Google Trends for Formulating GIS Mapping of Disease Outbreaks in India

Bhattacharya, I.,[1] Ramachandran, A.,[1] Bhattacharya, J.[2] and Dogra, N. K.[1]
[1]International Institute of Health Management Research, Plot No 3, Sector 18A, Dwarka
New Delhi 110075, India, Phone: +91-11-30418900, Fax +91-11-30418909
E-mail: indrajitbhat@gmail.com, anandhirama@gmail.com, nitish.dogra@gmail.com
[2]Indian Institute of Technology, New Delhi, India, E-mail: jaijit@gmail.com

Abstract
Use of online sources to track disease outbreaks and deliver real-time surveillance in emerging public health threats is becoming the need of the hour as resurgence of diseases challenge global health systems. Recent advances in Geographical Information System (GIS) and mapping technologies have enhanced diseases surveillance programs. The current paradigm shifts are moving towards greater adoption of the ubiquitous internet for GIS deployment. This paper attempts to highlight the applicability of three online disease mapping tools – Google trends (GT), HealthMap and Kazemill as newer paradigms of disease surveillance in India. The study has been limited to comparing the trends obtained from the online tools with freely available public health data for few diseases such as malaria, typhoid and cholera that commonly occur in India. We conclude that GT could possibly be used for providing first hand information of impending disease outbreaks subject to further validation against past trends of the disease to have more relevance.

1. Introduction

Increasing global disease burden in resource poor developing countries is forcing public health professionals and providers to urgently seek viable solutions for disease management and surveillance. Other than traditional virological or clinical sentinel surveillance,(Beatty et al, 2010) other approaches such as monitoring telephone triage calls (Yih et al., 2009) and school/work absenteeism (Besculides et al., 2005) provide a near real time information that can be used for predicting disease outbreaks. Increasingly the, use of online web searches for disease surveillance has been on the rise (Brownstein et al, 2011, Eisen et al., 2008 and Kamadjeu and Tolentino, 2006). As internet penetration increases drastically it has not only been seen as a successful source of health information for physicians and patients (Dickerson et al., 2004), but also as a tool by the general public to enquire about health challenges. Using internet for tracking and mapping diseases, decision-making and shaping policies are some of the key components of effective public health work. *"Infodemiology"* the study of the distribution and determinants of information on the Internet (Eysenbach, 2002), was the term introduced from information and epidemiology to track health demand and supply trends. *Infodemiology can be defined as "the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy"*. Potential infodemiology indicators and metrics include automatically aggregated and analyzed data on the prevalence and patterns of information on websites and social media; metrics on the "chatter" in discussion groups, blogs, and microblogs (eg, Twitter); and activities on search engines, etc., (Eysenbach, 2009). The large number of health related searches conducted through popular websites creates trend data which can be analyzed over time to identify disease outbreaks and supplement traditional surveillance methods (Eysenbach, 2006, Ginsberg et al., 2009 and Wilson and Brownstein, 2009). Google, a mainstream general search engine which also includes several additional tools and resources, represents an important gateway to online health information for patients and physicians (Wheeler, 2006 and Laurent and Vickers, 2009). Trend data generated by the number of Google searches over time in a particular geographic region have recently been made available by Google Trends. Google has launched experimental tools for near real-time detection of influenza and dengue outbreaks by monitoring and analyzing health care–seeking behavior in the form of queries to its online Google search engine. These tools mine the Google search query data to estimate the disease activity in near real-time. It was developed by matching trends in queries for flu and dengue related search terms to seasonal trends in the Centers for Disease Control and Prevention's (CDC) data for sentinel physician visits. Real-time surveillance would alert public health care practitioners in the early phases of an outbreak, enabling them to promptly institute control measures and case finding and to ensure adequate access to treatment, thereby reducing morbidity and mortality (Google trends, 2009; Johnson et al 2004). With international concerns about emerging infectious diseases, bioterrorism, and pandemics, the need for a real-time surveillance system is at an all-time high (Longini et al., 2005, Mandl et al., 2004 and NEDSS, 2001).The data generated would be useful for public health care practice, clinical decision making, and research (New York Times, 2008). In this paper, we describe the applicability of the more generic Google Trends (GT) methodology to disease surveillance. A brief overview of GT, how the data are processed, potential uses, and the tool's strengths and limitations are also highlighted. We provide initial results for few diseases that are of high priority in India and have compared them with other online mapping tools HealthMap and Kazemill to assess whether web search query data is a viable data source for the early detection and monitoring of disease epidemics in Indian scenario.

2. Methods

The web search tools utilized in the current paper were Google trends, Google Flu Trends, Google Dengue Trends, HealthMap and Kazemill. Surveillance data available in the public domain for the diseases specific to India were reviewed for comparison purposes.

2. 1 Google Trends (www.google.com/trends/)

We searched trends based on a list of keywords related to the infectious diseases particular to Hepatitis, Dengue, Chickenpox, Tuberculosis, Diarrhea, Cholera, H1N1 Flu, Typhoid and Malaria since these are relevant in the Indian context. The search can be carried out in global context for all countries or can be mapped on desired geographic location to provide trend data. The search provides trends from 2004 onwards but the graphs can be manipulated to restrict results to specific time frames.

