


# Can Users Search Trends Predict People Scares or Disease Breakout? An Examination of Infectious Skin Diseases in the United States

Infectious Diseases: Research and Treatment  
Volume 13: 1–7  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1178633720928356



Rand Obeidat<sup>1</sup> , Izzat Alsmadi<sup>2</sup>, Qanita Bani Bakr<sup>3</sup>  
and Laith Obeidat<sup>4</sup>

<sup>1</sup>Department of Management Information Systems, Bowie State University, Bowie, MD, USA.

<sup>2</sup>Department of Computing and Cyber Security, Texas A&M University-San Antonio, San Antonio, TX, USA. <sup>3</sup>Computer Science, Jordan University of Science and Technology, Irbid, Jordan.

<sup>4</sup>Jordanian Royal Medical Services, Amman, Jordan

## ABSTRACT

**BACKGROUND:** In health and medicine, people heavily use the Internet to search for information about symptoms, diseases, and treatments. As such, the Internet information can simulate expert medical doctors, pharmacists, and other health care providers.

**AIM:** This article aims to evaluate a dataset of search terms to determine whether search queries and terms can be used to reliably predict skin disease breakouts. Furthermore, the authors propose and evaluate a model to decide when to declare a particular month as Epidemic at the US national level.

**METHODS:** A Model was designed to distinguish a breakout in skin diseases based on the number of monthly discovered cases. To apply this model, the authors correlated Google Trends of popular search terms with monthly reported Rubella and Measles cases from Centers for Disease Control and Prevention (CDC). Regressions and decision trees were used to determine the impact of different terms to trigger the occurrence of epidemic classes.

**RESULTS:** Results showed that the volume of search keywords for Rubella and Measles rises when the volume of those reported diseases rises. Results also implied that the overall process was successful and should be repeated with other diseases. Such process can trigger different actions or activities to be taken when a certain month is declared as "Epidemic." Furthermore, this research has shown great interest for vaccination against Measles and Rubella.

**CONCLUSIONS:** The findings suggest that the search queries and keyword trends can be truly reliable to be used for the prediction of disease outbreaks and some other related knowledge extraction applications. Also search-term surveillance can provide an additional tool for infectious disease surveillance. Future research needs to re-apply the model used in this article, and researchers need to question whether characterizing the epidemiology of Coronavirus Disease 2019 (COVID-19) pandemic waves in United States can be done through search queries and keyword trends.

**KEYWORDS:** Disease breakout prediction, search queries, Google trends, epidemiology

**RECEIVED:** December 29, 2018. **ACCEPTED:** April 29, 2020.

**TYPE:** Original Research

**FUNDING:** The author(s) received no financial support for the research, authorship, and/or publication of this article.

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHOR:** Rand Obeidat, Department of Management Information Systems, Bowie State University, 14000 Jericho Park Rd., Bowie, MD 20715, USA. Email: robeidat@bowiestate.edu

## Introduction

Infectious skin diseases encompass a vast array of conditions that range in severity from mild to life-threatening. The clinical presentation of infectious skin diseases varies based on the type of pathogen involved, the skin layers and structures affected, and the underlying medical condition of the patient. Infectious skin diseases represent common diagnoses made by dermatologists, by primary care physicians, and in the emergency room.<sup>1</sup> Rubella, in particular, though a mild, vaccine-preventable skin disease, is of high public health importance owing to the teratogenic effects that can result from congenital rubella infection (CRI), leading to miscarriage, fetal death, or birth of an infant with congenital rubella syndrome (CRS).<sup>2,3</sup>

Rapidly identifying an infectious disease outbreak is critical, both for effective initiation of public health intervention

measures and timely alerting of government agencies and the general public.<sup>4</sup> A vast amount of real-time information about infectious disease outbreaks can be found in various forms of Web-based data streams. Studies show that health care providers rely on online search results in obtaining more information about diseases, symptoms, drugs, and other related information. Also, the research showed that doctors find searching online very helpful to get information about tracking geographical locations of disease. Google search queries are the most commonly used data source for search studies around the world. For example, Google's search engine has been used to detect influenza epidemics in areas with a large population of web search users because of its high correlation with the percentage of physician visits if a patient has influenza-like symptoms.<sup>5</sup>



