GLOBAL RAPID IDENTIFICATION TOOL SYSTEM

Text Samples
Infectious disease threats?

GRITS

Data feeds
Infectious disease threats?

METHODS

ABSTRACT

Global Rapid Identification Tool Set (GRITS) is a biosurveillance application that enables infectious disease analysts to monitor non-traditional information sources (e.g., online news outlets, ProMED reports, blogs) for infectious disease threats. GRITS analyzes these textual data sources by identifying, extracting, and succinctly visualizing critical epidemiologic information and suggests possible associated infectious diseases. Via GRITS' web-interface, infectious disease analysts can examine dynamic visualizations of GRITS' analyses, perform powerful queries of an index of over 250,000 infectious disease reports, and explore related historical infectious disease emergence events. The GRITS API can be used to continuously analyze information feeds and large collections of data and the API enables GRITS technology to be easily incorporated into larger surveillance systems. GRITS is a flexible pluripotent tool that can be modified to conduct sophisticated report triaging, expanded to include customized alert systems, or tailored to address other surveillance needs.

INTRODUCTION

Infectious diseases pose a significant threat to global health and economic stability. , Due to extensive globalization and urbanization, infectious diseases can spread at unprecedented rates. Small and localized infectious disease threats can rapidly become international catastrophes, as demonstrated by Severe Acute Respiratory Syndrome (SARS) in 2003, influenza (H1N1A) in 2009, and Ebola Virus Disease in 2014. , , Early detection of emerging threats through disease surveillance is critical to implementing effective responses.

Traditional disease surveillance relies predominantly on local clinicians, laboratory technicians, and public health practitioners to detect infectious disease outbreaks. Once detected, these outbreaks are communicated to regional, national and international public health organizations for evaluation and response. However, traditional disease surveillance faces a serious challenge.

The biosurveillance systems' ability to detect and collect information is dependent on regional healthcare infrastructure that is heterogeneous in quality and sophistication. Disease threats that emerge in regions with poor healthcare infrastructure may be detected slowly or missed completely. This problem was illustrated recently in the ongoing Ebola Virus Disease epidemic in West Africa, in which the etiological agent was not identified until 85 days after the first case.

The burgeoning field of digital disease detection (DDD) attempts to complement traditional disease surveillance by using software to expand data collection to non-traditional sources, and automate epidemiologic analyses.

DDD is an innovative, but innately challenging field. Although, the quantity of digital information is immense, the signal to noise ratio is very low. To curate data sources for significant public health information, most DDD tools require substantial human capital. This dependency may be limiting and makes it difficult to provide disease detection that is superior (faster and more accurate) to traditional surveillance systems. Conversely, some DDD tools make extensive use of NLP or machine learning algorithms to predict disease risk, but do not incorporate sufficient public health or medical expertise. In some cases, like early experiences with Google Flu Trends, this has led to inaccurate forecasts of infectious disease. DDD tools that enable analysts to examine non-traditional information sources more efficiently are critically needed.

GRITS Disease Classification System

GRITS uses binary logistic regression classifiers to identify infectious diseases that are most likely to be associated with a given text sample. Binary logistic regression is a classification method that predicts a binary response variable (whether a disease label applies to submitted text) based on a number of observed predictor variables (counts for each feature in submitted text).

A separate binary logistic regression classifier is built for each disease within the GRITS disease ontology. Each classifier is trained to estimate the probability that a text passage is associated with that particular disease, based on a vector of features extracted from processing the text. The classifiers return an estimated probability from 0 to 1, with a score of 0 predicting no likelihood, and a score of 1 indicating certainty. This score is referred to as the confidence disease classification.

Classifier Training Process

- Classifiers are trained separately for each disease, containing an index of 253,736 articles (subset of the total);
- HealthMap articles from a 2-3 year period that were assigned disease labels by HealthMap;
- Articles labeled with a disease classifier's particular disease are positive training examples and articles labeled with other diseases are negative examples;
- Each feature is assigned a real-numbered weight, the magnitude of which can be interpreted as how strongly predictive the presence of that feature is of the presence of that disease; and,
- SciKitLearn relies on LIBLINEAR's implementation to determine feature weights using a dual coordinate descent method to find the feature weights that optimize predictive accuracy. ,

Classifiers

- logistic regression
- OneVsRest from the Python package SciKitLearn

Features Ontology

Features used in the GRITS classification method consist of key words significant to public health compiled from available ontologies and libraries, or created specifically for GRITS (Table 1). Relationships between keywords are determined manually, and using the Wordnet ontology.

