

Original Paper

Predicting Norovirus in the United States Using Google Trends: Infodemiology Study

Kai Yuan^{1*}, MD, PhD; Guangrui Huang^{1*}, PhD; Lepeng Wang^{2*}, MD; Ting Wang¹, MD; Wenbin Liu¹, MD; Haixu Jiang¹, MD; Albert C Yang^{3,4,5}, MD, PhD

¹School of Life Sciences, Beijing University of Chinese Medicine, Beijing, China

²School of Humanities, Beijing University of Chinese Medicine, Beijing, China

³Digital Medicine Center, National Yang Ming Chiao Tung University, Taiwan, Republic of China

⁴Department of Medical Research, Taipei Veterans General Hospital, Taiwan, Republic of China

⁵Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA, United States

*these authors contributed equally

Corresponding Author:

Albert C Yang, MD, PhD

Digital Medicine Center

National Yang Ming Chiao Tung University

155 Li-Nong Street, Section 2

Peitou District

Taipei

Republic of China

Phone: 886 28267995

Email: accyang@gmail.com

Abstract

Background: Norovirus is a contagious disease. The transmission of norovirus spreads quickly and easily in various ways. Because effective methods to prevent or treat norovirus have not been discovered, it is important to rapidly recognize and report norovirus outbreaks in the early phase. Internet search has been a useful method for people to access information immediately. With the precise record of internet search trends, internet search has been a useful tool to manifest infectious disease outbreaks.

Objective: In this study, we tried to discover the correlation between internet search terms and norovirus infection.

Methods: The internet search trend data of norovirus were obtained from Google Trends. We used cross-correlation analysis to discover the temporal correlation between norovirus and other terms. We also used multiple linear regression with the stepwise method to recognize the most important predictors of internet search trends and norovirus. In addition, we evaluated the temporal correlation between actual norovirus cases and internet search terms in New York, California, and the United States as a whole.

Results: Some Google search terms such as gastroenteritis, watery diarrhea, and stomach bug coincided with norovirus Google Trends. Some Google search terms such as contagious, travel, and party presented earlier than norovirus Google Trends. Some Google search terms such as dehydration, bar, and coronavirus presented several months later than norovirus Google Trends. We found that fever, gastroenteritis, poison, cruise, wedding, and watery diarrhea were important factors correlated with norovirus Google Trends. In actual norovirus cases from New York, California, and the United States as a whole, some Google search terms presented with, earlier, or later than actual norovirus cases.

Conclusions: Our study provides novel strategy-based internet search evidence regarding the epidemiology of norovirus.

(*J Med Internet Res* 2021;23(9):e24554) doi: [10.2196/24554](https://doi.org/10.2196/24554)

KEYWORDS

norovirus; Google Trends; correlation; outbreak; predictors

Introduction

Norovirus, also named stomach bug or winter vomiting bug, is a kind of contagious virus leading to gastroenteritis [1]. It is responsible for almost 20% of acute gastroenteritis cases worldwide [2]. Especially in cruise ships, norovirus is the most common cause of gastrointestinal illness [3]. The symptoms of norovirus include diarrhea, nausea, and vomiting, and it can be transmitted by contaminated water, contaminated food, and person-to-person contact [4,5]. Polymerase chain reaction and enzyme-linked immunosorbent assay are useful methods to diagnose norovirus infection [6,7]. Thus far, there have been no specific drugs to treat norovirus infection. In the prevention of norovirus, hand washing and disinfectants are important and effective methods. In addition, avoiding contaminated water and contaminated food in restaurants are useful methods to prevent norovirus.

It is important to rapidly recognize and report norovirus infection. In the United States, the US Centers for Disease Control and Prevention (CDC) report and monitor the outbreak of gastroenteritis with an internet search system. In addition, other surveillance systems, such as the Foodborne Diseases Active Surveillance Network, can be used to report the trends of foodborne disease outbreaks [8]. However, these surveillance and reporting systems have a lag of 1 to 2 weeks after an outbreak due to the reporting and verification procedure. Internet search has been a useful method for people to access information immediately. Infodemiology is an emerging field based on public attention, knowledge, opinion, etc [9]. Data from infodemiology could be acquired and analyzed to enhance public health [10] and predict outbreaks of disease such as influenza [11]. In addition, infodemiology could also be used to monitor foodborne disease, gastroenteritis, dengue, etc [12]. Apart from Google search, search engines from other platforms including Facebook and Twitter can benefit infoveillance [13]. Yang et al [14-16] revealed the correlation between internet search trends and suicide incidence and depression. In addition, internet searches have been used to monitor infectious diseases, including influenza and Lyme disease [17,18]. Mavragani et al [19] analyzed 104 published papers published from 2006 to 2016 related to Google Trends and found that the monitoring of online queries can provide insight into human behavior. In addition, Nuti et al [20] analyzed 70 papers published from 2009 to 2013 and discovered the topic domains related to Google Trends were general population behavior, infectious disease, noncommunicable diseases, mental health, and substance use.

