the addition of Google Flu Trends confirming the predictive capabilities of search query based syndromic surveillance. This accessible and flexible forecast model can be used by individual medical centers to provide advanced warning of future influenza cases.

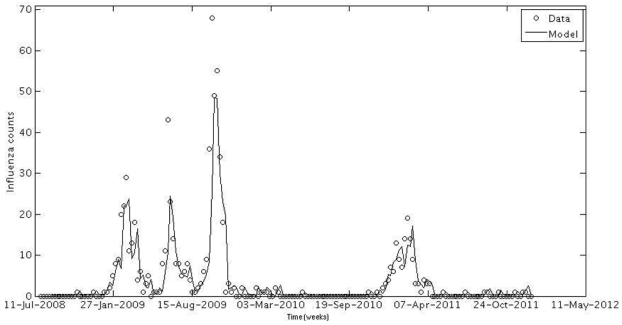

Figure 1: Number of weekly confirmed influenza cases during the verification period (2008-2011) comparing actual data (circles) and values forecasted by the final model [3rd order Generalized Autoregressive Poisson intercept model with Google Flu Trends] (solid line).

Keywords

Google Flu Trends; Crowding; Surveillance; Influenza; Forecasting

Acknowledgments

This work was supported by the Department of Homeland Security (PACER: National Center for Study of Preparedness and Response [grant number: 2010-ST-061-PA0001]); the National Science Foundation Systems Engineering and Design Program [grant number: NSF CMMI 0927207]; and the Natural Sciences and Engineering Research Council of Canada.

*Andrea F. Dugas
E-mail: adugas1@jhmi.edu

ISDS Annual Conference Proceedings 2012. This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (http://creativecommons.org/licenses/by-nc/3.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.