medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Title:

Monitoring COVID-19 related public Interest and population Health Literacy in South Asia: An Internet Search-Interest Based Model

Authors and Institutions

Hasan Symum [1*], Kh Mohammed Ali Sagor [2]

(1) Department of Industrial and Management System Engineering, University of South Florida, Tampa, Florida, USA

(2) Dhaka Medical College Hospital, Dhaka, Bangladesh

* Corresponding Author

Hasan Symum, MS
Graduate Research Assistant, Department of Industrial and Management System Engineering University of South Florida, Email: hsymum@usf.edu

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Background

Information epidemiology based on internet search data can be used to model COVID-19 pandemic progressions and monitor population health literacy. However, the applicability of internet searches to monitor COVID-19 infections and public health awareness in South Asian countries are unclear.

Objectives

To assess the association of public interest and health literacy in COVID-19 with the actual number of infected cases for countries in South Asia.

Methods

Google Trends data from January to March 2020 were used to correlate public interest and health literacy with official data on COVID-19 cases using the relative search volume (RSV) index. Public interest in COVID-19 was retrieved in RSV indices with the search term "Coronavirus (Virus)". Similarly, an OR combination of search terms "hand wash", "face mask", "hand sanitizer", "face shield" and "gloves" were used to retrieve RSV indices as a surrogate of population health literacy in COVID-19. Daily confirmed COVID-19 cases were obtained from the COVID-19 data repository managed by the Johns Hopkins University. Country-level time-lag correlation analyses were performed for a time lag between 30 and +30 days.

Results

COVID-19-related worldwide public interest reached a peak on March 16, 2020, right after the WHO announcement of coronavirus outbreak as a pandemic. COVID-19 related public interest reached the highest peak in South Asian countries a few days after each county reported 100th confirmed cases. There were significant positive correlations between COVID-19 related public interest and daily laboratory confirmed cases in countries expect Nepal, Bhutan, and Sri Lanka. The highest public interest in South Asian Countries was on average 12 days before the local maximum of new confirmed cases. Similarly, web searches related to personal hygiene and COVID-19 preventive measures in south Asia correlated to the number of confirmed cases as well as national restriction measures.

Conclusion

Public interest indicated by RSV indices can help to monitor the progression of an outbreak such as the current COVID-19 pandemic particularly in countries with a lack of diagnostic and surveillance capacity, and thereby distribute appropriate health information to the general public.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic represents an unprecedented global healthcare emergency with more than 20 million laboratory-confirmed cases and 730,000 deaths between February and July 2020.[1] The COVID-19 which is caused by the novel acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been the first time reported in Wuhan, China on December 31, 2019.[2] Since December 2019, the outbreak continues to spread exponentially worldwide due to the high virulence and prevalence of asymptomatic cases.[3,4] By March 2020, COVID-19 has spread globally with more than 400,000 confirmed cases which prompted the World Health Organization (WHO) to declare the COVID-19 as a global pandemic.[5] As of August 10, 2020, COVID-19 has affected more than 227 countries and territories with more than 7 million active cases and the number is still exponentially rising.[1,6] South Asia is one of the leading COVID-19 affected regions with more than 20 million (28.57 % of global cases) confirmed cases as of August 2020. India, Pakistan, and Bangladesh were the most COVID-19 affected countries in the South Asia region with 1695988, 278305, and 237661 confirmed cases, respectively.[6]

Over the last few years, with availability and usage increased worldwide, the internet has become the main source of information particularly for healthcare-related concerns.[7,8] With over 4.5 billion active internet users around the world, millions of people worldwide search online for health-related queries which make Web search queries a valuable source of public health infoveillance.[9,10] Understanding about Web search trends can provide valuable information about public interest and awareness in health emergencies as a proxy for public health risk perception.[8,11] Prior studies used internet search queries to model the outbreak of infectious disease (e.g., dengue, and influenza), track substance usages, and monitor public behavior.[12,13] Subsequently, internet search queries also used to investigate public interest, health awareness, and mental health in the COVID-19 pandemic situation.[14–16] However, most of the studies so far focused on China, the United States, and other European countries. To our knowledge, there is no Infodemiology study till now have explored the association of web queries to the COVID-19 public internet and lawlessness in South Asian countries. With more than 880 million (18.87% of global users) internet users in South Asia, tracking Web search queries can be a real-time health informatics tool to strengthen the public health surveillance in health emergencies such as

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

COVID-19 pandemic.[9,10,17] Therefore, this study explored the potential use of internet search trends for monitoring public interest and preventive health awareness towards COVID-19 infections in South Asia.

METHODS

Daily data on confirmed COVID-19 cases were obtained from the data repository managed by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.[6] Data were retrieved for worldwide total new cases as well as for the following individual countries namely the United States (US), Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka from the 22nd January 2020 to 29th July 2020.