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Research studies sought to study the association between disease outbreak and online search keywords and terms. For example, a study by Polgreen et al<sup>6</sup> examined the relationship between searches for influenza and actual influenza occurrence. Another study by Yom-Tov and Fernandez-Luque<sup>7</sup> collected data from a major Internet search engine, while people seek information about the Measles, Mumps, and Rubella (MMR) vaccine. The authors focused on developing an automated way to score Internet search queries and web pages to examine how people use Internet search engines to garner information on vaccination.

Recognizing the need for up-to-date data to inform researchers, policymakers, public stakeholders, and health care providers if search queries can be used to reliably predict skin disease breakouts, we correlated Google Trends popular search terms with monthly reported Rubella and Measles cases from 2004 to 2018. So, this study provides analysis and evaluation for the association between monthly reported Rubella and Measles cases and Google Trends popular search terms that can be used to predict a future outbreak of infectious skin disease case.

### Related Work

Several research studies used Google Trends to answer research questions within health care domains.<sup>8</sup> Some of these studies examined and confirmed the correlation between disease outbreaks and online search keywords and trends.<sup>9</sup> In 2009, Ginsberg et al<sup>5</sup> stated that the Google Flu Trends predictions were 97% accurate compared with the Centers for Disease Control and Prevention (CDC) data. CDC was also testing Google Flu Trends in the United States, and the preliminary finding suggests that Google Flu Trends can detect regional outbreaks of influenza 7 to 10 days earlier than conventional CDC surveillance.<sup>9,10</sup> The correlation between Google Trends and diseases surveillance was also assessed in several countries such as India,<sup>11</sup> South Korea,<sup>12</sup> South China,<sup>13</sup> and Spain.<sup>14</sup> However, those studies did not propose any unique search terms that can correlate with diseases predictions.

Google Trends were deployed to detect/estimate many disease outbreaks such as influenza,<sup>15,16</sup> Dengue,<sup>17,18</sup> Ebola,<sup>19</sup> and Lyme.<sup>20</sup> Few studies analyzed skin-related diseases using Google Trends. Bloom et al<sup>21</sup> extracted data from Google Trends to evaluate whether population inquisitiveness on melanoma and skin cancer was correlated with a lower incidence. They found that the general populations' interest in learning about skin cancer increases during the summer month. Hopkins and Secrest<sup>22</sup> used Google Trends data queried using several search terms (sunscreen, sunburn, skin cancer, and melanoma) in the United States. Then, time-matched search term data were correlated with melanoma outcomes data from Surveillance Epidemiology and End Results Program and United States Cancer Statistics. In another study, Hopkins and Secrest<sup>22</sup> explored international trends in English-speaking

countries including (United States, United Kingdom, Canada, Australia, and New Zealand) several search terms that are used to better guide skin cancer prevention campaigns. Hopkins and Secrest<sup>23</sup> assessed the correlations between search terms, time, and melanoma outcomes for each country. None of the previous studies correlated Google Trends popular search terms with certain infectious skin diseases including Rubella and Measles reported from CDC. In this work, Google Trends was used to propose unique search terms that can correlate with Epidemic disease prediction. For this purpose, we collected data and we used machine learning methods to evaluate a dataset of search terms to determine if search queries and terms can be used to reliably predict skin disease breakouts.

### Methods

In this study, it is important to use different classifiers to have more confidence in the results and compare those different classifiers based on accuracy. Therefore, we have created a new supervised classification model that uses a Support-Vector Machine (SVM) model, linear regression (LR), and Decision Tree (DT) to evaluate each disease breakout prediction. In the first part of this section, we provide an overview of the model we used. Then, we explain the data used in this study. We finish this section by presenting the proposed model for disease epidemic classification and algorithm.