In this study, we discovered a correlation between internet search terms and norovirus. The search trend of Google Insight was extracted and used in our study. Certain keywords in the Google search trend included “diarrhea,” “vomiting,” “fever,” etc. These keywords could reflect the characteristics of norovirus infection. We also conducted a monthly time series analysis to discover the relationship between internet search terms and norovirus. We hoped to provide new insight to reveal the epidemic and outbreak of norovirus using internet search data.

Methods

Internet Search Approach

The internet search data relating to norovirus were obtained from Google Trends [21]. Google Trends is a useful website to discover search queries from various regions and countries. This tracking system could also provide detailed analysis of Google searches. Google Trends could provide long time series and geographical search locations. Google Trends is in widespread use to investigate medical topics. We downloaded the data of various terms in monthly granularity. The normalization of data indicated that the values vary from 0 to 100. The value 0 indicated very low search volumes that are not included in the results, while the value 100 indicated very high search volumes. The data are retrieved directly from the Google Trends Explore page in .csv format after keywords are entered and the region, period, and category are selected (see [Multimedia Appendix 1](#)) [22]. We chose geographic locations for search volumes as the states of New York and California and the United States as a whole. A time range from January 1, 2004, to December 31, 2018, was chosen. The category was limited to Health to avoid unrelated results of Google searches. We used 50 search terms to reflect a broad sense of norovirus: “norovirus,” “gastroenteritis,” “diarrhea,” “vomiting,” “dehydration,” “contaminated,” “norovirus infection,” “watery diarrhea,” “contagious,” “contaminated water,” “Norwalk virus,” “noroviruses,” “acute gastroenteritis,” “stomach flu,” “viral gastroenteritis,” “winter vomiting disease,” “stomach bug,” “travel,” “party,” “barbecue,” “cruise,” “oyster,” “bar,” “restaurant,” “wedding,” “hotel,” “motel,” “virus,” “infectious,” “outbreak,” “rotavirus,” “coronavirus,” “influenza,” “food poisoning,” “incubation period,” “fever,” “poison,” “CDC,” “vaccine,” “Chipotle,” “ship,” “hand sanitizer,” “wash hand,” “flu symptom,” “streptococcus,” “antibiotics,” “candidiasis,” “otitis media,” “skin rash,” and “coxsackie virus.” These terms were chosen to indicate the related symptoms of and risk factors for norovirus. These terms could provide positive and negative information about norovirus. The Google Trends data were exhibited on a monthly basis. No patient records or personal information were included in the search results of Google Trends. Google Trends normalizes search data to make comparisons between terms easier. Search results are normalized to the time and location of a query. Each data point is divided by the total searches of the geography and time range it represents to compare relative popularity. The resulting numbers are then scaled on a range of 0 to 100 based on a topic’s proportion to all searches on all topics [23].

Actual Norovirus Cases Acquisition

Actual norovirus cases from California, New York, and the country as a whole were obtained from the National Outbreak Reporting System (NORS) at the CDC [24,25]. We limited the etiology to “norovirus” ranging from 2004 to 2018. In addition, the illnesses in every month varied from January 2004 to December 2018. Cross-correlation analysis was conducted to reveal the relationship between actual norovirus cases and internet search terms.

Statistical Analysis

SPSS for Windows (version 20.0, IBM Corp) was used for statistical analysis. A 2-stage modeling process was conducted to evaluate the correlation between norovirus and internet search trends. Cross-correlation analysis was conducted to reveal the relationship between 2 parameters. Cross-correlation is a method to discover the correlation between 2 time series parameters contemporaneously and with various lagged values. Cross-correlation could help researchers understand the relationship between 2 parameters. In addition, cross-correlation analysis could also help researchers understand whether one parameter was preceded by or followed another. Thus, cross-correlation could estimate the time lag of norovirus and other internet search trend terms. The results of cross-correlation were reported in our study.

We used a multiple linear regression method to recognize the most important predictors of internet search trends. Multiple linear regression was used in the second stage to analyze the internet search terms. A stepwise method in multiple linear regression was conducted to discover significant parameters

correlated with norovirus. It also could help researchers evaluate the variation in monthly norovirus counts explained with significant parameters. A stepwise method could reduce the collinearity that internet search might be intercorrelated. The variance inflation factor was used to evaluate the collinearity of internet search data. Those terms with variance inflation factor >10 were excluded from the model. The summary of multiple linear regression was reported as results. A 2-tailed $P < .05$ was regarded as significant in all analyses.

