Journal of Public Health: From Theory to Practice
https://doi.org/10.1007/s10389-020-01295-y

COMMENT

COVID-19 and digital epidemiology

Ehsan allah Kalteh[1] · Abdolhalim Rajabi[2]

Received: 3 April 2020 / Accepted: 22 April 2020

Since December 2019, a new viral pneumonia has emerged in Wuhan, China. This disease is a zoonotic coronavirus which is similar to SARS and MERS coronaviruses. The virus was called nCov – 2019 and finally named COVID-19 (Liu et al. 2020). As of 27 January 2020, a total of 4515 cases of COVID-19 infections and 106 deaths were reported (Velavan and Meyer 2020). One hundred and six cases of COVID-19 infections were also confirmed outside China, in other Asian countries, the USA, France, Australia, and Canada (Jung et al. 2020). Most of this population were individuals who lived or had a history of traveling to Wuhan. Therefore, the human-to-human transmission of this viral agent which mainly transfers through droplets was approved (World Health Organization 2020). According to the increasing growth rate of coronavirus cases, this outbreak was declared a Public Health Emergency of International Concern on 30 January 2020 by the WHO Emergency Committee (Velavan and Meyer 2020). According to the most recent health statistics released by the WHO, since the first infected cases in Wuhan, more than 735,875 cases of COVID-19 infections and 34,949 deaths were reported from all over the world. A total of 142,410 infected cases (19.35%) and 2,505 dead cases (7.16%) were from the USA. There were 593,465 infected cases (80.65%) and 32,444 dead cases (92.84%) from countries other than the USA (World Health Organization 2020).

For understanding and responding to the disease, the first efforts focus on describing the clinical course, case fatality rate, and treatment of patients. The epidemic of MERS, pandemic influenza, etc. have determined that recognition of disease epidemiology and its possible effects is a priority. Also, real-time analyses of epidemiological data are needed to increase situational awareness and inform interventions (Lipsitch et al. 2020). Epidemiology is the study of the occurrence, distribution, and determinants of health-related states or events in a given population, and the application of this study to the control and management of other health problems (Gordis 2009). As with the role of philosophy of science among other sciences, the introduction and methodological development of the study is one of the major objectives of epidemiology. Thus, epidemiology should be able to keep itself up-to-date on scientific developments and the emergence of new tools, and also provide a new field of expertise for researchers (Masodi et al. 2019). In the recent decade, the explosion of internet and mobile phone usage has led to a new type of epidemiology known as digital epidemiology (Masodi et al. 2019; Park et al. 2018; Salathé 2018).

The term digital epidemiology was defined by Marcel Salathé as epidemiology that uses data that was generated outside the public health system, i.e., with data that was not generated with the primary purpose of doing epidemiology (Salathé 2018; Eckmanns et al. 2019). Digital epidemiology uses digital methods from data collection to data analysis (Park et al. 2018). In the last few years, due to the increasing availability of digital data, and access to successful analytics methods, the number of scientific publications, events, and academic groups related to digital epidemiology has continually and rapidly increased (Salathé and policy 2018). Nowadays, there are more freely accessible tools that allow users access to digital data, which may provide deep insight into health-related phenomena and population behavior. However, there is limited knowledge of the uses and limitations of digital data for epidemiological studies (Park et al. 2018). Digital data include mobile phone network data, data generated by sensors, and data collected at call centers, social media posts (e.g., Twitter and Instagram), search terms and webpage access logs (e.g., Google Google Flu Trends (GFT) Google Trends, Wikipedia views) (Masodi et al. 2019; Park et al. 2018; Salathé 2018; Eckmanns et al. 2019; Lippi et al. 2019; Salathe et al. 2012). The potential application of this new branch is numerous and multifaceted, e.g., disease monitoring and detecting, and examining new substance use

✉ Abdolhalim Rajabi
rajabiepid@gmail.com

[1] Infectious Disease Research Center, Golestan University of Medical Sciences, Gorgan, Iran

[2] Environmental Health Research Center, Faculty of Health, Golestan University of Medical Sciences, Gorgan 1449614535, Iran

Published online: 30 April 2020

patterns (Lippi et al. 2019). A large proportion of sick people search for relevant health information using internet search engines, and many share their experience on social media. This type of data collection of health problems is provided time-stamped and geo-tagged (geographic location). Thus, we can study the general health of a population in real-time using such digital traces. Researchers have already started to use digital data to support public health surveillance and infectious disease monitoring or to understand public attitudes, perceptions, and behaviors towards health issues (Park et al. 2018). Thus, digital systems for early detection of infectious diseases such as ProMED-Mail, Global Public Health Intelligence Network (GPHIN), HealthMap and Google Flu Trends (GFT) are the key components of the Global Public Health System (Eckmanns et al. 2019). Google Flu Trends (GFT) is an early example of digital epidemiology, using search queries to track influenza-like illnesses (ILIs) (Park et al. 2018, Salathé 2018). In 2009, researchers from Google and the US Centers for Disease Control and Prevention (CDC) published a method to estimate flu activity by region using search queries. The search engine query data of GFT reflect influenza cases in real-time and optimize the detecting time and appropriate action against the epidemics. For many years, Google Trends (GT) has served as a prime data source for digital epidemiology (Park et al. 2018; Dugas et al. 2013; Dugas et al. 2012; Santillana et al. 2014). The digital data can be used to study infectious disease dynamics (Salathé et al. 2013; St Louis and Zorlu 2012). A group of researchers have used data from Twitter to track the level of disease activity and concern about influenza (Park et al. 2018; Salathé et al. 2013; St Louis and Zorlu 2012; Broniatowski et al. 2013).

Increasing numbers of epidemiological studies are using digital data generated for a purpose other than epidemiology. Park et al. 2018 reviewed digital epidemiological studies to characterize them by topic domain, study purpose, data source, and analytic method. There were six main topic domains: infectious diseases (58.7%), non-communicable diseases (29.4%), mental health and substance use (8.3%), general population behavior (4.6%), environmental, dietary, and lifestyle (4.6%), and vital status (0.9%). There were eight categories for the data sources: web search query (52.3%), social media posts (31.2%), web portal posts (11.9%), webpage access logs (7.3%), images (7.3%), mobile phone network data (1.8%), global positioning system data (1.8%), and others (2.8%). Of these, 50.5% used correlation analyses, 41.3% regression analyses, 25.6% machine learning, and 19.3% descriptive analyses (Jung et al. 2020).

Despite its growing importance, digital epidemiology is faced with various challenges, such as privacy issues, the potential nature of skewed data, information accuracy, and the effect of news coverage, media, and the general population's knowledge about diseases on the search volume. There are weak coverage, slow tracking systems, low reporting or lack of information on the social–cultural problems in Iran (Salathe et al. 2012); thus, a part of the high COVID-19 mortality rate in Iran can be related to the incomplete reports and contradictory description and diagnosis of the disease. Therefore, along with the increasing use of smartphones, expanding social networks among the public, developing online databases, and providing proper analytical methods double the importance of using digital epidemiology in tracking COVID-19.

Compliance with ethical standards

Conflict of interest None declared.

References

Broniatowsk DA, Paul MJ, Dredze M (2013) National and local influenza surveillance through Twitter: an analysis of the 2012–2013 influenza epidemic. PLoS One 8(12):e83672.

Dugas A F, Hsieh Y-H, Levin SR, Pines JM, Mareiniss DP, Mohareb A., Gaydos CA, Perl TM, Rothman RE (2012) Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. Clin Infect Dis 54(4):463–469

Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, Rothman RE (2013) Influenza forecasting with Google Flu Trends. PLoS One 2013 8(2):e56176

Eckmanns T, Füller H, Roberts SL (2019) Digital epidemiology and global health security; an interdisciplinary conversation. Life Sci Soc Policy 15(1):2

Gordis L (2009) Epidemiology. Saunders, Philadelphia

Jung S-M, Akhmetzhanov AR, Hayashi K, Linton NM, Yang Y, Yuan B, Kobayashi T, Kinoshita R, Nishiura HJ (2020) Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: Inference using exported cases. J Clin Med 9:523

Lippi G, Mattiuzzi C, Cervellin G (2019) Is digital epidemiology the future of clinical epidemiology? J Epidemiol Glob Health 9(2):146

Lipsitch M, Swerdlow DL, Finelli L (2020) Defining the epidemiology of Covid-19—studies needed. N Engl J Med 382(13):1194–1196

Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med 27(2):taaa021

Masodi S, Hashim M, Yavari P (2019) Digital epidemiology: a new perspective on modern epidemiology. 9th Iranian Congress of Epidemiology. 22-24 October 2019; Shahid Beheshti University of Medical Sciences School of Health and Safety.

World Health Organisation (2020) Coronavirus disease 2019 (COVID-19): situation report — 59. WHO, Geneva

Park H, Jung H, On J, Park SK, Kang H (2018) Digital epidemiology: use of digital data collected for non-epidemiological purposes in epidemiological studies. Healthc Inform Res 24(4):253–262

Salathe M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, Campbell EM, Cattuto C, Khandelwal S, Mabry PL (2012) Digital epidemiology. PLoS Comput Biol 8(7):e1002616

Salathé M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS (2013) Influenza A (H7N9) and the importance of digital epidemiology. N Engl J Med 369(5):401–404

Salathé M (2018) Digital epidemiology: what is it, and where is it going? Life Sci Soc Policy (2018) 14(1):1

Santillana M, Zhang DW, Althouse BM, Ayers JW (2014) What can digital disease detection learn from (an external revision to) Google Flu Trends? Am J Prev Med 47(3):341–347

St Louis C, Zorlu G (2012) Can Twitter predict disease outbreaks? BMJ344:e2353

Velavan TP, Meyer CG (2020) The COVID-19 epidemic. Trop Med Int Health 25(3):278–280