Using Search Volume for Surveillance of Medication Prescribing

Jacob E. Simmering*, Linnea A. Polgreen and Philip M. Polgreen

University of Iowa, Iowa City, IA, USA

Objective

To validate search volume estimation for outpatient medication prescribing.

Introduction

Investigators have used the volume of internet search queries to model disease incidence, especially influenza and general consumer behavior [1]. Our group has used search volume to model interest in FDA safety alerts and adverse drug event incidence. We found evidence of changes in search behavior following warnings and the expected relationship between search volume and adverse drug event incidence. Thus, search volume may help provide near real time surveillance of drug use patterns to help monitor and mitigate risk to the population from adverse drug events. However, the use of search query volume as a proxy for drug use has yet to be validated.

We attempt to validate search volume estimation of drug utilization in three ways: 1) explore seasonal variations in search volume and outpatient utilization, 2) monitor change between substitute drugs following patent expirations and 3) use search volume estimation methods to estimate TB incidence.

Methods

Google Insights normalized search share was used to characterize interest in a drug. The estimates of drug utilization were derived from the Medical Expenditure Panel Survey (MEPS), a nationally representative sample of the US population. Substitute drugs and notable patent expirations between 2004 and 2011 were obtained via pharmacist review. TB incidence was derived from the MMWR yearly Summary of Notifiable Diseases.

To validate the assumption that search volume relates to drug utilization, we estimated weekly utilization for 9 drugs (amoxicillin, azelastine, azithromycin, benzomatate, cefdinir, ciprofloxacin, levofloxacin, moxifloxacin and olopatadine) using MEPS for 2004-2009. The weekly utilization volume was cross-correlated with the Google Insights series with lags ranging from -6 to +6 months.

To compare the rate of substitution between name brand and generic drugs following the expiration of a patent, we treated the generic drug search volume as the independent variable and the name brand as the dependent variable. Using OLS, we calculated the marginal rate of substitution between the name brand and generic search queries. Preliminary work has focused on substitution of generic simvastatin for branded Zocor.

As TB treatment regimens usually include a fixed set of medications (isoniazid, rifampin, pyrazinamide, ethambutol), the utilization of these drugs should correspond with TB incidence. We modeled national TB incidence using OLS with search volume and an indicator for the month of December. The number of reported cases in December is inconsistent with the seasonality of TB in the US and is a significant departure from the expected value given the rest of the series. We suspect this is due to a reporting artifact and include the indicator variable in our model to mitigate the effects of this inconsistency.

Results

The seasonality of drug use is reflected in search volume. Only 3 of the 9 drugs (33%, amoxicillin, azithromycin and cefdinir) had enough volume in the MEPS to create a reasonable time series. All 3 drugs had statistically significant positive correlations at lags near 0 and significant negative correlations at lags of +/- 6 months. Amoxicillin, for example, had a significant correlation at lags around 0 of 0.55-0.60 and correlations at a lag of -5 or +5 months of -0.4. The magnitude of this correlation coefficient would suggest that the two series track closely.

Patent expirations (and the resulting emergence of generic medications with new names) are apparent in search volume as well. We find a strong negative relationship between search volume for simvastatin and Zocor. Specifically, a one unit increase in search volume for 'simvastatin' is associated with a 0.96 ($p < 0.0001$) unit decrease in the search volume for 'Zocor.'

The simple model for TB incidence demonstrates the utility of using drugs as queries for disease. Search volume was a significant ($p = 0.006$) and positive predictor of TB incidence controlling for the December aberrations.

Conclusions

The Google Insights search volume for a set of highly seasonal drugs is highly correlated with community utilization as measured by seasonal variance in utilization, change in search and prescribing patterns and expected prescribing following TB. The ability to estimate use of drugs from search volume presents a new method for keyword selection in search based incidence models and a method to monitor changes in the pharmaceutical market.

Keywords

pharmacovigilance; web search; prescribing behavior

References

1. Goel, S. et al. (2010) Predicting Consumer Behavior with Web Search. PNAS 104:17486-17490

*Jacob E. Simmering
E-mail: jacob-simmering@uiowa.edu