#### An SVM

Using SVM is a supervised machine learning technique that is widely used in classification and regression problems. The main objective of SVM algorithm is to find a hyperplane in an  $N$ -dimensional space that distinctly classifies the data points, where  $N$  is the number of features. To separate 2 classes of data points, infinite number of hyperplanes could be found. In the SVM, the main objective is to find a plane that has the maximum margin.

A separating hyperplane can be written as the following equation:

$W \bullet X + b = 0$ , where  $W = \{w_1, w_2, \dots, w_n\}$  is a weight vector and  $b$  a scalar (bias). For example, in two-dimensional (2D) it can be written as:

$$w_0 + w_1 x_1 + w_2 x_2 = 0$$

The hyperplane defining the sides of the margin is as follows:

$$H1: w_0 + w_1 x_1 + w_2 x_2 \geq 1 \quad \text{for } y_i = +1,$$

$$H2: w_0 + w_1 x_1 + w_2 x_2 \leq -1 \quad \text{for } y_i = -1$$

Any training tuples that fall on hyperplanes H1 or H2 are support vectors.

We used SVM as a classifier where in the applied SVM model, rows represent months and columns represent relevant Google Trends keywords that were extracted from several

**Table 1.** Proposed model for disease epidemic classification.

ORIGINAL DISEASE COUNTS	(CURRENT MONTH-PREVIOUS MONTH)/PREVIOUS MONTH	IF (SUM OF LAST 3MONTHS $\geq$ 0.5, 1, ELSE 0)
24	3.80	1.00
39	0.63	1.00
21	-0.46	1.00
36	0.71	1.00
5	-0.86	0.00

cycles. Dates used in SVM were matched for Google Trends keywords and CDC reported diseases. Table 1 shows the last 2 columns of the SVM matrix that we created from the number of reported cases, one of the columns as a continuous value and the last column as a binary target, based on our Epidemic formula.

### LR

In the LR, given  $N$  observations, where an output vector  $Y$  with dimension  $N \times 1$  and  $p$  inputs  $X_1, X_2, \dots, X_p$ , where each input vector being of dimension  $N \times 1$ , LR assumes that the regression function  $E(Y|X)$  is linear in the inputs.  $Y$  is computed based on the following equation:

$$Y = \beta_0 + \sum_{i=0}^p \beta_i X_i + \epsilon,$$

where  $\epsilon$  is the error term. LR model is applied to predict diseases, such as Measles, using many search terms (in terms of their significance and estimated values).

### Decision Tree

Decision Tree is a supervised learning algorithm which has a flowchart-like tree structure. In the DT, each internal node (non-leaf node) represents a test on an attribute. Each branch is an outcome of the tested condition of on an attribute, and each leaf node or terminal node holds a class label. Decision Tree classification model was employed to study the different terms impact on diseases predication.

### Data sources

*Centers for Disease Control and Prevention data.* Data were collected on reported cases for each disease in the United States over the period from 2004 to 2018. Reported cases of diseases by months were collected and maintained from public CDC publications. Then, the authors did some data transformation to aggregate data across all United States to monthly basis to match Google Trends data.

*Google Trends data.* Initial sets of relevant keywords were created to each disease and used them to extract Google Trends data. Specifically, in this article, the authors produced a dataset of popular search terms for 2 infectious skin diseases such as Rubella and Measles that can be used to predict future skin disease breakouts. Google reports search terms on a monthly basis were accumulated.