The Google Trends analytics platform was utilized to explore internet user search activities related to COVID-19 pandemic public interest and population health literacy. Google Trends is an online tracking system of internet hit□search volumes, which enables the researchers to study trends and temporal patterns of the popular search queries.[18] Google trend determines the proportion of the searches for user□specified terms among all web queries on the Google Search website and other affiliated Google sites for a given location and time.[19] Google trend then normalized the proportion by the highest query share of that term over the time-series and reports search interest as a unit of relative search volume (RSV) index. The retrieved RSV index values range from 0 to 100, with a value of 50 representing half the public interest as a value of 100.[20] This RSV indices have been used previously used to analyze public interest in various health-related issues as well as passive health surveillance and disease monitoring.[13,14,21] The search term “Coronavirus (Virus)” was used to retrieve Worldwide and nation-specific RSV indices in Google Trend as a representation of public interest on COVID-19 information. Also, a combination of the popular search phrases “hand wash”, “face mask”, “hand sanitizer”, “face shield” and “gloves” were used to retrieve RSV indices as a surrogate of public interest on the practice of personal hygiene and other COVID-19 preventive measures.

Changes in temporal trend of public interest and number of COVID-19 cases were analyzed graphically for nation-specific major events and policy initiatives (e.g., first coronavirus cases,1000 cumulative deaths, nationwide lockdown). Temporal association in public interest in COVDI-19 and the number of new confirmed cases were analyzed using time-lag correlations measures. The lag correlation assesses whether the increases in GT data were correlated with the

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

subsequent increase in COVID-19 cases and how far (days) the two series are offsets.[22] Pearson correlation coefficient was used to correlate between public interest and COVID-19 cases for a time lag between 30 and +30 days. Data of south Asian countries were also compared with the United States which was one of the most affected countries by the COVID-19 pandemic to investigate whether these searching terms objective for other communities or not. For the general worldwide interest and confirmed new case data, the period was set from January 22 to March 17, 2020, while for the countries it was set from February 15 to March 17. For the South Asian countries' time lag and correlation analysis, case data from February 22 to March 17 were used. Each country's data was examined individually, and no direct comparison was made between countries in COVID-19 data or RSV index data. For each time frame, a new google trend dataset for was retrieved and matched with the official COVID-19 confirmed case data for further analysis. All the statistical analyses were performed using R studio and p value less than 0.05 was considered significant. IRB approval was not required because this study did not involve human subjects.

RESULTS

Worldwide public interest in coronavirus started to increase from January 22nd,2020 and reached its first peak on January 31, 2020, when WHO declared Covid-19 as a national healthcare emergency (Figure 1). Then COVID-19-related worldwide searches remain low for some time and continuously increased after the word was spread on the outbreak in Wuhan, China. Worldwide COVID-19 related searches. COVID-19-related worldwide public interest continued to expand and reached a peak on March 16, 2020, as worldwide coronavirus cases and related death were reported and right after the WHO announcement to declare the coronavirus outbreak as a pandemic. Worldwide COVID-19 related searches remained steadily high for 1 week and again reached a peak on March 22, 2020, after the massive spread of coronavirus in Europe and the US. There are two small peaks, one sharp increase in numbers on February 12, 2020, when china adjusted coronavirus cases from confirmatory laboratory test and the other peak on April 12, 2020, due to cases around the globe. As of 29 July 2020, the daily number of confirmed coronavirus cases still increasing almost every day and has not reached its highest point yet.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The number of reported confirmed COVID-19 cases in Bhutan, Afghanistan, Maldives, Nepal, and Sri Lanka were at a low level compared to other South Asian countries. India, Pakistan, and Bangladesh were the most COVID-19 affected three South Asian countries that overtook China in terms of the number of coronavirus cases. The COVID-19 related searches in India, Pakistan, and Bangladesh researched its initial peak right after the 1st confirmed case reported in the respective country. After the first peak, public interest in COVID-19 related information increased rapidly after the dissemination of confirmed cases in mainland China and Europe. Public searches about the coronavirus in these countries reached its second major peak right after the announcement of strong lockdown measures and a few days after the numbers of newly locally confirmed COVID-19 started to increase exponentially. Internet COVID-19 related searches reached the highest peak in India, Pakistan, and Bangladesh on March 28, April 05, and April 03, 2020, a few days after each county reported 100th confirmed cases. Then public interest in these countries remained almost steady the same for two-three weeks with rapid increases in the number of COVID-19 cases. After the end of late April, COVID-19 related searched continuously declined similar to worldwide trends. The daily number of confirmed COVID-19 cases in Bangladesh and Pakistan reached a peak on 02 July 2020 and 06 June 2020, respectively. The daily number of confirmed COVID-19 cases have dipped from peak level in any South Asian Counties except India (Supplemental Table 1).