2.1.1 Methodology adopted by google trends

The trend is developed by analyzing a fraction of the total web searches in the Google search engine over a period of time for the term being searched (say "tuberculosis") and extrapolated to estimate the search volume relative to the total number of searches done on Google over time. Trends uses data aggregated over millions of users without personally identifiable information. Additionally, it only shows results for search terms that receive a significant amount of traffic, and enforces minimum thresholds for inclusion in the tool. This information is displayed in a search volume index graph, which is updated daily. The bottom part of the graph is the news reference volume graph that indicates the frequency with which the web search queries appeared in Google News stories. The data used to generate these graphs are scaled to the average search traffic for the selected term and are also normalized on a relative, rather than absolute, basis for example, queries for "Tuberculosis" (Figure 1).

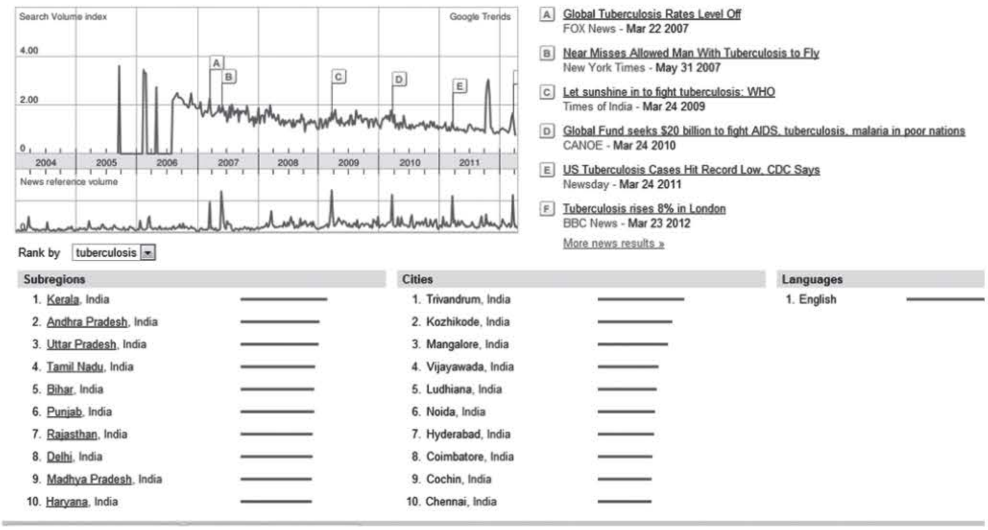


Figure 1: Google trends on Tuberculosis (Generated on 10/08/2012)

When a spike in the frequency is detected in the news reference volume graph, GT labels the graph with a headline of a relevant but randomly selected Google News story published near the time of the spike (Carneiro and Mylonakis, 2009). These headlines are shown to the right of the search volume index graph. Currently only English Language headings are considered. The regions, cities, and languages with the highest search volume are displayed on the bottom of the page.

2.1.2 Google flu trends / Dengue trends
(www.google.org/ flutrends & www.google.org/ denguetrends)

The tool, Google Flu Trends (GFT), is a sophisticated Web-based tool for detection of regional outbreaks of influenza by monitoring and analyzing health care–seeking behavior in the form of queries to its online search engine. This is based on the principle that a close relationship exists between the number of people searching for flu related terms and those who are actually having flu symptoms. Preliminary testing carried out by Centers for Disease Control and Prevention (CDC) suggests that GFT can detect regional outbreaks of influenza 7–10 days before conventional CDC surveillance (Frenk and Dantes, 2002). According to Lynette Brammer, an epidemiologist with CDC's influenza division, Google Flu Trends data supposed to be real-time measure of flu based on online searches correlates well with the CDC's trends (Amednews, 2013). The CDC uses laboratory and clinical data to publish national and regional weekly statistics, typically with a 1–2 week lag after it gets reports on flu cases. . GFT is based on the assumption that there exists high correlation between the relative frequency of certain queries in Google web search with the percentage of physician visits in which a patient presents with influenza-like symptoms leading to accurate estimation of influenza like symptoms at regional and state levels (Mohebbi et al., 2009 and Ferguson et al., 2005). Even though not every person who searches for "flu" is actually sick, but a pattern emerges when all the flu-related search queries are added together. Google Flu allows users to customize reported activity by region, state or large city. While GT compares world's interest in chosen topics over time and geographical location, GFT correlates with retrospective surveillance data from the CDC. Google Flu Trends has been validated by a John Hopkins Study (IHT[2], 2013) related to surges in flu-related emergency room visits a week before warnings came from the CDC (Figure 2). For India, the flu activity is not available in the Google Flu trends. In the same vein, Google Dengue Trends an extension of the flu trend show historical query-based dengue estimates for different countries and regions compared against official dengue surveillance data (Figure 3). Web search query data were found to be capable of tracking dengue activity in Bolivia, Brazil, India, Indonesia and Singapore (Chan et al 2010). Currently the trends are shown for 10 countries including India. The source of official case count for India is the WHO report on dengue cases (WHO, 2006).