*The proposed model for disease epidemic classification and algorithm.* The authors designed an “epidemic” model to distinguish a breakout of diseases based on the number of monthly discovered cases, as well as to decide whether a certain month counts as “Epidemic.” The authors observed final values collected from CDC and Google Trends and then decided to make the cut-off in Epidemic class based on the reported values for each disease. Epidemic class value will be 1, or else will be zero (ie, yes, 1 or no, 0). Table 1 shows that the level of increase in a month from previous month was calculated using this formula: (Current month-previous month)/previous month. Transient month increase or decrease was eliminated. As a result, this increase will be considered Epidemic if it occurs in 3 consecutive months. The authors made this algorithm as an approximation of how to flag a month as an epidemic month.

This model was applied and validated to distinguish a breakout in Rubella and Measles skin diseases based on the number of monthly discovered cases. The process of the creation of Rubella popular search terms has the following main steps:

- Start by evaluating the single term (for each skin disease) in Google Trends across the available period (from 2004 to 2018), with a total of 168 records. The numbers given for the popular terms are in the form of percentages (ie, from 0 to 100), rather than actual volumes of search terms.
- In the second step of keyword selection, we use Google Trends to expand our search terms. Starting Step 1, we analyze the “related queries” section in Google Trends and we extract the “Top” and “Rising” search terms as shown in Table 2.
- As shown in Table 2, Google Trends distinguishes terms breakout and rising keywords in terms of how quick and long such terms have been on the search rise.

For a search term to be selected in the collected dataset, it should have the following inclusion criteria: (1) Search terms are extracted only from the “related queries” section in Google Trends results. It should be listed in the “Top” or listed with more than 90% “Rising” terms. (2) It should be repeated for more than 4 times in the “related queries” from different results (ie, from different initial search terms). This was necessary to eliminate out term that are “outliers” (If they show up in only one related term).

**Table 2.** A sample of Rubella “Top” and “Rising” terms in Google Trends.

TOP	RISING
Rubella	Immunoglobulin G
Measles rubella	Titer
Measles	Blood
The measles	Side effect
Mumps	Scarlet fever
Mumps symptoms	Influenza
Skin rash	Blood test
Rubella in pregnancy	Fifth disease
Rubella blood test	Fetus
Symptoms of rubella	Varicella vaccine
Titer test	Medical diagnosis
Chickenpox	Varicella zoster virus
Varicella vaccine	Tuberculosis

### Experiments and analysis

An SVM model was used to evaluate each disease breakout prediction based on collected features in the different experiments. Correlations (*Pearson* and *Spearman*) were used between Google Trends of popular search terms and monthly reported Rubella and Measles cases from CDC. In addition, regressions and DTs were used to determine the impact of different terms to trigger the occurrence of epidemic classes.

### Results

In our SVM model (Table 3, Rubella SVM sample), rows represent months and columns represent relevant Google Trends keywords that were extracted from several cycles. Dates were matched for Google Trends keywords and CDC reported diseases. The count column was retrieved from Google Trends and represents the popularity of those relevant keywords in that particular month.

Records represent monthly data for both disease volume and Google Trends selected keywords. For Rubella disease, the volume of reported cases was small. In addition, in our dataset, we did not find reports for many other months (ie, missing values). This impacted overall prediction accuracy. Table 4 shows the results of LR prediction on Rubella SVM. We showed search terms with the lowest *P* values (ie, significant prediction results). However, their estimate values are low, which indicates a low overall impact on disease prediction.

Accuracy of prediction for Measles LR model is better for many search terms (in terms of their significance and estimate values; Table 5). One main reason for such better accuracy is

the large number of reported cases for Measles and also the fact that we have much fewer missing values for Measles' case.

For each one of the experiments to extract relevant keywords, we evaluated correlations (*Pearson* and *Spearman*) between popular terms of Google Trends and Disease arrays. In terms of correlation, no significant positive or negative correlation is shown in the volume of those terms and cases volumes. However, the highest keywords for Rubella in terms of correlation (negative or positive) were Titer, Rubeola, CRS (positive), rubella pregnancy, and rubella rash.