Ontology	Contents
The Disease Ontology	Human disease related terms, phenotypic characteristics, medical vocabulary disease concepts
Wordnet	English language ontology that maps word relatedness
Biocaster Ontology	General disease ontology
USGS Topographic Feature Vocabularies	Environmental factors
HealthMap Disease Labels	Diseases identified as significant by HealthMap and used for their disease labels
GRITS Ontology	Curated ontology of symptoms, control measures, descriptions of infected individuals, diseases, disease categories, environmental factors, hosts, host uses, modes of disease transmission, occupations, disease risks, vectors, and zoonotic types

Testing and Validation

Micro averaged f-scores across all disease labels are used to evaluate each disease classifier. This method weights labels equally, regardless of prevalence. This makes the method optimal for GRITS, because all diseases are weighted equally, regardless of their prevalence in the training set.

GRITS Search Feature

Elasticsearch is used for the GRITS search feature. The full text of the index of over 250,000 HealthMap articles is queried during each search. Elasticsearch assigns relevance scores to individual terms using TF-IDF based models. Metadata obtained through GRITS analysis of each article in the index is searched for specific purposes, including disease keyword searches, and sorting by date and country.

Translation

Prior to analysis, non-English text is translated using Bing Translate.

Equations describing the general form of the logistic regression model

Total feature contribution to a disease classification

$$\pi_i = x_i \cdot \beta$$

Probability value

$$\pi_i = \frac{e^{x_i \cdot \beta}}{1 + e^{x_i \cdot \beta}}$$

$$\pi_i = \frac{e^{x_i \cdot \beta}}{1 + e^{x_i \cdot \beta}}$$

Logit algorithm, logarithm of the odds ratio of the probability

$$\mathrm{logit}(\pi_i) = \log \frac{\pi_i}{1 - \pi_i},$$

Given:

π_i = estimated dependent variable value

x_i = feature value vector

β = vector of feature weights

Determining the relative contributions of features to a given diagnosis

Generate scores for each keyword involved in a given diagnostic analysis

Normalize all feature scores for each individual disease classifier analysis to unit vectors

Multiply by the estimated probability of each disease

$$score_{i,j} = \frac{x_i \cdot w_j}{||x_i \cdot w_j||}$$

Given:

article i and disease j.

pi hat = estimated probability of the disease

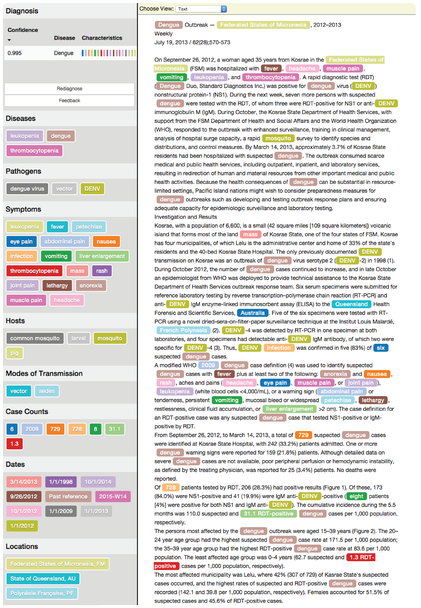
w = feature weight vector for the disease

x = feature count vector for the article

|| .. || Represents the norm (length) of the vector

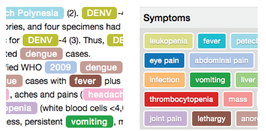
RESULTS

The results of GRITS analyses for a given text sample are displayed in the diagnostic dashboard view of the web-interface.



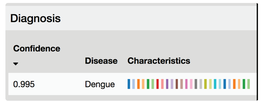
Diagnostic dashboard view showing the GRITS analysis of a report on an outbreak of Dengue Fever. The results of the GRITS disease classification are displayed in the top left. Here, Dengue fever has been identified as the disease the article is most likely concerned with.

- Provides the GRITS ranked diagnosis of potential infectious diseases associated with submitted text, ranked by confidence score.
- Succinctly visualizes the important information extracted from a text sample. This includes disease related keywords (disease, pathogens, symptoms, modes of transmission, and hosts), dates, locations and case counts.
- Information is highlighted in an annotated article view that can be filtered by category.
- Locations are plotted in a map view, and a histogram of temporal information is shown in a timeline



Important information is annotated in the text (left) and visualized by category (right).

Locations are plotted in a map view within the diagnostic dashboard. Here the location "Palikir" has been extracted from the text and is resolved on this map.



The results of a GRITS disease classification as seen in the dashboard view. Shows the disease most likely associated with submitted text, along with the confidence score for the classification. The "Characteristics" column provides a visualization of the keywords that contributed to the classification. Each colored column represents a distinct keyword identified and annotated in the submitted text with the corresponding color.