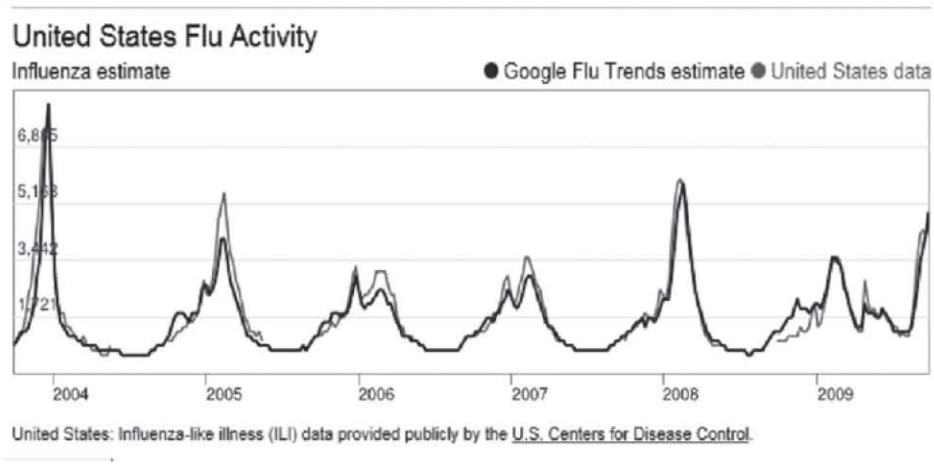


Figure 2: Google Flu Trends -to predict surges in flu-related emergency room visits a week before warnings came from the CDC (Johns Hopkins Study) (Generated on 26/04/2013)

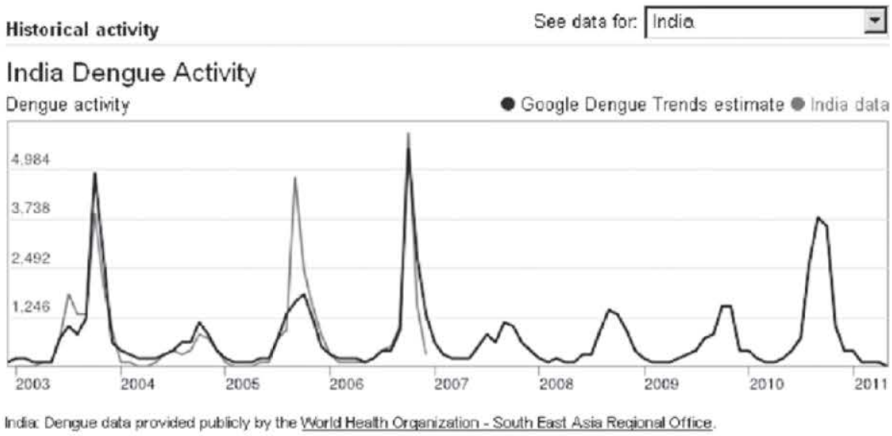


Figure 3: Google Dengue trends (Generated on 10/08/2012)

2.3 HealthMap (http://www.HealthMap.org/)

HealthMap, is a real time, automated electronic information system for monitoring, organizing, and visualizing reports of global disease outbreaks. It utilizes online informal sources for disease outbreak monitoring and real-time surveillance of emerging public health threats to achieve a unified and comprehensive view of current global state of infectious disease and their effect on human and animal health.

2.3.1 Methodology adopted by HealthMap

It brings together disparate data sources consisting of varied reliability, ranging from online news aggregators (such as Google news), eyewitness reports, expert- curated discussions (such as ProMED) to validated official reports / alerts (such as WHO reports) for providing disease foot prints, thus facilitating early detection of global public health threats (Figure 4). Through an automated text processing system, the data is aggregated by disease and displayed by location for user friendly access to the original alert.

2.4 Kazemill (http://www.mccann.co.jp/award/2011-/stac/index_en.html)

Kazemill is a new social media-powered HealthMap developed for a Japanese pharmaceutical company. The map monitors tweets that mention various cold and flu symptoms and applies that data to a map of the country to show how and where disease is spreading (Figure 5).

2.4.1 Methodology adopted by Kazemill

It is a predictive tool developed on inputs from the social networking site (twitter) that collects tweets on the six most common cold symptoms and displays back to the public on website using a color coded map. Nearly 18,000 daily mentions of specific symptoms – such as throat pain, runny nose or chills – are gathered along with geographic information and overlaid on a map of the country. Colors – such as orange for itchy throat, purple for a cough and red for fever – show viewers of the online map the spread of illnesses and the prevalence in particular areas. The recent version of the tool, Kazemill plus, combines historical information on cold symptoms and helps in providing predictive information of how the cold trend for the forth coming week would behave.

2.5 Secondary Data

To check whether GT can be used a viable tool for disease surveillance we mapped the trends obtained using it to the trend obtained from the government disease surveillance programs and annual health reports. The data obtained from the latter were either incidence data or mortality trends. The spikes in the GT search trend volumes were compared to the disease incidence/mortality rates obtained from the publically available reports to see if pattern provided by both could be compared with each other. The total search volume calculated using GT for a given year and the yearly data from the reports were plotted in scatter plots. Data from the National Vector Borne Disease Control Program (NVBDCP) reports were considered for malaria (www. nvbdcp.gov.in).