Decision Tree classification model was employed as the model has a categorical target class to study the different terms impact. With more than 95% accuracy, Figure 1 shows overall accuracy metrics. Figure 1 shows a high true positive (TP) rate and a very low false positive (FP) rate which implies very acceptable accuracy in all recorded performance metrics (ie, precision, recall, MCC, receiver operating characteristic [ROC] area, and precision-recall curve [PRC] area).

Due to size limitation, we show a summary snapshot from Measles DT in Figure 2. This figure summarizes search terms with a significant impact on our proposed Epidemic class. This figure shows also minimum weight for the search term to trigger the occurrence of the Epidemic class. In other words, if people are searching for more than this percentage on this particular term, then the rise in this disease is significant. The DT shows the keywords that decide the target class (whether a month is an epidemic or not), their cut-off value to switch the target class from Yes (epidemic) to No, and also how many instances in the dataset in that category.

### Discussion

This article aimed to evaluate a dataset of search terms to determine whether search queries and terms can be used to reliably predict disease breakouts. A model was proposed and evaluated to decide when to declare a particular month as Epidemic at the US national level. In this study, the authors applied the model on 2 infectious skin diseases such as Rubella and Measles.

By using LR as a regression method, we showed that the search terms with the lowest *P* values estimate values that are low, which indicates a low overall impact on disease prediction. By using the LR, we also found that the accuracy level of prediction for Measles is higher than the accuracy of prediction for Rubella using several Search Terms as shown in Tables 4 and 5. In addition, the DT classification model was employed as the model for classification with more than 95% accuracy. The DT model successfully shows that the keywords features can be used to classify whether a month is an epidemic or not with accuracy reach to 95%.

In this study, we found that people search for Rubella and Measles diseases throughout the year. Results showed that the volume of search keywords for Rubella and Measles rises when the volume of reported diseases rises.

**Table 3.** A sample of Rubella SVM model.

MONTH	COUNT	RUBELLA	MEASLES	RUBELLA MEASLES	RUBELLA VACCINE	MUMPS	RUBEOLA
Feb-05	1	52	5	12	24	7	16
Mar-05	1	51	5	21	25	8	15
Apr-05	2	62	5	22	42	8	15
May-05	1	48	5	23	34	8	9
Jun-05	1	39	4	22	35	6	13
Jul-05	2	36	3	15	17	6	9
Aug-05	1	35	3	27	24	7	17
Sep-05	0	43	4	19	22	5	11
Oct-05	0	42	4	14	38	6	9
Nov-05	0	47	5	33	29	7	13
Dec-05	2	47	3	12	24	5	9

**Table 4.** A sample of Rubella linear regression results.

SEARCH TERM	ESTIMATE	STD. ERROR	T-VALUE	P VALUE
Rubeola	0.016649	0.007614	2.186463	0.03069
rubella.antibody	-0.01164	0.005646	-2.06151	0.041379
rubella.vaccine	-0.02079	0.011871	-1.75111	0.082441
roseola.rash	-0.01143	0.006845	-1.67033	0.097418
measles.rubella.vaccine	-0.01286	0.009757	-1.31842	0.189834
rubella.symptoms	0.007534	0.005717	1.317845	0.190025
rubella.pregnancy	0.006707	0.005681	1.180734	0.240005

**Table 5.** A sample of Measles linear regression results.

SEARCH TERM	ESTIMATE	STD. ERROR	T-VALUE	P VALUE
measles.vaccine	8.627712241	2.0750631	4.157807174	6.03E-05
measles.rash	2.59555507	0.669714458	3.875614506	0.00017326
Mumps	0.424433485	0.202515751	2.095804811	0.038183852
measles.in.us	-4.64410727	1.291801125	-3.59506366	0.000470388
the.measles	-5.92620105	2.203501347	-2.68944744	0.008167587
symptoms.of.measles	2.138124834	1.577813253	1.355119074	0.177904054
measles.vaccination	0.98191531	1.263357557	0.777226767	0.43854188
chicken.pox.measles	0.846340881	0.541058991	1.564230324	0.120375129
red.measles	0.65549927	0.395612368	1.656923097	0.100125807
measles.pictures	0.64106948	0.399297903	1.605491729	0.110992415