Results

Basic Features Description

The results of Google Trends in the state of New York from January 2004 to December 2018 were downloaded in our study. Except for the search term “norovirus,” a total of 50 keywords from Google Trends were downloaded. Visual inspection indicated that the value of norovirus in Google Trends peaked around December to February in most years (Figure 1). The basic features of actual norovirus cases are shown in Figure 2.

Figure 1. The basic features of the norovirus Google search.

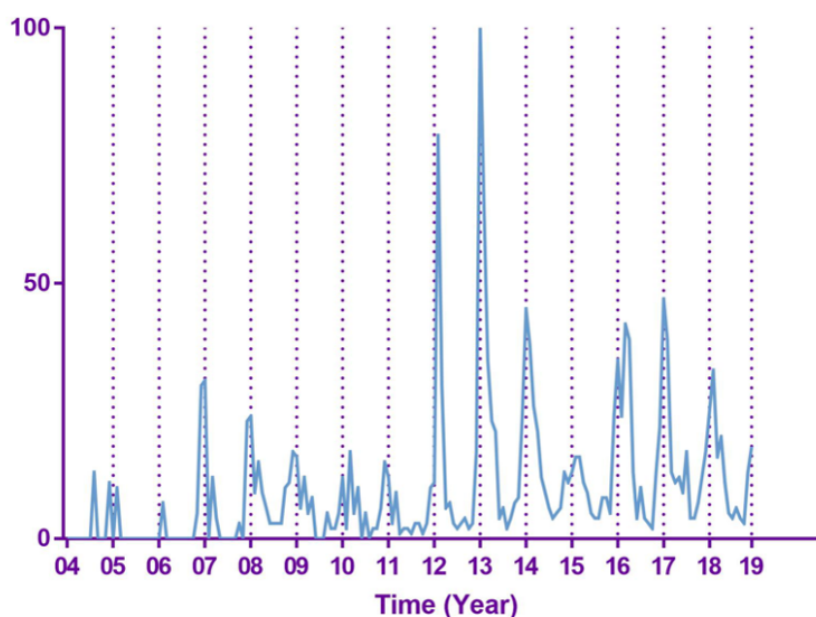
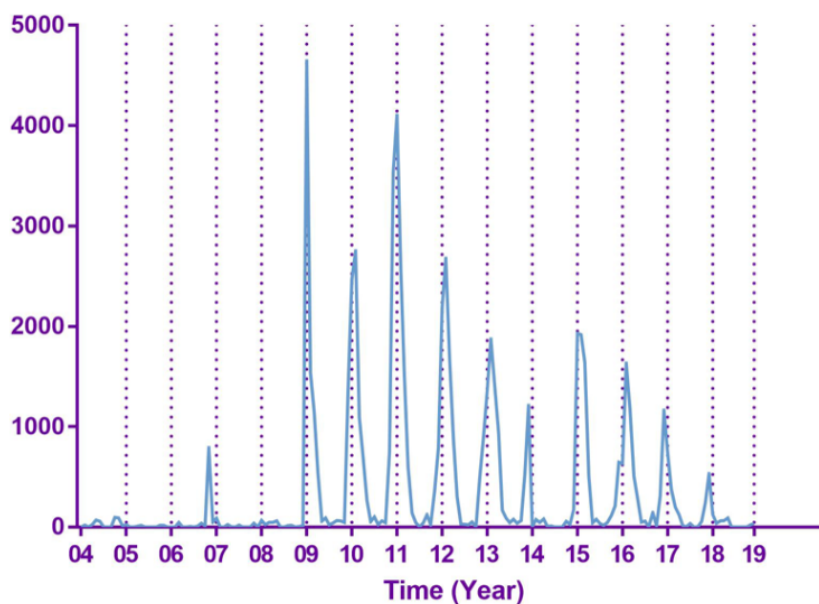


Figure 2. The basic features of actual norovirus cases.



Temporal Correlation in New York

We used cross-correlation analysis to discover the temporal correlation between norovirus and other terms. The internet search terms were included for cross-correlation analysis with norovirus. [Multimedia Appendix 2](#) indicates the results of the cross-correlation of norovirus and other internet search terms. The gray markup shows the lead pattern for norovirus and other internet search terms. The search terms “gastroenteritis,” “watery diarrhea,” “stomach flu,” “winter vomiting disease,” “stomach bug,” and “food poisoning” were coincided with norovirus in Google Trends. The search terms of “contagious,” “travel,” “party,” “restaurant,” “wedding,” “hotel,” “infectious,” and “poison” presented earlier than norovirus Google Trends. The search terms of “dehydration,” “bar,” “Chipotle,” and

“coronavirus” presented several months later than norovirus Google Trends.