From the correlation analysis, the highest public interest in COVID-19 was on average 12 days before the local maximum of new confirmed cases (Figure 3). High to moderate statistically significant ($p < 0.05$) correlation coefficient between COVID-19 related public interest and COVID-19 were found in south Asian Countries except for Bhutan, Nepal, and Sri Lanka. India and Bangladesh show a similar pattern with the highest correlation between RSV indices and newly confirmed cases found with time lags of -16 and -12 days, respectively. In addition, Pakistan shows a similar positive correlation pattern with the highest coefficient found with a slightly shorter time lag of -9 days than the two countries. All three countries show comparable patterns in correlation coefficient with the US, where the highest coefficient found with a slightly longer time lag of -19 days. According to Figure 3, the dynamics of web searches related to personal hygiene and COVID-19 preventive measures in India, Pakistan, and Bangladesh were related to vulnerability due to confirmation of local transmission of COVID-19 and public restlessness given the announcement of national restriction measures.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

Over the last few months, the coronavirus pandemic has become a global public health crisis and an unprecedented in the disruption of daily life around the world. With the rapid spread of coronavirus, people have faced extreme uncertainty due to evolving information and a strong desire for information about both coronavirus and its corresponding novel behavioral changes. In this study, we found substantial changes in coronavirus information-seeking countries all over the world, particularly in March 2020, suggesting a hyper-awareness and mass interest for information about COVID-19 pandemic. In addition, we found a varied level positive correlation between public interest and real-world cases in South Asian countries, with internet searches preceding real-world cases by on average 12 days. This lag time reflects the increases in the public attention weeks before the increases in COVID-19 cases occurred in most South Asian counties. This lag time biases further justices the importance of pursuing more real-time assessments of COVID-19 pandemic growth when there is an increase in internet activities as people develop symptoms.[16,23] These findings indicate that internet query searches can be a useful tool to predict and monitor the outbreak of infectious diseases, thus can provide a valuable picture of the outbreak of COVID-19 pandemic progression in real-time.

Close monitoring and continued evolution of enhanced communication strategies are urgently needed to provide general populations and vulnerable populations with actionable information for self-protection and clear guidance during an outbreak.[24] The application of electronic medium, specifically the internet data in health care research, known as infodemiology, is a promising new field that provides unmatched opportunities for the management of health information generated by the end-users.[10] Using this unique potential, previous researchers were able to correlate the internet searches with traditional surveillance data and can even predict the outbreak of infectious disease several days or weeks earlier.[11,13] Recent COVID-19 related infodemiology studies modeled daily laboratory-confirmed / suspected cases and associated death with internet search queries in the US, China, Iran and several European countries.[16–19,23,25] Also, the regression/ machine learning model using the Google search queries moderately predicted the incidence of COVID-19 in Iran. Similar attempts were subsequently made to predict the previous coronavirus related outbreaks (e.g., SARS, MERS) and other infectious diseases (e.g., Zika, dengue).[11,13,26]

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Despite the studies above, there is still a lack of research on this theme in South Asia and so fur there is no similar COVID-19 infodemiology study in the countries of this region. According to WHO, South Asia is considered highly vulnerable to any large-scale outbreak of an infectious disease being one of the most populous word region.[27] Therefore, the use of infodemiology particularly internet searches provides unique opportunities to monitor real-time public interest and awareness, particularly in countries with a lack of diagnostic and surveillance capacity, and thereby disseminate evidence-based health information to the people.[28] In South Asia, the first imported coronavirus case was reported in Nepal on 23 January 2020.[6] Starting from the March 2020 number of coronavirus cases started to increase rapidly in South Asian countries except for Bhutan (Supplemental Table 1). Countries in the region instituted various levels of national restriction measures and ban on international travel starting from the mid-March 2020.[29] The number of reported confirmed COVID-19 cases in Bhutan, Maldives, Nepal, and Sri Lanka were at a low compared to other South Asian countries. These countries are either completely sea locked or sparsely populated and completely sealed their international border, which may have contributed low COVID-19 infections compared with their neighbors.[30,31] According to the WHO's report on the 11th March 2020, that the outbreak of COVID-19 meets the criteria for a global pandemic and people's interest started to increase in all south-Asian Countries. However, People in India and Pakistan had a faster response speed toward COVID-19 than Bangladesh which is likely to be related to the number of local cases and lower virulence to COVID-19. As of the 15th March 2020, the number of patients confirmed with COVID-19 in Bangladesh (05 cases) was lower than in India (110 cases), and Pakistan (31 cases).[6] After the peaks within two-three weeks, internet searches continued to decline declined due to massive dissemination of information reported on the local/national news reporting, video news reporting, and health expert reporting in social media.[16,17] These findings suggest that internet searches can potentially help governments to define proper timing of risk communication, improve the public's vigilance, and strengthen the publicity of precautionary measures when facing any public health emergencies like COVID-19.[32]