The keywords that contribute to each diagnosis and their relative weights can be inspected from within the dashboard.

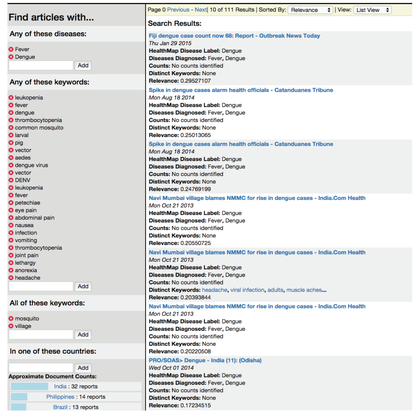
Disease	animal	bite	bleeding	blood	contact with infected	crimean congo hemorrhagic fever	endemic	exposure	fever	hemorrhagic fever	infected animal	male	rash	risk	surveillance	tick	tickborne	ticks	turkey	veterinarians	viral	virus
Crimean-Congo Hemorrhagic Fever	0.059	0.168	0.179	0.072	0.032	0.44	0.161	0.035	0.046	0.293	0.049	0.034	0.037	0.074	0.034	0.427	0.022	0.194	0.057	0.063	0.211	0.009

Detailed diagnosis report showing the relative contributions of different keywords to a disease classification. The disease being classified in this example is Crimean-Congo Hemorrhagic Fever.

The GRITS disease classification system has an overall precision of 64% and recall of 63%

Precision	Recall
64%	63%

A simple, customizable interface is used to query the over 250,000 GRITS analyzed infectious disease related articles.



- Articles can be searched by any combination of the following information; diagnosed disease, keyword, publication date range, and/or country.

  Search results can be viewed spatially or in a listview.
- Results can be displayed by relevance, oldest first, or newest first.
- A query of articles similar to the text sample is automatically conducted based on the results of the text sample's GRITS analysis.
- Meta-data is displayed for each search result.

List view of a GRITS search. Keywords identified in the submitted text are displayed in the left column. Histograms of the countries and dates associated with the search results are displayed below the keywords.

GRITS API applies diagnostics tools continuously on large collections of data and allows GRITS to be integrated into any biosurveillance application

Extracted keywords are annotated with their occurrences in the original text, so they may be easily highlighted in a user interface, indexed for search, used as features for statistical evaluation or machine learning, or subjected to further analysis.

DISCUSSIONS

In the hands of the astute public health analyst, GRITS is a powerful tool for infectious disease surveillance that allows users to efficiently monitor non-traditional data sources for infectious disease threats. GRITS can be expanded to incorporate a customizable triaging system that curates text sources temporally, spatially, by diagnosed disease, or by public health keyword. Additional ontologies can be created to train GRITS to make educated conclusions on additional complex variables besides disease, like pathogen class, report risk level, or the emergence of a novel pathogen. An alert system could be built into GRITS to warn users of potentially dangerous clusters of reports. Additionally, through the GRITS API, the tool can be incorporated into larger surveillance systems, like the Defense Threat Reduction Agency's prototype Biosurveillance Ecosystem, and run continuously on those systems data feeds. Through the web-interface users can evaluate the functionality, and public health foundations of GRITS, increasing the transparency of the tool. This will lead to scrutiny and the refinement, and continued development of GRITS. GRITS, and surveillance systems like it, may be able to fill in the gaps in traditional global biosurveillance systems.

EcoHealth Alliance

AUTHORS Andrew G. Huff[1], Amy M. Slagle[1], Russell S. Horton[1], Nathaniel Breit[1], Zachary S. Gold[1], Jonathan Goley[1]

PARTNERS International Society for Infectious Disease, Kitware, ProMED-mail, Epidemic, HealthMap

FUNDER Defense Threat Reduction Agency

[1]EcoHealth Alliance
460 West 34th Street - 17th Floor
New York, NY 10001-2320

[1]Morens DM, Fauci AS. PLoS Med. 2008. The challenge of emerging and re-emerging infectious diseases. Nature.430:242-249.

[2]Kahn LH. 2006. Philos Ethics Humanit Med Phil Pract Med Ethics Bioethics Humanities. 1:16. doi:10.1186/1747-5341-1-16.

[3]Morens DM, Folkers GK, Fauci AS. 2004. The challenge of emerging and re-emerging infectious diseases. Nature.430:242-249.

[4]Morens DM, Fauci AS. 2013. Emerging diseases in a complex world: the challenges of scientific and public health vigilance. J Infect Dis. 207(7):1062-65. doi:10.1093/infdis/jit008.