After discovering the temporal correlation between norovirus and other terms, we used multiple linear regression with the stepwise method to recognize the most important predictors of internet search trends and norovirus. The search terms that did not have a relationship with norovirus were excluded from the cross-correlation analysis. The internet search terms in [Multimedia Appendix 2](#) were included for multiple linear regression analysis with a norovirus search. The results indicated that searches for “stomach flu,” “stomach bug,” “winter vomiting disease,” “gastroenteritis,” and “cruise” were significant predictors of norovirus search and accounted for 55% (R^2) of the variation in norovirus search ([Table 1](#)).

Table 1. Stepwise regression model of norovirus and correlated internet search trends in New York State.

Search term	Model summary		
	Standardized β	SE	<i>P</i> value
Stomach flu	0.24	0.06	<.001
Stomach bug	0.29	0.06	<.001
Winter vomiting disease	0.24	0.05	<.001
Gastroenteritis	0.13	0.05	.01
Cruise	0.16	0.07	.03

Norovirus also refers to stomach flu, stomach bug, or winter vomiting disease. We excluded these terms to identify the possible predictors of norovirus. Therefore, “stomach flu,” “stomach bug,” “winter vomiting disease,” and “norovirus infection” were excluded in this multiple linear regression analysis. The results showed that searches for “fever,” “gastroenteritis,” “poison,” “cruise,” “wedding,” “watery

diarrhea,” and “acute gastroenteritis” accounted for 40% (R^2) of the variation in the norovirus search ([Table 2](#)). We found that the symptoms of gastroenteritis, including vomiting and watery diarrhea, were important factors that were significantly correlated with norovirus. In addition, activities such as cruise and wedding were important in norovirus outbreak.

Table 2. Stepwise regression model of norovirus and correlated internet search trends excluded alternate names of norovirus in New York State.

Search term	Model summary		
	Standardized β	SE	<i>P</i> value
Fever	0.33	0.08	<.001
Gastroenteritis	0.19	0.06	.001
Poison	-0.16	0.05	.002
Cruise	0.32	0.10	.001
Wedding	-0.34	0.11	.002
Watery diarrhea	0.09	0.04	.02
Acute gastroenteritis	0.09	0.04	.04

Identifying Predictors of Actual Norovirus Cases

In this study, we evaluated the temporal correlation between actual norovirus cases in New York and internet search terms in New York. The data from actual norovirus cases were downloaded from the CDC. As shown in [Multimedia Appendix 3](#), the search terms “norovirus,” “gastroenteritis,” “norovirus infection,” “contagious,” and “acute gastroenteritis” were coincided with actual norovirus cases in New York. The search terms “travel,” “party,” “barbecue,” and “cruise” presented earlier than actual norovirus cases in New York. The search terms “dehydration,” “winter vomiting disease,” “outbreak,” and “rotavirus” presented one or several months later than actual norovirus cases in New York.

Furthermore, we evaluated the temporal correlation between actual norovirus cases in California with internet search terms in California. The search terms including “barbecue,” “oyster,” “bar,” and “infectious” were presented one or several months earlier than actual norovirus cases in California. The search terms including “diarrhea,” “vomiting,” “contagious,” “contaminated water,” and “Norwalk virus” were presented one or several months later than actual norovirus cases in California. The detailed information is shown in [Multimedia Appendix 4](#).

In addition to New York and California, we also evaluated the temporal correlation between actual norovirus cases with internet search terms within the United States. The search terms “vomiting,” “contaminated,” “norovirus infection,” and “contagious” coincided with actual norovirus cases in the United States. Search terms including “travel,” “party,” “cruise,” “restaurant,” and “wedding” presented one or several months earlier than actual norovirus cases in the United States. The search terms “gastroenteritis,” “diarrhea,” “dehydration,” and “watery diarrhea” presented one or several months later than actual norovirus cases in the United States. The detailed information is shown in [Multimedia Appendix 5](#).

Discussion

Principal Findings

In our study, we found that norovirus was correlated with the search terms “gastroenteritis,” “diarrhea,” “vomiting,” “dehydration,” “contagious,” “stomach flu,” “stomach bug,” “restaurant,” and “food poisoning.” In the total related search terms, gastroenteritis, diarrhea, watery diarrhea, fever, vomiting,

and dehydration were typical symptoms of norovirus. Stomach flu, winter vomiting disease, and stomach bug are other names of norovirus. Restaurant and bar were food service settings that occurred with norovirus outbreaks. Chipotle Mexican Grill is a specific restaurant that has experienced norovirus outbreaks in the United States. Apart from the previous search terms, streptococcus and coxsackievirus were related to norovirus. Infections with streptococcus and coxsackievirus have characteristics similar to those of norovirus, such as fever and contagious features.