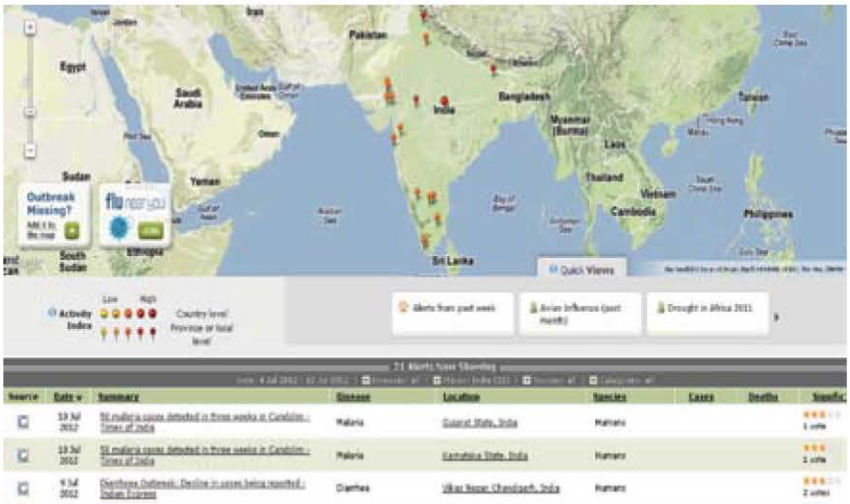


Figure 4: HealthMap footprint on India (accessed on 10th July 2012)

Figure 5: Cold Map displayed by Kazemill

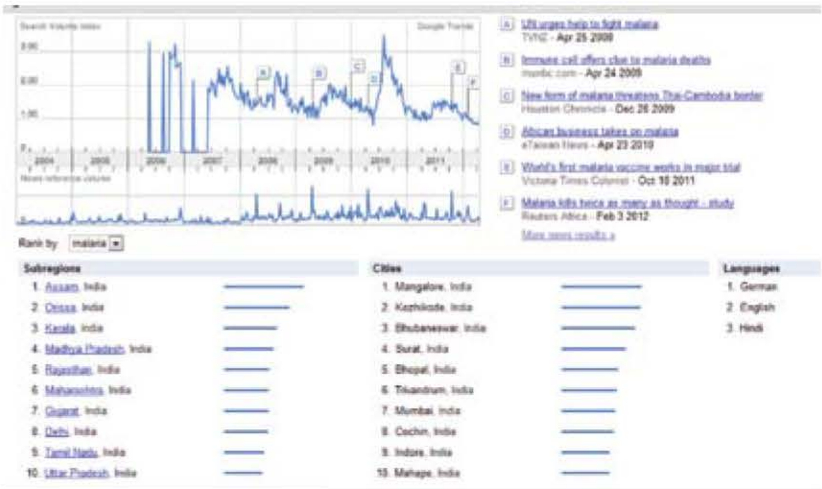


Figure 6: Google trends on Malaria

3. Results and Discussions

The GT trends for the diseases considered in the study are illustrated in Figures 6 – 11. Earlier studies relating GFT and GT with sentinel physician networks (SPNs) have shown that GFT and GT models are robust and could help in epidemiological surveillance by providing more rapid estimates of disease incidence even before conventional methods of detection (Valdivia and Lopez-Alcalde, 2010, Valdivia and Corella, 2010 and Pelat et al., 2009). GT is in the early phases of development, and its data "may contain inaccuracies for a number of reasons, including data sampling issues and approximation methods used to compute the results" (New York Times, 2008). Currently, GT search criteria are not standardized. Users may enter symptoms differently, depending on their level of education and their cultural and language backgrounds. For example, users may enter fever, pyrexia, chills, and rigors for the same symptom. Detailed analyses are required to find search query proxies that correlate well with specific diseases. In a recent study in South China, it has been shown that Google Trends data using Chinese search terms were generally well correlated with conventional methods of surveillance (Kang et al., 2013).

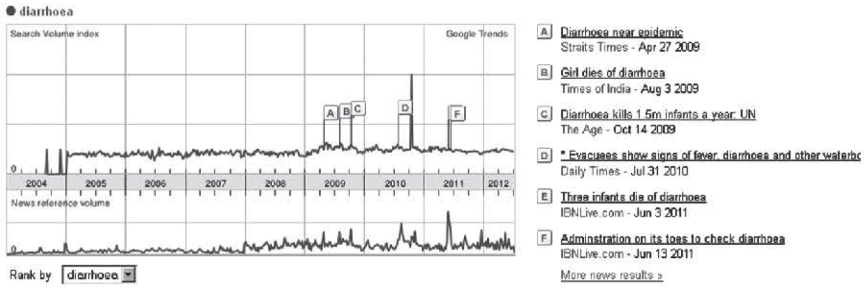


Figure 7: Google trends on Diarrhoea

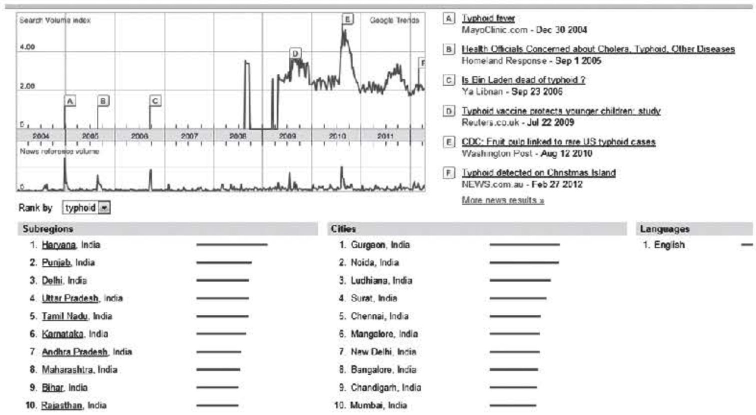


Figure 8: Google trends on Typhoid (Enteric Fever)