This study has some limitations that should be acknowledged. First, this study used single search engines, Google, to retrieve population interest data form the South Asian Countries. Thus, there might be selection bias since people who use other search engines are not included in this investigation. However, since more than 880 million people in South Asia use the internet and

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

google as the major search engine (more than 98% market share), google search queries can be a strong tool to estimate public interest.[9,33] Second, Google trend do not report search query result the form of a relative search value instead of absolute search volumes which might have limited more in-depth and precise investigations. In addition, the Google trend excludes all the search results with any typographical error in the query terms. Third, although the number of studies based on google trend is increasing, Google does not provide the detailed information about the procedures by which they generate search data, and the study population responsible for the searches remain unclear.[13] Finally, search volumes can be influenced by the dissemination of information through the news media and it is still unclear whether changes in online activity translate to changes in health behavior. This requires caution when analyzing results and making interpretations from the analyses.

Funding

This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*. 2020;20(5):533-534.

2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020;395(10223):470-473.

3. Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *J Adv Res*. 2020;24:91-98.

4. Bai Y, Yao L, Wei T, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA - J Am Med Assoc*. 2020;323(14):1406-1407.

5. Timeline: WHO's COVID-19 response. World Health Organization. Published 2020. Accessed November 8, 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline/

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

6. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE). Johns Hopkins University. Published 2020. Accessed August 8, 2020. https://github.com/CSSEGISandData/COVID-19

7. Ybarra ML, Suman M. Help seeking behavior and the Internet: A national survey. *Int J Med Inform*. 2006;75(1):29-41.

8. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. *Heal Inf Libr J*. 2018;35(2):91-120.

9. Digital 2020: Global Digital Overview. DataReportal. Accessed November 8, 2020. https://datareportal.com/reports/digital-2020-global-digital-overview

10. MAvragani A. Infodemiology and infoveillance: Scoping review. *J Med Internet Res*. 2020;22(4):e16206.

11. Barros JM, Duggan J, Rebholz-Schuhmann D. The application of internet-based sources for public health surveillance (infoveillance): Systematic review. *J Med Internet Res*. 2020;22(3):e13680.

12. Carneiro HA, Mylonakis E. Google trends: A web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis*. 2009;49(10):1557-1564.

13. Nuti S V., Wayda B, Ranasinghe I, et al. The Use of Google Trends in Health Care Research: A Systematic Review. Voracek M, ed. *PLoS One*. 2014;9(10):e109583.

14. Jella TK, Samuel LT, Acuña AJ, Emara AK, Kamath AF. Rapid Decline in Online Search Queries for Hip and Knee Arthroplasties Concurrent With the COVID-19 Pandemic. *J Arthroplasty*. Published online 2020:1-7.

15. Kutlu Ö. Analysis of dermatologic conditions in Turkey and Italy by using Google Trends analysis in the era of the COVID 19 pandemic. *Dermatol Ther*. 2020;(June):1-6.

16. Husnayain A, Fuad A, Su ECY. Applications of Google Search Trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. *Int J Infect Dis*. 2020;95:221-223.

17. Keller M, Blench M, Tolentino H, et al. Use of unstructured event-based reports for global

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

infectious disease surveillance. *Emerg Infect Dis*. 2009;15(5):689-695.

18. Arora VS, McKee M, Stuckler D. Google Trends: Opportunities and limitations in health and health policy research. *Health Policy (New York)*. 2019;123(3):338-341.

19. Google Trends. Accessed August 8, 2020. https://trends.google.com/trends/?geo=US

20. Mavragani A, Ochoa G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. *JMIR public Heal Surveill*. 2019;5(2):e13439.

21. Alexander J, Seth J, Christopher E. Silver lining of COVID-19: Heightened global interest in pneumococcal and influenza vaccines, an infodemiology study Joseph. 2020;(January).

22. Mavragani A, Ochoa G, Tsagarakis KP. Assessing the methods, tools, and statistical approaches in Google trends research: Systematic review. *J Med Internet Res*. 2018;20(11):1-20.

23. Effenberger M, Kronbichler A, Shin J Il, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with Internet Search Volumes: A Google TrendsTM Analysis. *Int J Infect Dis*. 2020;95:192-197.

24. Heymann DL, Shindo N. COVID-19: what is next for public health? *Lancet*. 2020;395(10224):542-545.

25. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, R Niakan Kalhori S. Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. *JMIR Public Heal Surveill*. 2020;6(2):e18828.

26. 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;15(7).

27. Gupta I, Guin P. Communicable diseases in the South-East Asia Region of the World Health Organization: Towards a more effective response. *Bull World Health Organ*. 2010;88(3):199-205.