Infodemiology could be discovered and analyzed in near real time. Many disease outbreaks could be predicted and monitored by infodemiology. The applications of infodemiology included queries analysis with internet search engines, people’s status updates monitor, publications related to public health monitors, etc. Patthi et al [26] used Google Trends from 2004 to 2016 to investigate the global search trends of oral problems. Their method provided an insight to analyze and compare oral disease over time on a mass population. Shimkhada et al [27] used Twitter Chat to identify barriers and responsive policy of patients with metastatic breast cancer care. Klein et al [28] conducted Twitter data to track the spread of COVID-19. Johnson et al [29] performed Google Trends to monitor sexually transmitted infections Chicago. Thus, infodemiology could provide valuable insights into health-related behavior of populations [9].

In this study, we discovered the correlation between norovirus and other search terms in Google search trends. Norovirus is a kind of infectious disease leading to diarrhea and vomiting [30]. Extremely or chronic diarrhea and vomiting can lead to dehydration in some cases of norovirus illness patients [31,32]. There are various transmission methods for norovirus [33]. First, noroviruses are transmitted to healthy people by contaminated food [34,35]. Foods such as oysters or fruits might be contaminated by water containing norovirus particles [36]. Second, noroviruses can be transmitted to healthy people by contaminated water [37]. Last, the norovirus could spread to healthy people by surfaces contaminated by infected people [38]. Therefore, norovirus particles easily infect healthy people via contaminated food, water, and surfaces. Because an effective vaccine to prevent norovirus has not been discovered, it is important for healthy people to prevent norovirus infection by washing hands thoroughly [39]. The transmission of norovirus could be effectively decreased by practicing proper hand

hygiene, preparing food safely, sterilizing surfaces, and washing laundry thoroughly. In this study, we used reasonable prior knowledge of norovirus-related risks to define our list of search terms. In order to cover a wider range of norovirus-related search, we adopted a loose criteria to choose the terms used in this paper that can be used to predict norovirus outbreak.

This study discovered the temporal correlation between norovirus and other terms. Dehydration would be caused by chronic diarrhea or vomiting in older adults or young children. Thus, the search terms including “dehydration” were presented several months later than norovirus in Google Trends. Interestingly, the search term “contagious” was presented earlier than “norovirus” in Google Trends. This indicated that people might contact some contaminated water or food. In addition, people might contact directly with other people who are infected with norovirus. Therefore, they search the term “contagious” before “norovirus.” Apart from typical symptoms, the stepwise regression model showed that internet searches for “cruise” and “wedding” accounted for predictors in the norovirus search. Thus, the internet search terms of some activities might be useful indicators of norovirus outbreaks. In these cases, internet search might occur several months earlier than actual activity for the purpose of reservation and planning ahead of these activities.

Apart from norovirus Google Trends, we also used the data of actual norovirus cases to validate the conclusion. In New York, the cross-correlation analysis of actual norovirus cases and internet search terms showed that the results were similar to norovirus in Google Trends. The Google Trends results of “travel,” “party,” and “wedding” presented one or several months earlier than actual norovirus cases. This result indicated that the search terms “party” and “travel” might be factors of norovirus outbreaks. “Party” and “travel” would lead to people gathering and cause norovirus transmission. This search strategy could be a useful method to predict and monitor outbreaks of norovirus in New York and the United States. Later, we also evaluated the temporal correlation between actual norovirus

cases in California with internet search terms in California. However, the results of California were different from those of New York and the United States. The geographic area of California was much bigger than New York. The California illnesses in every norovirus outbreak might be less than New York. It could explain the reason for the difference between New York and California.

Limitations

There are some limitations in this work. First, we tried to discover the correlation between internet search terms and norovirus infection in New York, California, and the United States. Other countries in the world were not included in this study. The results of the United States with Google Trends might not be applied to other countries worldwide. Second, the internet search engine Google was used in this study, other internet search engines and social networking services such as Bing, Twitter, and Facebook were not included in this study. In the future, the social networking services could be included to evaluate outbreaks of norovirus. Last, the norovirus data and Google Trends data were available monthly. Thus, we have not calculated the correlation between daily data on norovirus infections or hospitalizations with Google Trends data in this study. In the future, the correlation between daily data on norovirus infections or hospitalizations with Google Trends data could be explored if the daily data are available.

Conclusions

In this study, we used Google search trends to investigate the correlation between internet searches and norovirus. The data were downloaded from Google Trends and NORS at the CDC. We used cross-correlation analysis to discover the temporal correlation between norovirus and other terms. We found that search trend data from Google is useful to predict norovirus outbreak. Our study provides a novel strategy based on internet search data to investigate the infodemiology of norovirus and monitor the outbreak of norovirus in the future.

Authors' Contributions

KY and GH performed statistical analysis and drafted the manuscript. LW, TW, WL, and HJ contributed to database building. ACY conceived the design of the study and revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The average CSV about the search terms in Google Trends.

[\[DOCX File, 18 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Cross-correlation analysis of norovirus and Internet search terms–New York.

[\[DOCX File, 25 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Cross-correlation analysis of actual norovirus cases and Internet search terms–New York.