Due to the small volume of reported cases for rubella, it is found that the accuracy level of prediction for Measles is higher than the accuracy of prediction for Rubella. Despite some

challenges related to missing values in certain months, the results implied that the overall process was successful and should be repeated with other diseases. Such a process can

Correctly Classified Instances	136	95.1049 %
Incorrectly Classified Instances	7	4.8951 %
Kappa statistic	0.8989	
Mean absolute error	0.082	
Root mean squared error	0.2025	
Relative absolute error	16.7656 %	
Root relative squared error	40.952 %	
Total Number of Instances	143	
Ignored Class Unknown Instances	35	

=== Detailed Accuracy By Class ===

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
	0.988	0.098	0.931	0.988	0.959	0.901	0.851	0.738
	0.902	0.012	0.982	0.902	0.940	0.901	0.936	0.828
Weighted Avg.	0.951	0.062	0.953	0.951	0.951	0.901	0.887	0.777

Figure 1. Accuracy of Measles Decision Tree.

```

| | | | | | | | | | rubella <= 38: 0 (7.0/1.0)
| | | | | | | | | | rubella > 38: 1 (3.0)
| | | | | | | | | | measles in us > 2: 1 (3.0)
| | | | | | | | | | the measles > 3
| | | | | | | | | | symptoms of measles > 5
| | | | | | | | | | measles vaccination > 3: 0 (16.0/2.0)
| | | | | | | | | | measles cdc > 3
| | | | | | | | | | varicella > 56: 1 (4.0)
| | | | | | | | | | the measles > 5: 0 (8.0)
| | | | | | | | | | chicken pox measles > 20
| | | | | | | | | | varicella <= 54: 1 (9.0)
| | | | | | | | | | varicella > 54
| | | | | | | | | | red measles > 17
| | | | | | | | | | measles rash > 9: 1 (9.0)
measles symptoms > 5
| | | | | | | | | | mmr vaccine <= 26: 1 (13.0)
| | | | | | | | | | mmr vaccine > 26
| | | | | | | | | | the measles > 17: 1 (2.0)

```

Figure 2. A summary sample of Measles Decision Tree.

trigger different actions or activities to be taken when a certain month is declared as “Epidemic.”

One interesting observation is that the query volumes considerably vary according to the searched term. However, this research has shown great interest in vaccination against Measles and Rubella.

This study has some limitations. At first, we were weighing our options to use US data at the national level or state by state. However, based on data availability, we reported analysis only at US national level. In the future, and based on data availability in the CDC, we will analyze historical data on several years per state. For Google Trends, one major limitation we have to deal with in Google Trends is that Google Trends aggregates

relative not absolute data. All data reported are relative (ie, in percentage from 0% to 100%) rather than actual volumes of search terms.

## Conclusions

In the era of online information overload, can users search trends predict diseases outbreak? To address this question, this study aimed at evaluating a dataset of search terms from 2004 to 2018, by developing and evaluating a model to decide when to declare a particular month as Epidemic at the US national level. The findings suggest that the search queries and keyword trends can be truly reliable to be used for the prediction of disease outbreaks, and search-term surveillance can provide an

additional tool for infectious disease surveillance. Future research needs to re-apply the model used in this article, and researchers need to question whether characterizing the epidemiology of Coronavirus Disease 2019 (COVID-19) pandemic waves in the United States can be done through search queries and keyword trends.

### Author Contributions

Data Collection: RO, IA. Designed the experiments: IA, RO, QBB. Analyzed the data: IA, RO, QBB. Wrote the first draft of the manuscript: RO, IA, QBB. Contributed to the writing of the manuscript: RO, IA, QBB, LO. Agree with manuscript results and conclusions: RO, IA, QBB, LO. Jointly developed the structure and arguments for the paper: RO, IA, QBB. Made critical revisions and approved final version: RO, IA, QBB. All authors reviewed and approved the final manuscript.