28. Ortiz-Martínez Y, Garcia-Robledo JE, Vásquez-Castañeda DL, Bonilla-Aldana DK, Rodriguez-Morales AJ. Can Google® trends predict COVID-19 incidence and help

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

preparedness? The situation in Colombia. *Travel Med Infect Dis*. Published online 2020.

29. Kutlu Ö. Analysis of dermatologic conditions in Turkey and Italy by using Google Trends analysis in the era of the <scp>COVID</scp> 19 pandemic. *Dermatol Ther*. Published online July 27, 2020.

30. Coronavirus in South Asia, June 2020: Cases in India, Bangladesh, and Pakistan Spike | Council on Foreign Relations. Accessed August 12, 2020. https://www.cfr.org/blog/coronavirus-south-asia-june-2020-cases-india-bangladesh-and-pakistan-spike

31. Paul A, Chatterjee S, Bairagi N. Prediction on Covid-19 epidemic for different countries: Focusing on South Asia under various precautionary measures. *medRxiv*. Published online April 11, 2020:2020.04.08.20055095.

32. Hu D, Lou X, Xu Z, et al. More effective strategies are required to strengthen public awareness of COVID-19: Evidence from google trends. *J Glob Health*. 2020;10(1).

33. Search Engine Market Share Worldwide | StatCounter Global Stats. Accessed August 13, 2020. https://gs.statcounter.com/search-engine-market-share/all/

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

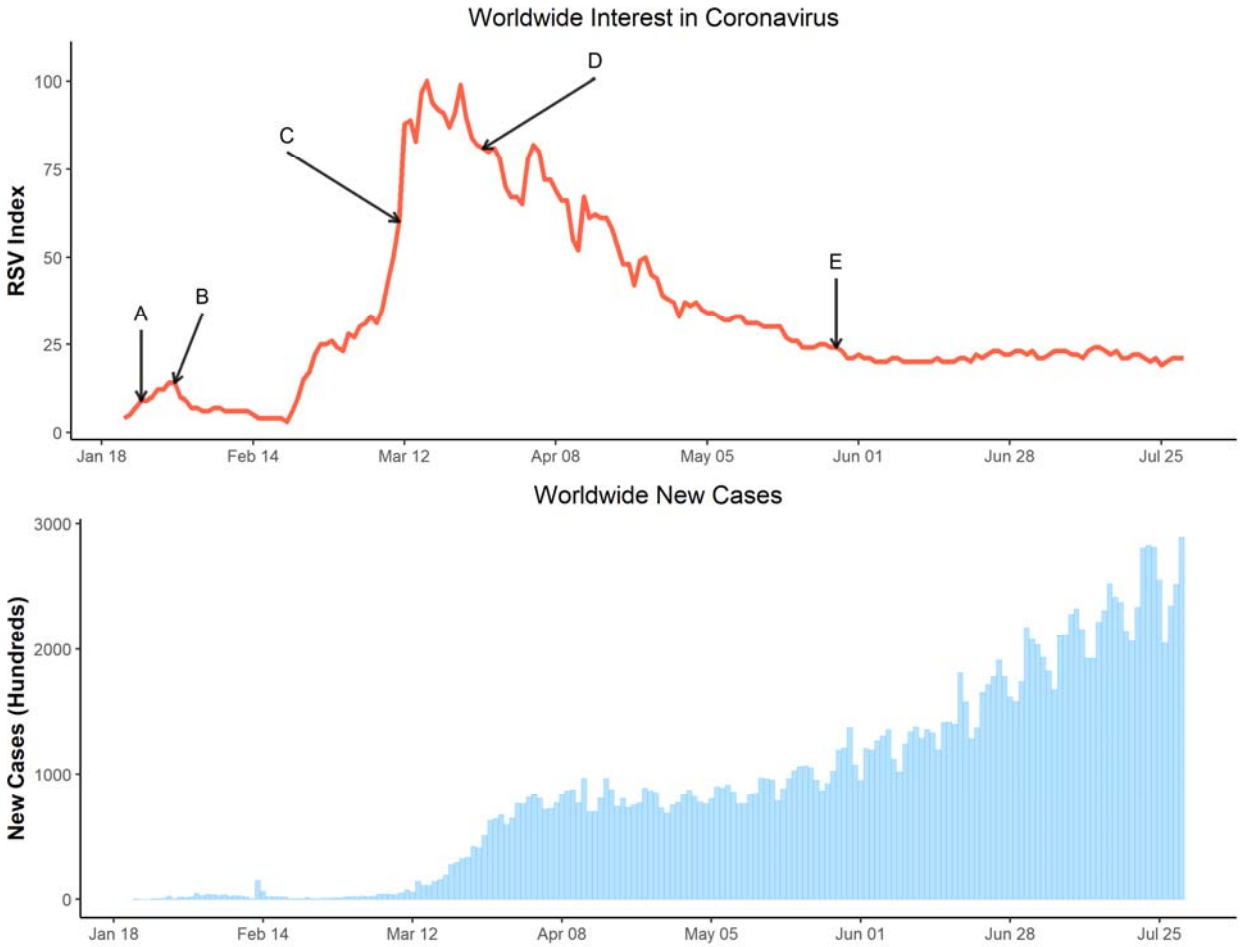
Table 1: Timeline Of COVID-19 Progression and National Restriction Measures