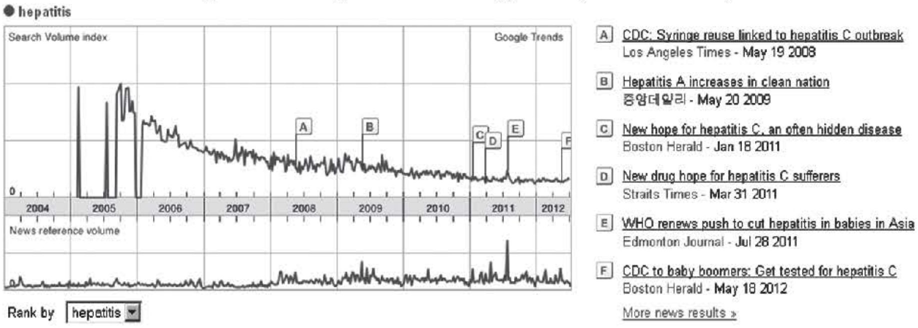


Figure 9: Google trends on Hepatitis

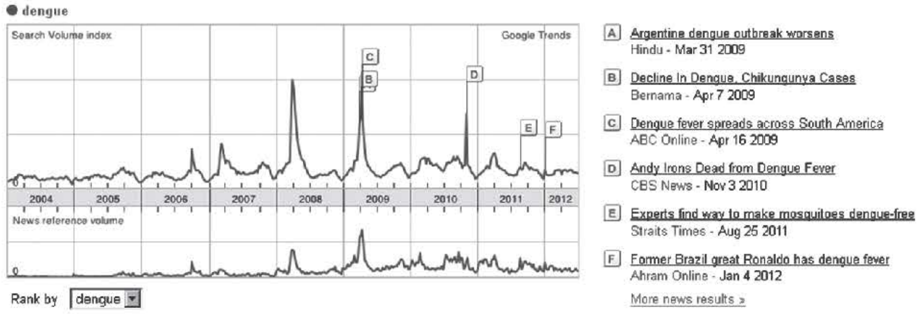


Figure 10: Google trends on Dengue & Chickenpox

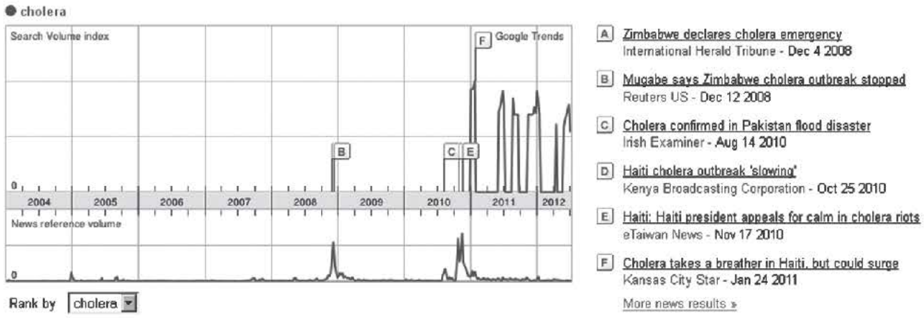
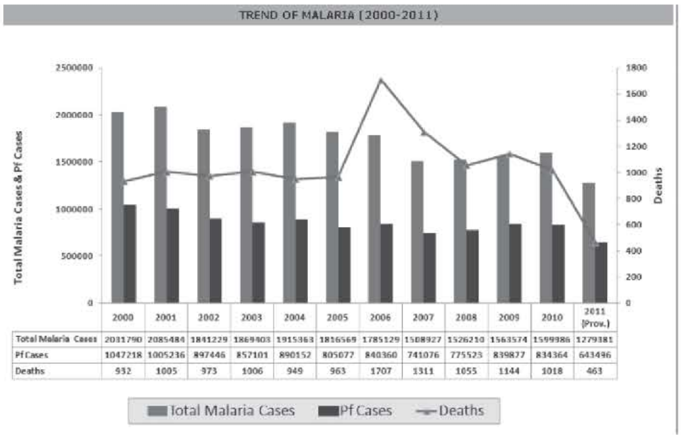


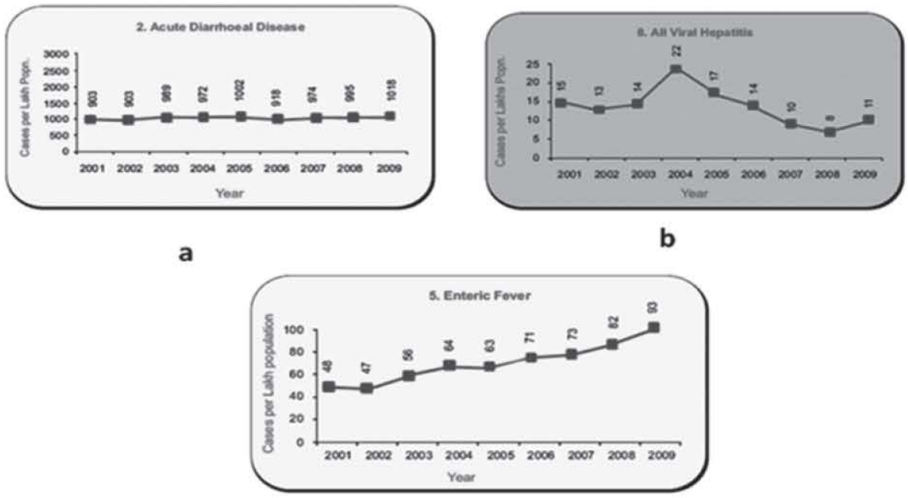
Figure 11: Google Trends for Cholera & H1N1