[\[DOCX File, 22 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Cross-correlation analysis of actual norovirus cases and Internet search terms—California.

[\[DOCX File, 25 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Cross-correlation analysis of actual norovirus cases and Internet search terms—US.

[\[DOCX File, 28 KB-Multimedia Appendix 5\]](#)

References

1. Robilotti E, Deresinski S, Pinsky BA. Norovirus. *Clin Microbiol Rev* 2015 Jan;28(1):134-164 [[FREE Full text](#)] [doi: [10.1128/CMR.00075-14](https://doi.org/10.1128/CMR.00075-14)] [Medline: [25567225](https://pubmed.ncbi.nlm.nih.gov/25567225/)]
2. Ahmed SM, Hall AJ, Robinson AE, Verhoef L, Premkumar P, Parashar UD, et al. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis* 2014 Aug;14(8):725-730 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(14\)70767-4](https://doi.org/10.1016/S1473-3099(14)70767-4)] [Medline: [24981041](https://pubmed.ncbi.nlm.nih.gov/24981041/)]
3. Isakbaeva ET, Widdowson M, Beard RS, Bulens SN, Mullins J, Monroe SS, et al. Norovirus transmission on cruise ship. *Emerg Infect Dis* 2005 Jan;11(1):154-158 [[FREE Full text](#)] [doi: [10.3201/eid1101.040434](https://doi.org/10.3201/eid1101.040434)] [Medline: [15705344](https://pubmed.ncbi.nlm.nih.gov/15705344/)]
4. Karst SM, Tibbetts SA. Recent advances in understanding norovirus pathogenesis. *J Med Virol* 2016 Nov;88(11):1837-1843 [[FREE Full text](#)] [doi: [10.1002/jmv.24559](https://doi.org/10.1002/jmv.24559)] [Medline: [27110852](https://pubmed.ncbi.nlm.nih.gov/27110852/)]
5. de Graaf M, van Beek J, Koopmans MPG. Human norovirus transmission and evolution in a changing world. *Nat Rev Microbiol* 2016 Jul;14(7):421-433. [doi: [10.1038/nrmicro.2016.48](https://doi.org/10.1038/nrmicro.2016.48)] [Medline: [27211790](https://pubmed.ncbi.nlm.nih.gov/27211790/)]
6. Switaj TL, Winter KJ, Christensen SR. Diagnosis and management of foodborne illness. *Am Fam Physician* 2015 Sep 01;92(5):358-365 [[FREE Full text](#)] [Medline: [26371569](https://pubmed.ncbi.nlm.nih.gov/26371569/)]
7. Thorne L, Goodfellow I. Norovirus gene expression and replication. *J Gen Virol* 2014 Feb;95(Pt 2):278-291. [doi: [10.1099/vir.0.059634-0](https://doi.org/10.1099/vir.0.059634-0)] [Medline: [24243731](https://pubmed.ncbi.nlm.nih.gov/24243731/)]
8. Tack DM, Ray L, Griffin PM, Cieslak PR, Dunn J, Rissman T, et al. Preliminary incidence and trends of infections with pathogens transmitted commonly through food - foodborne diseases active surveillance network, 10 U.S. Sites, 2016-2019. *MMWR Morb Mortal Wkly Rep* 2020 May 01;69(17):509-514 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6917a1](https://doi.org/10.15585/mmwr.mm6917a1)] [Medline: [32352955](https://pubmed.ncbi.nlm.nih.gov/32352955/)]
9. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009 Mar 27;11(1):e11 [[FREE Full text](#)] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
10. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](https://pubmed.ncbi.nlm.nih.gov/21521589/)]
11. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc* 2006:244-248 [[FREE Full text](#)] [Medline: [17238340](https://pubmed.ncbi.nlm.nih.gov/17238340/)]
12. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res* 2013 Jul 18;15(7):e147 [[FREE Full text](#)] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
13. Mavragani A. Infodemiology and infoveillance: scoping review. *J Med Internet Res* 2020 Apr 28;22(4):e16206 [[FREE Full text](#)] [doi: [10.2196/16206](https://doi.org/10.2196/16206)] [Medline: [32310818](https://pubmed.ncbi.nlm.nih.gov/32310818/)]
14. Yang AC, Tsai S, Huang NE, Peng C. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004-2009. *J Affect Disord* 2011 Jul;132(1-2):179-184. [doi: [10.1016/j.jad.2011.01.019](https://doi.org/10.1016/j.jad.2011.01.019)] [Medline: [21371755](https://pubmed.ncbi.nlm.nih.gov/21371755/)]
15. Yang AC, Huang NE, Peng C, Tsai S. Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. *PLoS One* 2010 Oct 28;5(10):e13728 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0013728](https://doi.org/10.1371/journal.pone.0013728)] [Medline: [21060851](https://pubmed.ncbi.nlm.nih.gov/21060851/)]
16. Yang AC, Tsai S, Yang C, Shia B, Fuh J, Wang S, et al. Suicide and media reporting: a longitudinal and spatial analysis. *Soc Psychiatry Psychiatr Epidemiol* 2013 Mar 10;48(3):427-435. [doi: [10.1007/s00127-012-0562-1](https://doi.org/10.1007/s00127-012-0562-1)] [Medline: [22878833](https://pubmed.ncbi.nlm.nih.gov/22878833/)]
17. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014. [doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634)] [Medline: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)]
18. Seifter A, Schwarzwalder A, Geis K, Aucott J. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospat Health* 2010 May 01;4(2):135-137. [doi: [10.4081/gh.2010.195](https://doi.org/10.4081/gh.2010.195)] [Medline: [20503183](https://pubmed.ncbi.nlm.nih.gov/20503183/)]
19. Mavragani A, Ochoa G, Tsagarakis KP. Assessing the methods, tools, and statistical approaches in Google Trends research: systematic review. *J Med Internet Res* 2018 Nov 06;20(11):e270 [[FREE Full text](#)] [doi: [10.2196/jmir.9366](https://doi.org/10.2196/jmir.9366)] [Medline: [30401664](https://pubmed.ncbi.nlm.nih.gov/30401664/)]
20. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of google trends in health care research: a systematic review. *PLoS One* 2014 Oct 22;9(10):e109583 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0109583](https://doi.org/10.1371/journal.pone.0109583)] [Medline: [25337815](https://pubmed.ncbi.nlm.nih.gov/25337815/)]
21. Google Trends. URL: <https://trends.google.com/> [accessed 2021-07-26]