[5]Brownstein JS, Freifeld CC, Madoff LC. 2009. Digital disease detection—harnessing the Internet for public health surveillance. N Engl J Med. 360(2):2353-2355. doi:10.1056/NEJMp0900702.

[6]Brownstein JS, Freifeld CC, Reis BY, Mandl KD. 2008. Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. PLoS Med. 5(7):e151. doi:10.1371/journal.pmed.0050151.

[7]Hay SI, George DB, Moyes CL, Brownstein JS. 2013. Big data opportunities for global infectious disease surveillance. PLoS Med. 10(4):e1001415. doi:10.1371/journal.pmed.1001415.

[8]Lee EG, Boseman L, Chen M, Olsen A, Wilson K, Morin C, Mykhalovskiy E, Smith G, Lem W. 2007. Surveillance for emerging infectious diseases of small regional significance using informal sources. Can J Infect Dis Med Microbiol. 18(5):377-382.

[9]Madoff LC. 2004. ProMED-mail: an early warning system for emerging diseases. Clin Infect Dis. 39(227):227-232. doi:10.1086/422002.

[10]Madoff LC, Woodall JP. 2005. The Internet and the global monitoring of emerging diseases: lessons from the first 10 years of ProMED-mail. PLoS Med. 2(2):e162. doi:10.1371/journal.pmed.0020162.

[11]Carrigan PJ, Jr, Pavlin JA, Hyman WR, Wasieloski LP, Jr, Wasieloski HP, Proctor JX, Nichols DK, Nelson EL, Kruszon-Moran D, Shaffer RA, Scherer MJ, Blair PJ. 2007. Emerging infectious disease surveillance: guidance for strengthening international health surveillance through improved reporting. Emerg Health Threats J. 2:e10. doi:10.3134/ehtj.07.010.

[12]Yan SJ, Chu A, Lim TM. 2007. Predicting infectious disease for biofeedback and gaming. Proceedings of the 18th International Conference on Pattern Recognition (ICPR'08). Tampa, Florida: IEEE. 978-1-4244-2174-9/08/25.00.2008 IEEE. 2:1-4. doi:10.1109/ICPR.2008.4761149.

[13]Chan EH, Sahai V, Conrad C, Brownstein JS. 2011. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. PLoS Negl Trop Dis. 5(8):e1090. doi:10.1371/journal.pntd.0001090.

[14]Polgreen PM, Chen Y, Pennock DM, Nelson FD, Weinstein RA. Clin Infect Dis. 2008. Using internet searches for influenza surveillance. 47(2):443-444. doi:10.1086/589754.

[15]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[16]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[17]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[18]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[19]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[20]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[21]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[22]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[23]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[24]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[25]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[26]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[27]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[28]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[29]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[30]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[31]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[32]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[33]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[34]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[35]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[36]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[37]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[38]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[39]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[40]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[41]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[42]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[43]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[44]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[45]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[46]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[47]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[48]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[49]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[50]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[51]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[52]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[53]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[54]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[55]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[56]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[57]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[58]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[59]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[60]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[61]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[62]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[63]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[64]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[65]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[66]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[67]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[68]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[69]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[70]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[71]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[72]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[73]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[74]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[75]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[76]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[77]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[78]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[79]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[80]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[81]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[82]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[83]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[84]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[85]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[86]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[87]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[88]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[89]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[90]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[91]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[92]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. 2014. Using clinical logistic models to improve syndromic detection of influenza epidemics. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.d1b0ba82f5343f6a2f0a0f6f2f8e2ed5.

[93]Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature.457:1012-1014. doi:10.1038/nature07634.

[94]Polgreen PM, Chen Y, Pennock DM, Nelson FD. 2008. Using internet searches for influenza surveillance. Clin Infect Dis. 47(2):443-444. doi:10.1086/589754.

[95]Wilson K, Brownstein JS. 2009. Early detection of disease outbreaks using the Internet. CMAJ. 180(771):771-775. doi:10.1503/cmaj.090026.

[96]Eysenbach G. 2009. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and transaction data. J Med Internet Res. 11(11):e11001. doi:10.2196/jmir.1157.

[97]Eysenbach G. 2011. Infodemiology and Infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 40(5S2):S228-S230. doi:10.1016/j.amepre.2011.02.006.

[98]Paul MJ, Dredze M, Broniatowski D. 2014. Twitter improves influenza forecasting. PLOS Currents Outbreaks. doi:10.1371/currentoutbreaks.90743b0bf8958f8fabf66a23a7f9d96f.

[99]Lampos V, Yom-Tov E, Pebody R, Cox IJ. 2015. Assessing the impact of a health intervention in the crowdsourced detection of influenza epidemics. arXiv:1502.00928v1[stat.AP].

[100]Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brown