Country	Confirmed 1st case (date)	Total 1000 cases (date)	Total 10,000 cases (Date)	Highest daily cases	lockdown measures (date)	Lockdown eases (date)	Total Confirmed cases (07-30-20)
Afghanistan	24 Feb-20	20 Apr-20	24 May-20	05 June-20	14 Mar-20	21 May-20	36675
Bangladesh	08 Mar-20	14 Apr-20	04 May-20	02 July-20	23 Mar-20	30 May-20	237661
Bhutan	20 Mar-20	-	-	06-June-20	22-Mar-20	-	101
India	01 Jan-20	29 Mar-20	14 Apr-20	-	25 Mar-20	31 May-20	1695988
Maldives	08 Mar-20	15 May-20	-	30 Apr-20	-	-	3793
Nepal	23 Jan-20	28 May-20	23 June-20	07 July-20	24 Mar-20	24 July-20	19771
Pakistan	26 Feb-20	24 Mar-20	21 Apr-20	06 June-20	24 Mar-20	09 May-20	278305
Sri Lanka	27 Jan-20	19 May-20	-	10 July-20	20 Mar-20	11 May-20	2815

medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

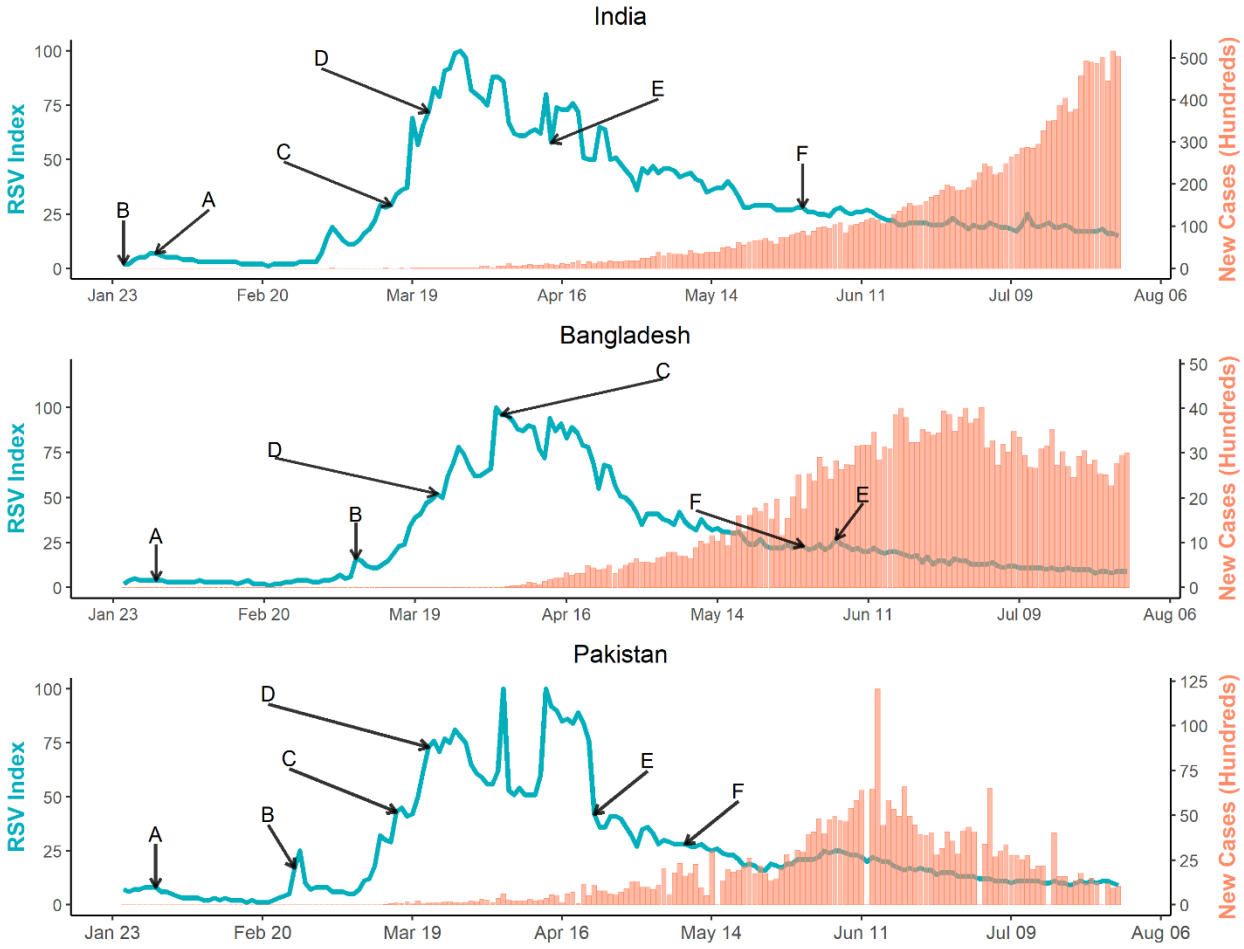