22. Mavragani A, Ochoa G. Google Trends in Infodemiology and infoveillance: methodology framework. *JMIR Public Health Surveill* 2019 May 29;5(2):e13439 [FREE Full text] [doi: [10.2196/13439](https://doi.org/10.2196/13439)] [Medline: [31144671](https://pubmed.ncbi.nlm.nih.gov/31144671/)]
23. FAQ about Google Trends data. URL: <https://support.google.com/trends/answer/4365533?hl=en> [accessed 2021-07-26]
24. Hall AJ, Wikswo ME, Manikonda K, Roberts VA, Yoder JS, Gould LH. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Dis* 2013 Aug;19(8):1305-1309 [FREE Full text] [doi: [10.3201/eid1908.130482](https://doi.org/10.3201/eid1908.130482)] [Medline: [23876187](https://pubmed.ncbi.nlm.nih.gov/23876187/)]
25. NORS Dashboard. URL: <https://wwwn.cdc.gov/norsdashboard/> [accessed 2021-07-26]
26. Patthi B, Kumar J, Singla A, Gupta R, Prasad M, Ali I, et al. Global search trends of oral problems using Google Trends from 2004 to 2016: an exploratory analysis. *J Clin Diagn Res* 2017 Sep;11(9):ZC12-ZC16 [FREE Full text] [doi: [10.7860/JCDR/2017/26658.10564](https://doi.org/10.7860/JCDR/2017/26658.10564)] [Medline: [29207825](https://pubmed.ncbi.nlm.nih.gov/29207825/)]
27. Shimkhada R, Attai D, Scheitler A, Babey S, Glenn B, Ponce N. Using a Twitter chat to rapidly identify barriers and policy solutions for metastatic breast cancer care: qualitative study. *JMIR Public Health Surveill* 2021 Jan 15;7(1):e23178 [FREE Full text] [doi: [10.2196/23178](https://doi.org/10.2196/23178)] [Medline: [33315017](https://pubmed.ncbi.nlm.nih.gov/33315017/)]
28. Klein AZ, Magge A, O'Connor K, Flores Amaro JI, Weissenbacher D, Gonzalez Hernandez G. Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set. *J Med Internet Res* 2021 Jan 22;23(1):e25314 [FREE Full text] [doi: [10.2196/25314](https://doi.org/10.2196/25314)] [Medline: [33449904](https://pubmed.ncbi.nlm.nih.gov/33449904/)]
29. Johnson AK, Bhaumik R, Tabidze I, Mehta SD. Nowcasting sexually transmitted infections in Chicago: predictive modeling and evaluation study using Google Trends. *JMIR Public Health Surveill* 2020 Nov 05;6(4):e20588 [FREE Full text] [doi: [10.2196/20588](https://doi.org/10.2196/20588)] [Medline: [33151162](https://pubmed.ncbi.nlm.nih.gov/33151162/)]
30. Romero C, Tinoco Y, Loli S, Razuri H, Soto G, Silva M, et al. Incidence of norovirus-associated diarrhea and vomiting disease among children and adults in a community cohort in the Peruvian Amazon basin. *Clin Infect Dis* 2017 Sep 01;65(5):833-839. [doi: [10.1093/cid/cix423](https://doi.org/10.1093/cid/cix423)] [Medline: [29017284](https://pubmed.ncbi.nlm.nih.gov/29017284/)]
31. Godoy P, Alsedà M, Bartolomé R, Clavería D, Módol I, Bach P, et al. Norovirus gastroenteritis outbreak transmitted by food and vomit in a high school. *Epidemiol Infect* 2016 Jul;144(9):1951-1958. [doi: [10.1017/S0950268815003283](https://doi.org/10.1017/S0950268815003283)] [Medline: [26759924](https://pubmed.ncbi.nlm.nih.gov/26759924/)]