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011 (Prov.)
Total Malaria Cases	2031790	2085484	1841229	1869403	1915363	1816569	1785129	1508927	1526210	1563574	1599986	1279381
Pf Cases	1047218	1005236	897446	857101	890152	805077	840360	741076	775523	839827	834364	643496
Deaths	932	1005	973	1006	949	963	1707	1311	1055	1144	1018	463

Figure 12: Trend of Malaria (2000-2011) from NVBDCP

However limitations were reported due to limited proportion of common terms used in the web searches which were insufficient to be included in the Google Trends and limited identification of exact search trends generated due to true cases. To be most effective, GT requires large populations of Web search users, which means that GT is currently better suited to tracking disease activity in developed countries. On comparing the malaria trend obtained from GT with the data predicted by NVBDCP (Figure 12), the spikes predicted by GT were found to be supported with the similar such trend in the data. Since GT provides only trend as predicted by the relative search volume and not the exact number of occurrences, the comparison was done to determine if the spikes in GT suggest increased incidence of the disease. Similarly the trend in the search queries for diarrhea, enteric fever (Typhoid) and viral hepatitis were mapped with the pattern in the number of death cases reported in the National Health Profile (Figures 13). Even though the GT trend cannot be directly correlated to the mortality data, the comparison is based on the assumption that the search queries will be greater if there are greater incidence and mortality events due to the disease.



a

b

c

Figure 13: Mortality Trends of Diarrhoea, Hepatitis &Typhoid from National Health Profile