### ORCID iD

Rand Obeidat  <https://orcid.org/0000-0001-8271-3829>

### REFERENCES

- Dawson AL, Dellavalle RP, Elston DM. Infectious skin diseases: a review and needs assessment. *Dermatol Clin*. 2012;30:141-151.
- Cutts FT, Robertson SE, Diaz-Ortega J, Samuel R. Control of rubella and congenital rubella syndrome (CRS) in developing countries, Part 1: burden of disease from CRS. *Bull World Health Organ*. 1997;75:55.
- Hahné S, Macey J, van Binnendijk R, et al. Rubella outbreak in the Netherlands, 2004–2005: high burden of congenital infection and spread to Canada. *Pediatr Infect Dis J*. 2009;28:795-800.
- Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. *Can Med Assoc J*. 2009;180:829-831.
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457:1012.
- Polgreen PM, Chen Y, Pennock DM, Nelson FD, Weinstein RA. Using Internet searches for influenza surveillance. *Clin Infect Dis*. 2008;47:1443-1448.
- Yom-Tov E, Fernandez-Luque L. Information is in the eye of the beholder: seeking information on the MMR vaccine through an Internet search engine. *AMLA Annu Symp Proc*. 2014;2014:1238-1247.
- Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google Trends in health care research: a systematic review. *PLoS ONE*. 2014;9:e109583.
- Carneiro HA, Mylonakis E. Google Trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis*. 2009;49:1557-1564.
- Dugas AF, Jalalpour M, Gel Y, et al. Influenza forecasting with Google flu trends. *PLoS ONE*. 2013;8:e56176.
- Verma M, Kishore K, Kumar M, Sondh AR, Aggarwal G, Kathirvel S. Google search trends predicting disease outbreaks: an analysis from India. *Healthcare Inform Res*. 2018;24:300-308.
- Cho S, Sohn CH, Jo MW, et al. Correlation between national influenza surveillance data and google trends in South Korea. *PLoS ONE*. 2013;8:e81422.
- Kang M, Zhong H, He J, Rutherford S, Yang F. Using Google Trends for influenza surveillance in South China. *PLoS ONE*. 2013;8:e55205.
- Valdivia A, Monge-Corella S. Diseases tracked by using Google Trends, Spain. *Emerg Infect Dis*. 2010;16:168.
- Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google Flu Trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PLoS ONE*. 2011;6:e23610.
- Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol*. 2013;9:e1003256.
- Gluskin RT, Johansson MA, Santillana M, Brownstein JS. Evaluation of Internet-based dengue query data: Google Dengue Trends. *PLoS Neglect Trop Dis*. 2014;8:e2713.
- Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Neglect Trop Dis*. 2011;5:e1206.
- Alicino C, Bragazzi NL, Faccio V, et al. Assessing Ebola-related web search behaviour: insights and implications from an analytical study of Google Trends-based query volumes. *Infect Dis Pover*. 2015;4:54.
- Seifter A, Schwarzwalder A, Geis K, Aucott J. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospatial Health*. 2010;4:135-137.
- Bloom R, Amber KT, Hu S, Kirsner R. Google search trends and skin cancer: evaluating the US population's interest in skin cancer and its association with melanoma outcomes. *JAMA Dermatol*. 2015;151:903-905.
- Hopkins ZH, Secrest AM. Public health implications of Google searches for sunscreen, sunburn, skin cancer, and melanoma in the United States. *American Journal of Health Promotion*. 2019;33:611-615.
- Hopkins ZH, Secrest AM. An international comparison of Google searches for sunscreen, sunburn, skin cancer, and melanoma: current trends and public health implications. *Photodermatol Photoimmunol Photomed*. 2019;35:87-92.