32. Shah MP, Hall AJ. Norovirus illnesses in children and adolescents. *Infect Dis Clin North Am* 2018 Mar;32(1):103-118 [FREE Full text] [doi: [10.1016/j.idc.2017.11.004](https://doi.org/10.1016/j.idc.2017.11.004)] [Medline: [29406972](https://pubmed.ncbi.nlm.nih.gov/29406972/)]
33. Gaythorpe KAM, Trotter CL, Lopman B, Steele M, Conlan AJK. Norovirus transmission dynamics: a modelling review. *Epidemiol Infect* 2018 Jan;146(2):147-158 [FREE Full text] [doi: [10.1017/S0950268817002692](https://doi.org/10.1017/S0950268817002692)] [Medline: [29268812](https://pubmed.ncbi.nlm.nih.gov/29268812/)]
34. Hardstaff JL, Clough HE, Lutje V, McIntyre KM, Harris JP, Garner P, et al. Foodborne and food-handler norovirus outbreaks: a systematic review. *Foodborne Pathog Dis* 2018 Oct;15(10):589-597 [FREE Full text] [doi: [10.1089/fpd.2018.2452](https://doi.org/10.1089/fpd.2018.2452)] [Medline: [30109958](https://pubmed.ncbi.nlm.nih.gov/30109958/)]
35. Centers for Disease Control and Prevention (CDC). Norovirus activity—United States, 2002. *MMWR Morb Mortal Wkly Rep* 2003 Jan 24;52(3):41-45 [FREE Full text] [Medline: [12570319](https://pubmed.ncbi.nlm.nih.gov/12570319/)]
36. Wang J, Deng Z. Modeling and prediction of oyster norovirus outbreaks along Gulf of Mexico Coast. *Environ Health Perspect* 2016 May;124(5):627-633 [FREE Full text] [doi: [10.1289/ehp.1509764](https://doi.org/10.1289/ehp.1509764)] [Medline: [26528621](https://pubmed.ncbi.nlm.nih.gov/26528621/)]
37. El-Senousy WM, Costafreda MI, Pintó RM, Bosch A. Method validation for norovirus detection in naturally contaminated irrigation water and fresh produce. *Int J Food Microbiol* 2013 Oct 01;167(1):74-79. [doi: [10.1016/j.ijfoodmicro.2013.06.023](https://doi.org/10.1016/j.ijfoodmicro.2013.06.023)] [Medline: [23866911](https://pubmed.ncbi.nlm.nih.gov/23866911/)]
38. Repp K, Hostetler T, Keene W. A norovirus outbreak related to contaminated surfaces. *J Infect Dis* 2013 Jul 15;208(2):295-298. [doi: [10.1093/infdis/jit148](https://doi.org/10.1093/infdis/jit148)] [Medline: [23559462](https://pubmed.ncbi.nlm.nih.gov/23559462/)]
39. Assab R, Temime L. The role of hand hygiene in controlling norovirus spread in nursing homes. *BMC Infect Dis* 2016 Aug 09;16(1):395 [FREE Full text] [doi: [10.1186/s12879-016-1702-0](https://doi.org/10.1186/s12879-016-1702-0)] [Medline: [27507065](https://pubmed.ncbi.nlm.nih.gov/27507065/)]

Abbreviations

CDC: Centers for Disease Control and Prevention

NORS: National Outbreak Reporting System

Edited by R Kukafka; submitted 24.09.20; peer-reviewed by A Mavragani, A Sudaryanto, JK Kumar; comments to author 21.12.20; revised version received 26.02.21; accepted 05.07.21; published 29.09.21

Please cite as:

Yuan K, Huang G, Wang L, Wang T, Liu W, Jiang H, Yang AC

Predicting Norovirus in the United States Using Google Trends: Infodemiology Study

J Med Internet Res 2021;23(9):e24554

URL: <https://www.jmir.org/2021/9/e24554>

doi: [10.2196/24554](https://doi.org/10.2196/24554)

PMID:

©Kai Yuan, Guangrui Huang, Lepeng Wang, Ting Wang, Wenbin Liu, Haixu Jiang, Albert C Yang. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 29.09.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.