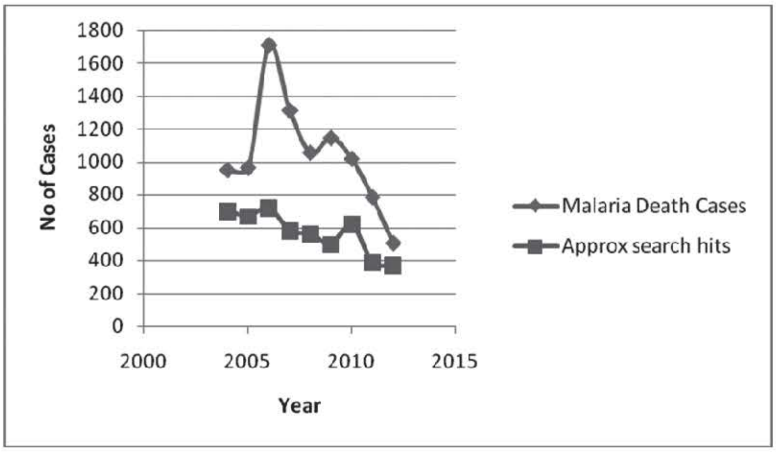


Figure 14: Mapping of spikes from Google Trends with Annual Health Data

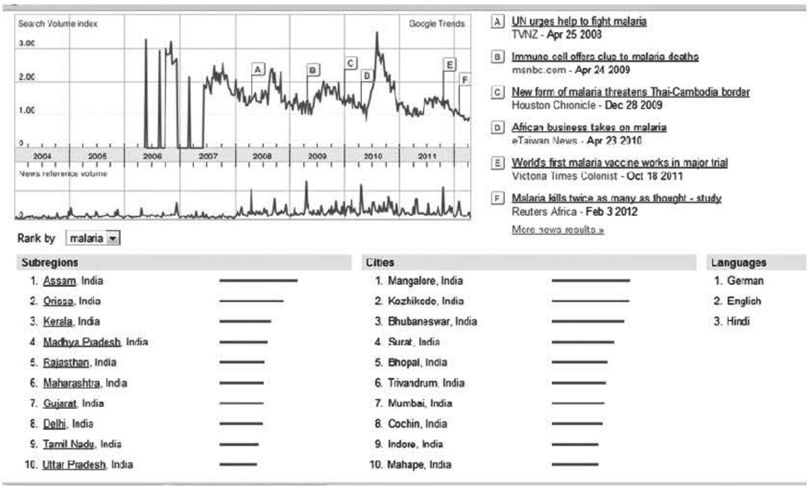


Figure 15: HealthMap foot print of South India

Wherever the trends showed appreciable fluctuations in GT we found a corresponding increase in number of incidence/ death cases. Figure 14 illustrates the comparison made between the malaria deaths as reported in the reports and the search volume pattern obtained from GT. All this goes to suggest that GT can be used as a preventive tool to gather prior information of occurrence and spread of infectious diseases. It is a convenient, easy, and accessible source of search data. In its current stage of development, GT can be used in conjunction with traditional surveillance systems to improve the efficacy of disease surveillance. Research is needed to find suitable Web search query proxies that correlate well to actual cases of diseases of interest. These proxies then can be used to establish specialized tools for infectious diseases, using Google Flu Trends as a blueprint. On the other hand the limitations are that GT is available only in a limited number of languages and region-specific versions. GT is an exciting tool with enormous potential, as shown by Google Flu Trends Additionally, the number of internet users in India is also rising dramatically along with the socio-economic profile of users (Sharma and Thoppil, 2011 and I-Cube Report, 2011). These two factors must also be adjusted while commenting on trends. Also, there is privacy issues involved in using Google web search data. There are news that Google stores and uses data from personal web searches for public research, often without the consent and knowledge of Internet users (Dover, 2008). In some cases, Google search data may be traced to individuals if they are signed into their accounts when they conduct online searches even though Google provides assurance to the users that personal search data will remain safe and private. HealthMap's data source is based on the freely accessible online information from ProMED Mail, World Health Organization, GeoSentinel, EuroSurveillance, Google News and the like. The HealthMap footprint on South India showed Swine flu H1N1 reported in parts of Andhra Pradesh, Karnataka and TamilNadu ; cholera reported in Bangalore and dengue reported in Madurai ; Hepatitis A,B,E, Typhoid reported in Kochi (Figure 15). It uses the technique of data mining and has been successful in monitoring health situations like meningitis, whooping cough, equine herpes virus, a measles outbreak in Utah, and antibiotic resistance worldwide. But if HealthMap is to successfully predict all relevant disease it is necessary for information regarding incidence and prevalence to be validated to reduce chances of misinterpretation. Also there should be a real time disease surveillance program that can accurately report the incidence on real-time to be freely available in public domain which is generally not the case. In India even though Integrated Disease Surveillance Program (IDSP), other numerous government sponsored health programs and the District Level Health Survey are part of the public health system, they are not sufficient enough for providing inputs to tools like HealthMap.

4. Conclusion

Web-based search data in particular can be successfully used to detect various disease outbreaks through spikes in Internet searches which can be an early predictor. Also in the case of an outbreak, how effectively information is disseminated among the population can be tracked through population search behavior during the pandemic. Developing countries like India, having resource constraints, where major population is isolated by geography, either completely lack timely data about impeding infectious and pandemic diseases or have a time lag in obtaining the information. Under such situations, meaningful data can be envisaged using online tools even though it has limitations. It can serve an important purpose in relation to Information, Education and Communication (IEC) activities for the masses, reducing the expense of unnecessary treatment, potentially avoiding fatalities and for taking preventing measures. The disease surveillance systems highlighted in this paper have both advantages and disadvantages (Table 1). All these systems need to be further evaluated thoroughly for the potential to track epidemics and other diseases with high prevalence. GT depends upon the number of users querying about a disease, news reports and on the health seeking behavior like visits to clinics, emergency department and sale of medicines related to drugs. Information and communication patterns in the internet are not only related to queries regarding information seeking, there is an equal chance of the internet searches being burdened due to misreporting. The challenge in utilizing these systems does not lie so much on the availability of the information but on the aggregation and analysis of the data which determines the final trend that is displayed. Depending upon the nature or prevalence of the disease, availability of the internet and information seeking behavior of the users, the web search traffic would differ. Hence it is imperative that metrics and indicators need to be developed that takes into consideration the above.

Table 1: Comparison table of Online Disease Surveillance Systems

Sr. No.	Health Surveillance Tools	Popular since	Accessibility	Advantages	Disadvantages
1	**Google Trends**	2005	Freely accessible	A probable timely, robust, and sensitive surveillance system. It is best used for surveillance of epidemics and diseases with high prevalence's Currently better suited to track disease activity in developed countries where there is large web search traffic relative to health and disease.	Spikes in search volume are currently hard to interpret but have the benefit of increasing vigilance. Herman Anthony Carneiero (2009)
2	**Health Map**	1993	freely accessible since 2006	It has ability to extract useful, customizable messaging and views from a mass of unstructured data.	Still requires many improvements to be made for it to be a truly useful resource for both public health professionals and the general public. Clark C.et.al (2008)
3	**Kazemill**	2012	freely accessible	Effective in predicting cold and flu outbreaks.	Limited testing in Japan for mobilizing over-the-counter products for ailments including cold and flu remedies.

Considering all this, so far GT has more correlation with rates of disease like incidence symptoms than actual prevalence. Initiatives like Google correlates, a new, experimental service designed to connect search analysis with real-life data are being taken (Mohebbi et al., 2011). With the global health perspectives changing and society becoming more dependent on the Internet, a wealth of information that reflects the "collective intelligence" of populations is available making it possible to track infectious disease activity faster by tools like GT than by traditional surveillance systems (Chen et al., 2010, Hulth et al., 2009 and Pelat et al., 2009). Although GFT, GT and other internet resources cannot replace the conventional surveillance methods, these unique and innovative technologies takes us one step closer to true reliable real-time outbreak surveillance that could be used as an early-warning or just-in-time tracking systems by public health professionals and physicians to provide the necessary assistance to patients.

References

About Google Trends, 2009, www.google.com/intl/en/trends/about.html (accessed on 16 July 2009).

Amednews, March 4, 2013 , Early Alarms Sound Online When Illness Go Viral, Posted by Christine S. Moyer, www.amednews.com, accessed on April 23, 2013.

Beatty, M. E., Stone, A., Fitzsimons, D. W., Hanna, J. N. and Lam, S. K., 2010, Best Practices in Dengue Surveillance: A Report from the Asia-Pacific and Americas Dengue Prevention Boards, *PLoS Negl Trop Dis* , 4, 11.

Besculides, M., Heffernan, R., Mostashari, F. and Weiss, D., 2005, Evaluation of School Absenteeism Data for Early Outbreak Detection, New York City, *BMC Public Health*, 5, 105–105.

Carneiro, H. A. and Mylonakis, E., 2009, Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks, *Clinical Infectious Diseases,* 49, 1557–64.

Chan, E. H., Sahai, V., Conrad, C. and Brownstein, J. S., 2011, Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance, *PLoS Negl Trop Dis* , 5(5), e1206.

Dickerson, S., Reinhart, A. M., Feeley, T. H., Bidani, R., Rich, E., Garg, V. K. and Hershey, C. O., 2004, Patient Internet Use for Health Information at Three Urban Primary Care Clinics. *J Am Med Inform Assoc,* 11, 499-504.

Dover, D., 2008, The Evil Side of Google? Exploring Google's User Data Collection, Posted on June 24th 2008, to SEO Blog at http://www.seomoz.org/blog/.

Eysenbach, G., 2006, Infodemiology: Tracking flu-related searches on the Web for Syndromic Surveillance. *AMIA Annu Symp Proc,* 244–8.

Eysenbach, G., 2009, 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*, 11, e1.

Ferguson, N. M., Cummings, D. A. and Cauchemez, S., 2005, Strategies for Containing an Emerging Influenza Pandemic in Southeast Asia, *Nature*, 437, 209–14.

Frenk, J. and Gomez-Dantes, O., 2002, Globalization and the Challenges to Health Systems, *Health Aff (Millwood)*, 21,160–5.

Mohebbi MH, G. J., Patel, R. S., Brammer, L., Smolinski, M. S., 2009, Detecting Influenza Epidemics using Search Engine Query Data, *Nature* 457, 1012–1014.

Hulth, A., Rydevik, G. and Linde, A., 2009, Web Queries as a Source for Syndromic Surveillance. *PLoS One 4*, e4378-e4378.

I-Cube Report, 2011, Internet in India, http://www.iamai.in, accessed on July 12, 2012.

IHT2 Big Data Report, 2013, Transforming Healthcare through Big Data, ihealthtran.com/wordpress/2013/03/iht²-releases-big-data-research-report-download-today/ accessed on April 23, 2013Johnson, H. A., Wagner, M.M., Hogan, W.R., et al., 2004, Analysis of Web access logs for surveillance of influenza, *Stud Health Technol Inform,* 1107, 1202–6.

Kamadjeu, R., and Tolentino, H, 2006, Web-based Public Health Geographic Information Systems for Resources-Constrained Environment using Scalable Vector Graphics Technology: A Proof of Concept Applied to the Expanded Program on Immunization Data, *International Journal of Health Geographics* , 5, 24.

Kang, M., Zhong, H. and Rutherford, S., 2013, Using Google Trends for Influenza Surveillance in South China, *PLOS One*, 8(1).

Longini, I. M. Jr, Nizam, A. and Xu, S., 2005, Containing Pandemic Influenza at the Source, *Science,* 309, 1083–7.

Mandl, K. D., Overhage, J. M. and Wagner, M. M., 2004, Implementing Syndrome Surveillance: A Practical Guide Informed by the Early Experience, *J Am Med Inform Assoc*, 11,141–50.

Mohebbi, M., Vanderkam, D. and Kodysh, J., 2011, Google Correlate Whitepaper, accessed from http://www.google.com/trends/correlate/ on July 10, 2012.

Mohebbi, M., Ginsberg, J., Matthew, H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant, L., 2009, Detecting Influenza Epidemics using Search Engine Query Data, *Nature*, 457, 1012-1014.

National Electronic Disease Surveillance System Working Group, 2001, National Electronic Disease Surveillance System (NEDSS): A Standards Based Approach to Connect Public Health and Clinical Medicin, *J Public Health Manag Pract*, 7, 43–50.

National Vector Borne Disease Control Program, www.nvbdcp.gov.in/ malaria-new,html/ Lastly accessed on April 23, 2013.

New York Times, 2008, Google uses Searches to Track Flu's Spread, available at: http://www.nytimes.com/2008/11/12/technology / internet/12flu.html?rp1. Accessed 19 January 2009.

Pelat, C., Turbelin, Cm., Bar-Hen, A., Flahault, A. and Valleron, A. J., 2009, More diseases tracked by using Google Trends, *Emerg Infect Dis*, 15, 1327-1328.

Sharma, A. and Thoppil, D., 2011, Google Sees India Web Explosion - Internet Use Poised to Balloon but Some Business Habits Die Hard , The Wall Street Journal Technology , Published on September 16, http://online.wsj.com/article/, accessed on July 12, 2012.

Valdivia, A. and Lopez-Alcalde, 2010, Monitoring Influenza Activity in Europe with Google Flu Trens: Comparision with Findings Of Sentinel Physician Networks – Results for 2009-10, Euro.Surveill, 15(29).

Valdivia, A. and Corella, S. M, 2010, Diseases Tracked by using Google Trends, Spain, *Emerg Infect Dis*. 16(1), 168.

Wheeler, D., 2006, Google as a Pathology Portal, *Adv Anat Pathol*, 13, 275-276.

Yih, W. K., Teates, K. S., Abrams, A., Kleinman, K. and Kulldorff, M, 2009,, Telephone Triage Service Data for Detection of Influenza-Like Illness, *PLoS ONE* , 4(4), e5260.