List of Figures

Figure 1: Time series of relative search volumes (RSVs) related to COVID-19 and worldwide daily cases. Note, A: 1000 case reported worldwide [January 25,2020], B: WHO declared health emergency [January 31,2020], C: WHO declared COVID-19 as pandemic [March 11,2020], D: 1000 death reported in USA [March 26, 2020], E: 100000 death reported in USA [May 28 , 2020]



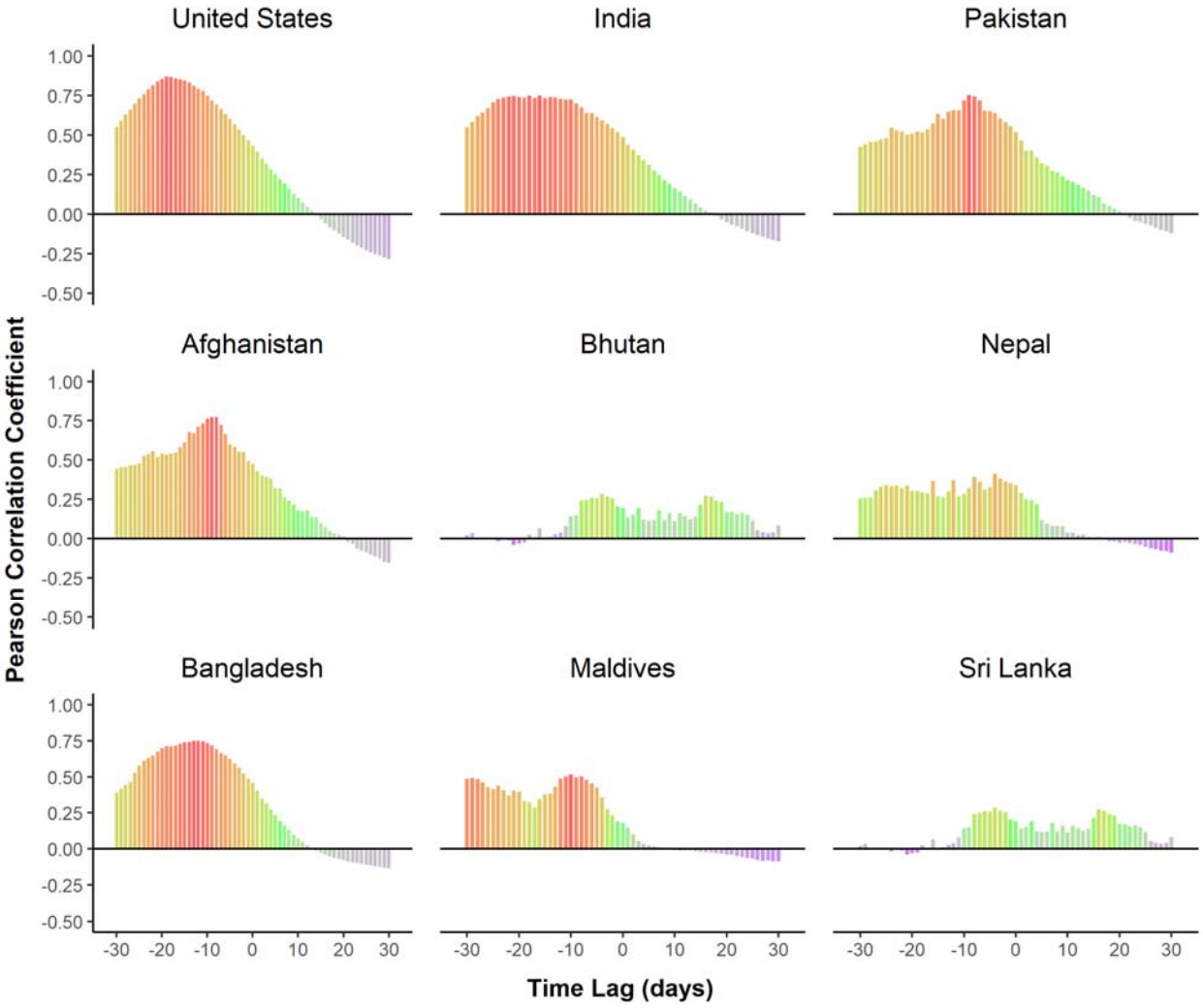
medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2: Time series of relative search volumes (RSVs) related to COVID-19 and daily cases in India, Bangladesh, and Pakistan. Note, A: WHO declared health emergency, B: 1st COVID-19 case reported, C: WHO declared COVID-19 as pandemic, D: National lockdown restriction implemented, E: 10000th COVID-19 cases reported, F: National lockdown restrictions lifted



