Public Risk Perception Explains the Mitigation of COVID-19

Siyu Lai
College of Systems Engineering, National University of Defense Technology, Changsha, China
e-mail: laisiyu@nudt.edu.cn

Tong Wang
College of Systems Engineering, National University of Defense Technology, Changsha, China
e-mail: wangtongnudt@163.com

Ziqiang Cao
POWERCHINA Zhongnan Engineering Corporation Limited, Changsha, China
e-mail: ziqiangcao07@gmail.com

Mengsi Cai
College of Systems Engineering, National University of Defense Technology, Changsha, China
e-mail: mengsi.cai92@gmail.com

Suoyi Tan
College of Systems Engineering, National University of Defense Technology, Changsha, China
e-mail: tansuoyi_cn@outlook.com

Jianguo Liu
Institute of Accounting and Finance, Shanghai University of Finance and Economics, Shanghai, China
e-mail: jianguo@sufe.edu.cn

Xin Lu[1,2*]
[1] College of Systems Engineering, National University of Defense Technology, Changsha, China
[2]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden
[*]Corresponding author: xin.lu.lab@outlook.com

***Abstract*—Public awareness of self-protection (PASP) was of vital importance in predicting the spread of infectious diseases. Based on large-scale Weibo and Twitter datasets, we analyzes the temporal patterns of PASP for COVID-19 and develop improved models integrating PASP to predict the spread of COVID-19. Results indicate that in the first two months of local outbreaks with mitigation actions, the rate of online users with PASP in China and UK increased by 53% and 26%, respectively. And the integrated models yield an improvedof 96.57% and 95.12% for predicting outbreaks in China and UK. Additionally, we implement the models to evaluate non-pharmaceutical intervention strategies such as travel restrictions and prove that PASP and timely travel interventions are both effective in containing the spread of COVID-19. Our study reveals that measuring public response had instructional significance in epidemiological models and was important in infectious disease prevention and control.**

Keywords- COVID-19; PASP; Sina Weibo; Twitter; risk perception; infodemiology

I. INTRODUCTION

Until February 18, 2022, the novel coronavirus (COVID-19) pandemic has caused 420,166,191 confirmed cases and 5,865,242 deaths in over 222 countries, areas, or territories globall.[1] The pandemic has changed human psychology, the educational system, the global economy, and fundamentally impacted the world.[2,3]

Massive research efforts have been devoted to accurately predicting the development trend of COVID-19 ever since its emergence. Previous models can be divided into three categories: infectious disease models, infodemiology models and machine learning models. The infectious disease models, or compartment models, divide the population into a number of categories according to their states of infection, and describe the transmission process of infectious agents in the host population. Susceptible-exposed-infectious-removed (SEIR) models[4,5] are the most common classic models. With the Exposed (E) state of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the SEIR models consider the complicated situation of COVID-19 and portray its transmission mechanism more practically.[6] Cai et al. studied the dynamic evolution of COVID-19 caused by the Omicron variant via a fractional SEIR model and used the model to make short-time prediction.[7]

With the ultimate aim to inform public health and policy, infodemiology is mainly the science of distribution and determinants of information in the Internet.[8] The analysis of queries from Internet search engines such as Google,[9-11] and monitoring public status updates on social media such as Twitter[12,13] and Sina Weibo[14-17] are the two primary ways for infodemiology studies to predict the outbreaks and trends of COVID-19. Google mobility data was used to assess the relationship between different community activities and the pandemic transmission rate. There was a strong negative correlation between the number of COVID-19 cases and the number of COVID-19 searches on Google Trends, showing a statistical significance.[18] Besides, with the advancement of time-series modeling and prediction, machine learning model can also be used to predict the spread of COVID-19.[19-21] Prediction models that combine several features to estimate the risk of infection have been developed. Zoabi et al. established

2024 Sixth International Conference on Next Generation Data-driven Networks (NGDN) | 979-8-3503-8837-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/NGDN61651.2024.10744181

a machine-learning approach that detects COVID-19 cases by simple features accessed by asking basic questions, and their framework can be used to prioritize testing for COVID-19 when testing resources are limite.[22]

During this pandemic, people have been extensively using social media such as Twitter, Facebook and Sina Weibo to share their awareness, knowledge and behavior about COVID-19. Texts like “cloud greetings”, “travel refrain”, “home isolation”, “vaccination” as well as “wearing masks”, reflecting the self-protection measures initiative taken by the Internet users to resist the COVID-19 pandemic. As a result of self-protection awareness, which conceptually refers to a preemptive attempt to resist the potential threat,[23] these measures such as social distancing, isolating, vaccination, and wearing masks covered by massive text data have made us defensively resistant to COVID-19.[24] Public awareness of Self-protection (PASP) can change the way people travel and socialize, thereby curbing the spread of the novel coronavirus and mitigating its impact. PASP of internet users is not only a reflection of their offline responses, but also serves as a driving force to take precautions to prevent infection. Such influence has been validated by a number of studies. For example, Kaushik et al. determined the role of public awareness in preventing the spread of the COVID-19 outbreak in India, which would be automatically reflected in the societal behavioral response of rigorous precautionary measure.[25] Similar roles were also characterized with Twitter[26] and online survey[27]. However, to the best of our knowledge, little work has qualitatively considered PASP for predicting the pandemic.[28,29] Therefore, to fill this knowledge gap, in this study, we developed two improved models, namely the SEIR-w and SEIRP model, by considering two rational mechanisms of PASP. First, we quantify the evolutionary patterns of PASP through a computational framework of nature language processing, based on massive empirical data from the online social media. Then, we proposed improved models by incorporating the information of PASP into the SEIR model to predict the spreading of COVID-19. Lastly, we adopted the best-fit SEIRP model for policy evaluation of non-pharmaceutical intervention strategies such as travel restrictions, and prove that PASP and timely travel interventions are both vital for containing the epidemic in China and UK.

The rest of the paper is organized as follows: Section 2 describes the data and method of our work, and presents the design and implementation of our proposed SEIR-w and SEIRP models. The experimental results of our framework and models are demonstrated and analyzed in Section 3. Finally, we conclude our work and discuss future research in Section 4.

II. Materials and Methods

A. Data description

To understand PASP and its interplay with the transmission of the pandemic, multisource data including the complete Wuhan Weibo data, COVID-19 related Twitter data, and COVID-19 reporting data were synergistically analyzed in this paper.

1) Sina Weibo Dataset: During the COVID-19 outbreak, large, sudden onset, and disturbing discussions began spreading extensively online. We collected a comprehensive dataset containing all publicly available Sina Weibo posts in Wuhan from December 1, 2019, when symptoms of the first recorded case of COVID-19 in Wuhan appeared,[30] to March 20, 2020. The data consists of 63,077,475 microblogs and 3,514,206 unique users, including key fields: title/content, original/forwarding, nickname (used as an identifier for unique users and anonymized to protecting users’ privacy), posting time, original content, authentication type, etc.

2) Twitter Dataset: To demonstrate the generalizability of our model in different countries, we also collected COVID-19 related English tweets through mature APIs[31] from February 26, 2020, when the epidemic began to accelerate spreading outside China, to October 31, 2021. The massive number of tweets were extracted by the following keywords: “COVD19”, “CoronavirusPandemic”, “COVID-19”, “2019nCoV”, “CoronaOutbreak,” “coronavirus” and “WuhanVirus,” including 471,449,768 tweets and 28,478,123 unique users in 154 countries. Furthermore, we map each tweet to the country where it was posted by identifying its geo-referenced information through the Nominatim API,[32] and successfully retrieved about 61.3% of tweets.

3) COVID-19 reporting data: Official reporting data on the confirmed cases of COVID-19 in Wuhan and the United Kingdom (UK) were combined. Reporting data in Wuhan covers the period of January 23, 2020 to April 16, 2020, and it is obtained from Chinese Center for Disease Control and Prevention. Additionally, we collected reporting data in the UK from February 26, 2020 to August 5, 2020, which also covers the initial outbreak of COVID-19 in the UK and is sufficient for evaluating the effectiveness of improved prediction capability through integrating PASP.

B. PASP extraction

In order to quantify PASP in Wuhan, we first constructed the *COVID-19 dictionary*, with 60 well-documented keywords in Appendix A Table S1 (translated in English) and Table S2 (in Chinese),[33] and the *self-protection awareness dictionary*, with 30 well-documented keywords in Table S3 (translated in English) and Table S4 (in Chinese),[34] and identified texts of self-protection behaviors in response to the COVID-19 pandemic. Then we performed time-series analysis and proposed two indicators to effectively describe PASP.

Let $\Omega_{cov}(t)$ denotes the number of users who have posted at least one COVID-19 related microblog according to the *COVID-19 dictionary* on day t, and $\Omega_{sp}(t)$ be the set of users who had explicitly stated self-protection awareness among $\Omega_{cov}(t)$ according to the *self-protection awareness dictionary* on day t, then the proportion of users with self-protection awareness among those who were concerned about COVID-19 on day *t* is defined as

$$R_{sp}(t)=\frac{|\Omega_{sp}(t)|}{|\Omega_{cov}(t)|} \tag{1}$$

and the cumulative proportion of users with PASP can be calculated by

$$R_{tsp}(t)=\frac{|\Omega_{sp}(1)\cup\cdots\cup\Omega_{sp}(t)|}{|\Omega_{cov}(1)\cup\cdots\cup\Omega_{cov}(t)|} \tag{2}$$

with max(t)=111 days being the length of the study period.

In the following, R_{tsp} was fitted by the least square method,[35] and the growth rate of $R_{tsp}(t)$ provides a foundation for constructing the epidemiological models discussed below.

C. SEIR-w model

1) SEIR-w model: The traditional SEIR model ignores the infectivity of latent individual,[36] and existence of asymptomatic latent infections and self-recovered populations. Furthermore, residents' self-protection awareness will strengthen the protective measures of transmission and reduce mutual contact between susceptible individuals (S) and infected patients (I), which will in turn result in lower contagion rate β. Based on this rationale, we propose the SEIR-w model, which was presented in Figure 1 (a). Compared with the SEIR model, three major improvements have been made: First, the R_{tsp} operates at infectious rate β, suggesting that β would decrease since residents reduced their risk of transmission due to their self-protection awareness. Second, we assume that the contagion rate of incubation (E) varies directly with the contact number and can be quantified as $k\beta$ to account for their infectiousness. Third, the self-recovering rate of E (especially asymptomatic latent infections among E) is supposed to be μ. Its differential expression can be derived as Eq. (3).

2) SEIRP model: Unlike the SEIR-w model, which assume that the self-protection awareness directly influences β, the SEIRP model sets a new compartment P for individuals with self-protection awareness, who have taken strict precautions to prevent infection and are not susceptible. Based on this assumption, we incorporate the following three quantifications for a well-mixed SEIRP model: First, with the development of the epidemic, the vulnerable population (S) gradually become safeguarded and transfer to P at the rate λ. Additionally, in order to illustrate the infectiousness and self-recovery of latent population (E), the ratio of the two can also be quantified as $k\beta$ and μ respectively like the SEIR-w model. The diagram and differential equations of the SEIRP model are shown in Figure 1 (b) and Eq. (4), respectively.

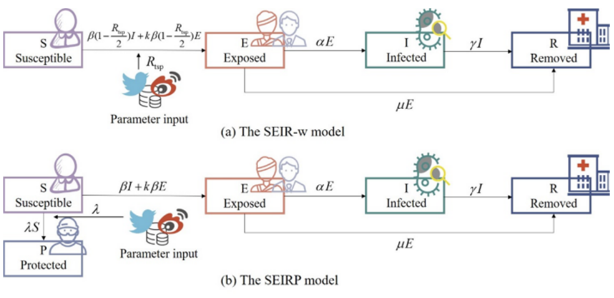

Figure 1. Diagram of the SEIR-w and SEIRP models

D. Experimental settings

To verify the effectiveness of the infodemiology models proposed in this paper, a comparative evaluation was designed combined with the numbers of active infections per day during the epidemic. Initially, the time-series data of cumulative infections was a non-smooth sequence, therefore we eliminated the outlier and realized a stable sequence using an x-day sliding window; Next, the initial parameters of three models, i.e., the classical SEIR and the proposed SEIR-w and SEIRP models, were determined according to existing research and official epidemic data, and then the gradient descent algorithm, a preferred way to minimize the cost function,[37] was adopted to achieve the optimal parameters using the above-mentioned data; Third, the development of COVID-19 was predicted, and the goodness-of-fit[38] of three models was calculated respectively; Lastly, we investigated the effectiveness of non-pharmaceutical intervention strategies such as travel restrictions, which illustrated the impact of PASP and government interventions on the containment of the new coronavirus outbreak.

III. RESULTS

A. PASP extraction and analysis

First, we analyze and describe the temporal patterns of PASP for COVID-19 in Wuhan and UK. Evident trends of $R_{sp}^{(WH)}$ and $R_{tsp}^{(WH)}$ (of Wuhan), and $R_{sp}^{(UK)}$ and $R_{tsp}^{(UK)}$ (of UK) during the study period are demonstrated in Figure 2. For China, in the 60 days after the announcement of human-to-human transmission of SARS-CoV-2 (January 20, 2020), $R_{tsp}^{(WH)}$ had increased from 6.77% to 60.08%; and for UK, in the 60 days after the announcement of the Coronavirus Action Plan (March 3, 2020),[39] $R_{tsp}^{(UK)}$ had increased from 0.52% to 26.96%.

There was a sudden rise for $R_{sp}^{(WH)}$ on December 31, 2019, when Wuhan Municipal Health Commission first reported 27 cases of unexplained pneumonia.[40] What stands out is that the $R_{sp}^{(WH)}$ successively rose from January 20, 2020 and peaked 4 days after. This period coincides with the time when the human-to-human transmission announced by Nanshan Zhong, and lockdown in Wuhan city should also be an important determinant. In addition, during the period from January 24 to February 19, 2020, when Wuhan was severely affected by the outbreak,[41] the daily PASP in Wuhan maintained high ratios. Afterward, it remained when this major outbreak of infectious disease in Wuhan was contained after February 20, 2020. $R_{tsp}^{(WH)}$ reached approximately 0.6 on March 20, 2020, indicating that by the end of that day, about 60% of Weibo users in Wuhan who were concerned about the COVID-19 epidemic had posted relevant microblogs about protecting themselves from the epidemic and calling on others to defensively resist this threatening pandemic.

Based on the least-squares method,[35] we performed a sigmoid fitting for the $R_{tsp}^{(WH)}$, and obtained the best fit with $R_{tsp}^{(WH)}(t)=0.57/1+e^{-0.13\times(t-60.23)}-0.01$ with $R^2 = 0.9881$. For PASP in UK, as influenced by the epidemic in China and the Diamond Princess outbreaks,[42] $R_{sp}^{(UK)}$ was growing at first of the outbreak of COVID-19. UK implemented coronavirus lockdown measures on March 23, 2020[43] and the Queen delivered a televised address on April 5,[44] when $R_{sp}^{(UK)}$ all received phenomenal growth. The optimal curve for $R_{tsp}^{(UK)}$ is $R_{tsp}^{(UK)}(t) = 0.67/1 + e^{-0.03\times(t-147.75)} - 0.13$ with $R^2 = 0.9974$.

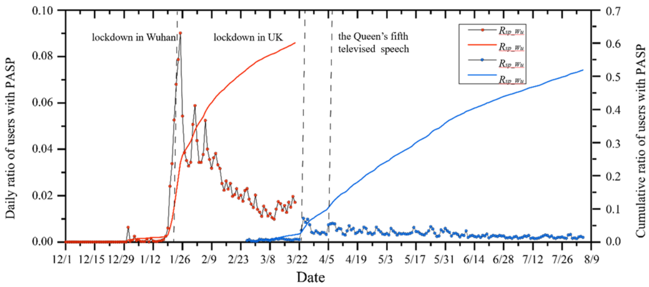

Figure 2. Daily and cumulative ratio of users with PASP

B. Model comparison

To prove the validity of our proposed infodemiology models, we applied them to the complete Wuhan Weibo data and COVID-19 related Twitter data, simulating the early spread in China and UK and comparing the performance of three models.

Although COVID-19 outbroke in the late December of 2019, due to insufficient public knowledge of it, until January 20, 2020, when the academician Nanshan Zhong confirmed its human-to-human transmission and urged people to avoid traveling to Wuhan, did the public realized the severity of this epidemic. Simultaneously, COVID-19 related microblogs have also emerged. Since the implementation of control measures on January 23, 2020, the spread of the epidemic has been affected by government interventions and PASP. Therefore, this paper adopted the daily number of active infections published by CCDC[45] from January 23, 2020 to April 16, 2020 as empirical analysis data, including data from January 23 to February 2 for model fit and data from February 3 to April 16 for the test.

As previously mentioned, after preliminary selection of parameters and optimization of gradient descent method, we ultimately found the optimal parameters of three models for Wuhan epidemic. It is worth noting that the reciprocal of the incubation period was set as the transition rate of latent individuals becoming infected (i.e., α), and the reciprocal of the duration of the disease was estimated as the average rate of recovering from an infection (i.e., γ)[46]. According to Li et al.[47] and Guan et al.[48], these two parameters were determined as 1/5.4 and 1/13, respectively. In addition, the daily growth rate of PASP λ and cumulative rate R_{tsp} were measured according to the sigmoid fitting presented in Section 3.1. Based on the growth rate of the sigmoid fitting line $R_{tsp}^{(WH)}(t) = 0.57/1 + e^{-0.13\times(t-60.23)} - 0.01$ and 0.6 of the $R_{tsp}^{(WH)}$ on March 20, 2020 obtained in Section 3.1, we set λ as $R_{tsp}^{\prime(WH)}(t)$ and R_{tsp} as 0.6.

After selecting the optimal parameters of each model, the SEIR model, SEIR-w model and SEIRP model were trained to predict the spreading of COVID-19 in Wuhan from February 3, 2020 to April 16, 2020, the results are shown in Figure 3 (a).

Following similar methods and steps, we extracted PASP from COVID-19 related tweets according to our *self-protection awareness dictionary*, which was summarized and adjusted according to the characteristics of twitter and the COVID-19 situation. For UK, we also investigated its early epidemic, using tweets from February 26 to April 8 to fit, and tweets from April 9 to August 5 to test, as shown in Figure 3 (b).

From Figure 3, we can see that the SEIR-w model forecasts reliably on the trend in the early stage of the epidemic, and effectively captures the peak time and infections, but it makes an obvious overestimation in the later stage; SEIRP model estimates the peak time and the trend in the mid-term with higher accuracy, with a slight deviation between the predicted and the actual infections in the later stage; For the classic SEIR model, despite a reliable prediction on the late-stage epidemic data than the two models proposed in this paper, it has significant discrepancy in predicting the number of infections of the outbreak peak.

By further calculating the goodness-of-fit of the three models, the R^2 for prediction in Wuhan are 92.44%, 96.57%, and 81.46%, for the SEIR-w model, SEIRP model and SEIR model, respectively; and the R^2 for UK is 91.25%, 95.12%, and 79.87%, respectively. Apparently, the proposed two models are superior to the classic SEIR model in predicting the early development of the epidemic, indicating that PASP forms a crucial part in explaining the spreading of COVID-19.

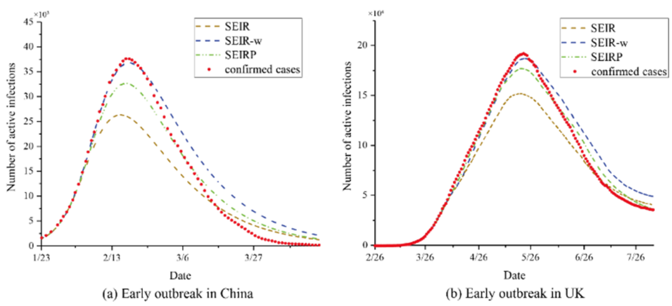

Figure 3. Comparison of prediction by the SEIR, SEIR-w, and SEIRP models in (a) Wuhan and (b) UK.

C. Policy evaluation

As the SEIRP model yields remarkable fit for the reported infections both in Wuhan and in UK, we choose it as the benchmark model for the evaluation experiment to generate simulated scenarios under different strategies. We then perform a policy evaluation about the prevention and control measures, focusing especially on the non-pharmaceutical intervention strategies for COVID-19. More specifically, we considered the situation where travel restrictions such as quarantine, isolation and lockdown ceased, allowing regular migration.

Expecting that some form of control measure such as gathering limitation and wearing masks would continue to be in place to reduce effective social contact, we set the contact-related parameters $\beta = 10$ and $k = 0.8$, with other parameters being fixed. To explore the possible impact of travel interventions on COVID-19 spread, the trend of the COVID-19 epidemic in China and UK was predicted by the estimated parameter values. The results suggest that, had the travel interventions not been implemented, the epidemic in Wuhan city would have peaked by February 8 with 82,519 cases and a final epidemic size of 220,987 cases should have been expected. Similarly, epidemic in UK would have peaked by April 13 with 380,025 cases. It would receive a more phenomenal growth with the peak time 39 days earlier and 551,842 more total infections. The comparison between actual infections and simulated situation of non-travel restrictions in these two countries are shown in Figure 4.

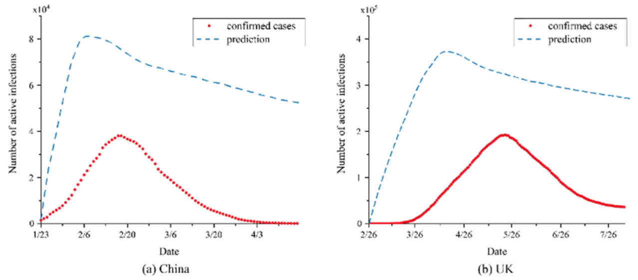

Figure 4. Policy evaluation of infections in confirmed cases and prediction

IV. Conclusion

Risk conception to the infections of COVID-19 and the adoption of self-protection behaviors are of critical importance for connecting the online opinion and offline behavior of the public to understand the epidemic curve. Based on large-scale Weibo and Twitter datasets, this paper first monitored the temporal patterns of PASP for COVID-19. Additionally, we proposed the SEIR-w and SEIRP model, and evaluated the effectiveness of non-pharmaceutical intervention strategies such as travel restrictions in China and UK.

Integrating self-protection awareness of individuals, epidemiological characteristics of COVID-19, and non-pharmaceutical interventions, the proposed SEIR-w and SEIRP model was effective in predicting the epidemic trend of the outbreaks for different countries. In summary, this study presents a new attempt to quantify PASP and exploring it to predict the epidemic trend with massive online social media data. A generalizable and easy-to-use computational framework is provided for communicating the online opinion with offline behavior of internet users, which can portray offline how people react to public health emergencies and allay their concerns and scarce. In addition, our studies serve as a proof-of-concept that PASP and government interventions are essential for the containment of COVID-19 and show promise for future prediction of the epidemic. More generally, our methodology could be extended to other epidemics besides COVID-19.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (72025405, 72088101), the National Social Science Foundation of China (22ZDA102), the Hunan Science and Technology Plan Project (2020TP1013, 2020JJ4673, 2023JJ40685), and the Innovation Team Project of Colleges in Guangdong Province (2020KCXTD040).

References

[1] Coronavirus Resource Center. (2020) Johns Hopkins: Baltimore. https://coronavirus.jhu.ed.

[2] Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. (2020) The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int J Surg., 78:185-93.

[3] Bianchi F, Bianchi G, Song D. (2023)The long-term impact of the COVID-19 unemployment shock on life expectancy and mortality rates. JEDC., 146:104581.

[4] Hatami F, Chen S, Paul R, Thill JC. (2022) Simulating and Forecasting the COVID-19 Spread in a U.S. Metropolitan Region with a Spatial SEIR Model. Int J Env Res Pub He., 19(23):15771.

[5] Ning H, Li Z, Qiao S, Zeng C, Zhang J, Olatosi B, et al. (2023) Revealing geographic transmission pattern of COVID-19 using neighborhood-level simulation with human mobility data and SEIR model: A case study of South Carolina. Int J Appl Earth Obs., 118:103246.

[6] Yang W, Zhang D, Peng L, Zhuge C, Hong L. (2021) Rational evaluation of various epidemic models based on the COVID-19 data of China. Epidemics-Neth., 37:100501.

[7] Cai M, Em Karniadakis G, Li C. (2022) Fractional SEIR model and data-driven predictions of COVID-19 dynamics of Omicron variant. Chaos., 32(7):071101.

[8] Eysenbach G. (2002) Infodemiology: The epidemiology of (mis)information. Am J Med., 113(9):763-5.

[9] Rovetta A, Castaldo L. (2022) A new infodemiological approach through Google Trends: longitudinal analysis of COVID-19 scientific and infodemic names in Italy. BMC Med Res Methodol., 22(1):33.

[10] Nagy B, Gabor MR, Bacoş IB. (2022) Google Mobility Data as a Predictor for Tourism in Romania during the COVID-19 Pandemic—A Structural Equation Modeling Approach for Big Data. Electronics-Switz., 11(15):2317.

[11] Wang B, Liang B, Chen Q, Wang S, Wang S, Huang Z, et al. (2023) COVID-19 Related Early Google Search Behavior and Health Communication in the United States: Panel Data Analysis on Health Measures. Int J Env Res Pub He., 20(4):3007.

[12] Sunitha D, Patra RK, Babu NV, Suresh A, Gupta SC. (2022) Twitter sentiment analysis using ensemble based deep learning model towards COVID-19 in India and European countries. Pattern Recogn Lett., 158:164–70.

[13] Cheng IK, Heyl J, Lad N, Facini G, Grout Z. (2021) Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. Sci Rep., 11(1):19009.

[14] Han XH, Wang JL, Zhang M, Wang XJ. (2020) Using Social Media to Mine and Analyze Public Opinion Related to COVID-19 in China. Int J Env Res Pub He., 17(8):2788.

[15] Huang CM, Xu XJ, Cai YY, Ge QM, Zeng GW, Li XP, et al. (2020) Mining the Characteristics of COVID-19 Patients in China: Analysis of Social Media Posts. J Med Internet Res., 22(5).

[16] Li LF, Zhang QP, Wang X, Zhang J, Wang T, Gao TL, et al. (2020) Characterizing the Propagation of Situational Information in Social Media During COVID-19 Epidemic: A Case Study on Weibo. IEEE Tcss., 7(2):556-62.

[17] Duan YJ, Liu LB, Wang Z. (2021) COVID-19 Sentiment and the Chinese Stock Market: Evidence from the Official News Media and Sina Weibo. Res Int Bus Financ., 58.

[18] Alshahrani R, Babour A. (2021) An Infodemiology and Infoveillance Study on COVID-19: Analysis of Twitter and Google Trends. Sustainability-Basel., 13(15):8528.

[19] Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. (2020) Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ., 369.

[20] Alyasseri ZAA, Al-Betar MA, Abu Doush I, Awadallah MA, Abasi AK, Makhadmeh SN, et al. (2021) Review on COVID-19 diagnosis models based on machine learning and deep learning approaches. Expert Syst., 39(3).

[21] Lalmuanawma S, Hussain J, Chhakchhuak L. (2020) Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos Soliton Fract., 139.

[22] Zoabi Y, Deri-Rozov S, Shomron N. (2021) Machine learning-based prediction of COVID-19 diagnosis based on symptoms. Npj Digit Med., 4(3).

[23] Sherman DK, Hartson KA. (2011) Reconciling self-protection with self-improvement. In: Alicke MD, Sedikides C. (Eds.), Handbook of Self-Enhancement and Self-Protection. The Guilford Press, New York . p. 128-51.

[24] Weissleder R, Lee H, Ko J, Pittet MJ. (2020) COVID-19 diagnostics in context. Sci Transl Med., 12(546):eabc1931.

[25] Kaushik M, Agarwal D, Gupta AK. (2021) Cross-sectional study on the role of public awareness in preventing the spread of COVID-19 outbreak in India. Postgrad Med J., 7(1154):777-81.

[26] Boon-Itt S, Skunkan Y. (2020) Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study. Jmir Public Hlth Sur., 6(4):245-61.

[27] Sun CX, He B, Mu D, Li PL, Zhao HT, Li ZL, et al. (2020) Public Awareness and Mask Usage during the COVID-19 Epidemic: A Survey by China CDC New Media. Biomed Environ Sci., 33(8):639-45.

[28] Jang HJ, Rempel E, Roth D, Carenini G, Janjua NZ. (2021) Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis. J Med Internet Res., 23(2).

[29] Ku D, Yeon C, Lee S, Lee K, Hwang K, Li YC, et al. (2021) Safe traveling in public transport amid COVID-19. Sci Adv., 7(43).

[30] Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al.(2020) A new coronavirus associated with human respiratory disease in China (vol 579, pg 265, 2020). Nature., 580(7803):E7.

[31] Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. (2021) A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration. Epidemiologia., 2(3):315-24.

[32] Imran M, Qazi U, Ofli F. (2022) Tbcov: two billion multilingual covid-19 tweets with sentiment, entity, geo, and gender labels. Data., 7(1):8.

[33] Chang J, Li J. (2021) Spatiotemporal distribution of COVID-19 and public anxiety: Analysis based on micro-blog data. Human Geogr., 36(3):47-57.

[34] Neuberg SL, Kenrick DT, Schaller M. (2011) Human threat management systems: Self-protection and disease avoidance. Neurosci Biobehav R., 35(4):1042-51.

[35] Ding F. (2023) Least squares parameter estimation and multi-innovation least squares methods for linear fitting problems from noisy data. J Comput Appl Math., 426:115107.

[36] Tang B, Wang X, Li Q, Bragazzi NL, Tang SY, Xiao YN, et al. (2020) Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J Clin Med., 9(2).

[37] Desideri JA. (2012) Multiple-gradient descent algorithm (MGDA) for multiobjective optimization. Cr Math., 350(5-6):313-8.

[38] Piepho H P. (2019) A coefficient of determination (R2) for generalized linear mixed models. Biometrical J., 61(4): 860-872.

[39] Susan M. (2016) Behavioural strategies for reducing covid-19 transmission in the general population. https://www.comminit.com/content/behavioural-strategies-reducing-covid-19-transmission-general-population.

[40] XinhuaNet. (2019) Bulletin of wuhan municipal health commission. http://www.xinhuanet.com/local/2019-12/31/c_1125409031.htm.

[41] Lian X, Huang J, Huang R, Liu C, Wang L, Zhang T. (2020) Impact of city lockdown on the air quality of COVID-19-hit of Wuhan city. Sci Total Environ., 742:140556.

[42] Mizumoto K, Kagaya K, Zarebski A, Chowell G. (2020) Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance., 25(10):2-6.

[43] Iacobucci G. (2020) Covid-19: UK lockdown is "crucial" to saving lives, say doctors and scientists. BMJ., 368.

[44] Lian X, Huang J, Huang R, Liu C, Wang L, Zhang T. (2020) Impact of city lockdown on the air quality of COVID-19-hit of Wuhan city. Sci Total Environ., 742:140556.

[45] National Health Commission of the People's Republic of China. (2022) Notification of epidemic situation. http://www.nhc.gov.cn/xcs/yqtb/list_gzbd_2.shtml.

[46] Hou C, Chen J, Zhou Y, Hua L, Yuan J, He S, et al. (2020) The effectiveness of quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): A well-mixed SEIR model analysis. J Med Virol., 92(7):841-848.

[47] Li Q, Guan XH, Wu P, Wang XY, Zhou L, Tong YQ, et al. (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med., 382(13):1199-207.

[48] Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. (2003) Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. Science., 300(5627):1966-70.