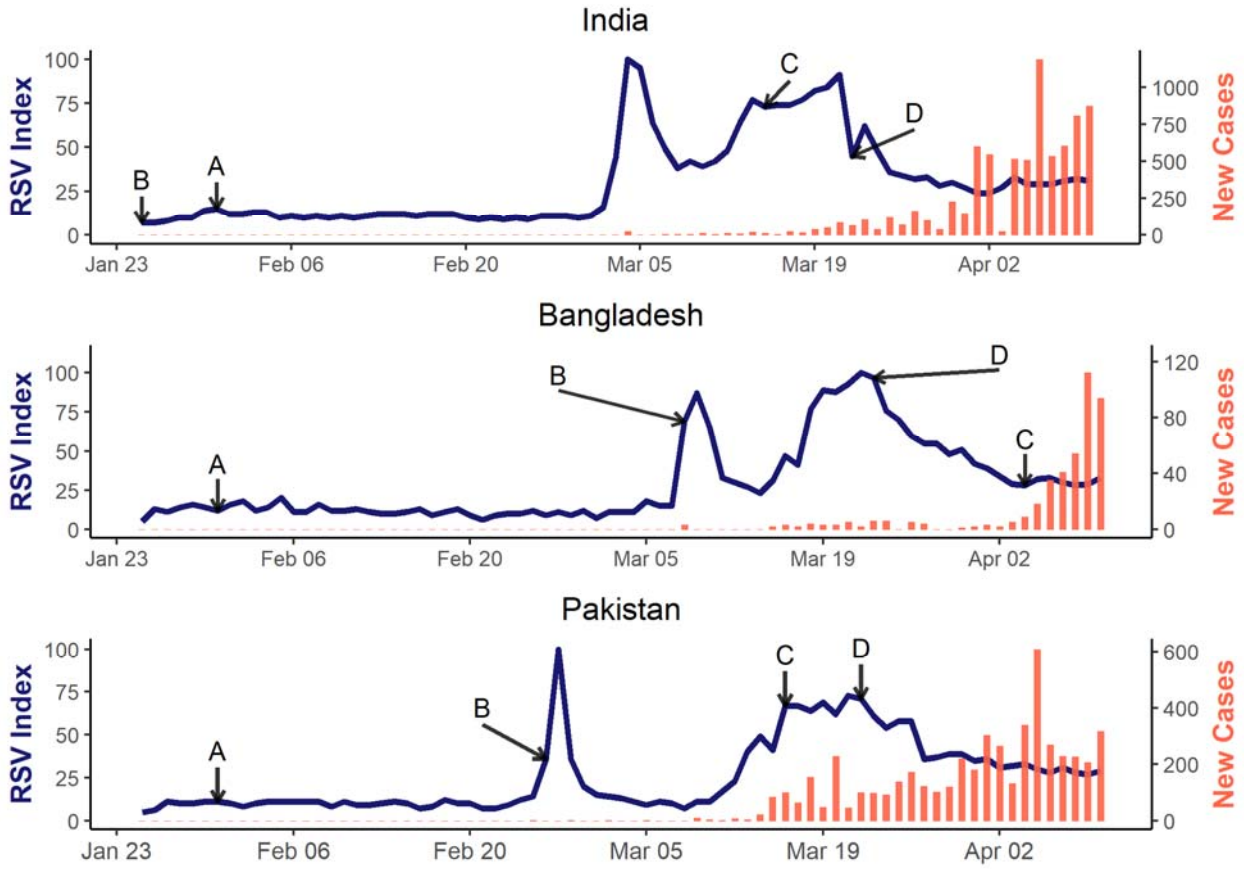
medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3. Time-lag correlations of Relative Search Volume (RSV) indices for COVID-19 and daily new confirmed COVID-19 cases for South Asian Countries



medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 4: Time series of relative search volumes (RSVs) related to COVID-19 health literacy and daily cases in India, Bangladesh, and Pakistan. Note, A: WHO declared health emergency, B: 1st COVID-19 case reported, C: 1000th COVID-19 case reported, and D: National lockdown restriction implemented



medRxiv preprint doi: https://doi.org/10.1101/2020.08.24.20180943; this version posted August 26, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .