A Comparative NLP-Based Study on the Current Trends and Future Directions in COVID-19 Research

PRIYANKAR BOSE[1], **SATYAKI ROY**[2], **AND PREETAM GHOSH**[1]

[1]Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA
[2]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27515, USA

Corresponding author: Priyankar Bose (bosep@vcu.edu)

ABSTRACT COVID-19 is a global health crisis that has altered human life and still promises to create ripples of death and destruction in its wake. The sea of scientific literature published over a short time-span to understand and mitigate this global phenomenon necessitates concerted efforts to organize our findings and focus on the unexplored facets of the disease. In this work, we applied natural language processing (NLP) based approaches on scientific literature published on COVID-19 to infer significant keywords that have contributed to our social, economic, demographic, psychological, epidemiological, clinical, and medical understanding of this pandemic. We identify key terms appearing in COVID literature that vary in representation when compared to other virus-borne diseases such as MERS, Ebola, and Influenza. We also identify countries, topics, and research articles that demonstrate that the scientific community is still reacting to the short-term threats such as transmissibility, health risks, treatment plans, and public policies, underpinning the need for collective international efforts towards long-term immunization and drug-related challenges. Furthermore, our study highlights several long-term research directions that are urgently needed for COVID-19 such as: global collaboration to create international open-access data repositories, policymaking to curb future outbreaks, psychological repercussions of COVID-19, vaccine development for SARS-CoV-2 variants and their long-term efficacy studies, and mental health issues in both children and elderly.

INDEX TERMS COVID-19, natural language processing, coefficient of variation, mean squared error.

I. INTRODUCTION

Many virus-borne diseases like Ebola, Influenza, and now COVID, have threatened mankind. Amongst these, variants of the coronavirus have caused global pandemics, such as MERS, SARS, and COVID-19, by mainly manifesting as respiratory infections in humans [1]–[3]. The coronavirus is an RNA virus that pushed the world to a socio-cultural and economic standstill. It has also inspired scientific research in many different domains beyond the realm of medicine. Understandably, scientific literature on COVID-19 skyrocketed after January 2020 and keeping track has become a challenge, especially when the body of literature on existing diseases like Influenza, Ebola, MERS is still growing. Over the past two decades, more than 35 million articles on virus-borne diseases have been published [4]. A comparative

analyses on how the scientific practices and findings differ for COVID-19 with respect to other virus-borne diseases may shed light on the similarities and dissimilarities between these virus-borne diseases, and identify top strengths and shortcomings in clinical methods, practices, and treatment besides also innovations in public policies, for each disease. Thus analyzing textual information of published articles in these areas can highlight research advances in virus-borne diseases in general, and COVID-19 in particular, and inform future directions on scientific research, clinical trials, course of treatment, socioeconomic implications and administrative decision-making.

A. BACKGROUND

Enormous efforts have gone into COVID-19 research within a short span of time. Research works on COVID-19 symptoms [5], [6] and screening [7], [8] were particularly conducted, including techniques like telephone based screening.

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

COVID-19 testing [9], [10] and its spread [11], [12] have also been a very popular research direction. Medications for COVID-19 [12], [13] have been addressed in some works which can be helpful in drug discovery. COVID-19 vaccination has emerged as a very important topic where substantial research has been conducted [14], [15]. A plethora of research papers discuss the impact of the pandemic on mental health [16]–[20]. The socioeconomic, political and cultural aspects of COVID-19 outbreak [21], [22] have also been extensively covered in prior research; these aspects were analyzed in [11], [23]–[25] to study the effects of lockdown, social distancing and US job losses in different sectors. Specifically, in [24], the authors used topic modeling to analyze the effect of COVID-19 in the job sector. Artificial intelligence and machine learning approaches for building predictive models on COVID-19 data were reviewed in [26], [27], while machine learning and deep learning methods on noisy chest X-ray images [28] were employed to detect COVID-19 in [29]–[31]; data augmentation techniques for both X-ray and CT image data [32] were also investigated [31].

There have also been some recent efforts in textual analysis on COVID-19 related research articles. Analysis of bibliometric aspects of the studies and a scoping analysis led to the identification of the main safety-related topics in COVID-19 [33]. Similarly, bibliometric analysis on COVID-19 related articles was accomplished using text mining approaches [34]–[36]. Textual analysis of social media data has the potential to study public perceptions, attitudes and trends related to COVID-19 [42]–[44]. The topics, key terms, and features of the COVID-19 tweets were also analyzed using Topic Modeling, UMAP, and DiGraphs [37]. Text mining techniques have been extremely popular for exploratory analysis on coronavirus-related research. Although a considerable amount of such exploratory analysis is only related to COVID-19, a number of research works have also studied the coronavirus in general. A bibliometric analysis of 395 journal articles from the field of social sciences related to coronavirus was performed by using the 'biblioshiny' package [38]. Text mining techniques like topic modeling using LDA have been widely used to extract the research hotspots and other related information from the articles on different coronavirus-related diseases such as COVID-19, SARS, and MERS, etc. [40], [41]. Also, interesting recurring patterns have been identified by using scientometric comparisons across various coronavirus literature [39]. The different areas of related works and the key-takeaway from them are shown in Table 1. In these works, the authors have considered the bibliometric data of articles to perform their analysis.

In this paper, however, we compare the abstracts on other virus-borne diseases like MERS, Ebola, and Influenza against that of COVID-19 related studies for the first time. The objective of this work is to report and analyze the implications of the semantic and statistically significant words in COVID-19 research compared to those on MERS, Ebola and Influenza.

TABLE 1. Summary of the different facets of COVID-19 research.

Area	Dataset	Methods	Key Takeaways
COVID-19 detection [29]–[31]	Combined COVID-19 Dataset [32] and Noisy COVID-19 X-ray Dataset [28]	Machine Learning and Deep Learning models	LWL-SOM performed better than the current machine learning models on this data
Bibiliometric analysis of COVID-19 [33]–[36]	COVID-19 bibliometric data	Packages like VOSviewer, Bibliometrix and R software	Identified some research themes but unable to address other themes like the long-term impact
Textual analysis of COVID-19 tweets [37]	COVID-19 tweet data	Topic Modeling, UMAP, and DiGraphs	Identified the topics, key terms, and features of COVID-19 tweets
Comparative analysis of COVID-19 research with respect to other coronavirus research [33], [38], [39]	SARS, MERS and COVID-19 publications dataset	Packages like VOSviewer, biblioshiny and R software	Identified the influential aspects of different areas in coronavirus research
Topic Modeling: COVID-19 research with respect to other coronavirus research [40], [41]	SARS, MERS, COVID-19 and other CoV publications (CORD-19) dataset	Latent Dirichlet Allocation	Identified inter-relationships between various coronavirus research to address knowledge gaps

Our goal is to identify the research gaps in COVID-19 that can motivate future research. We discuss the salient contributions of this study next.

B. CONTRIBUTIONS

We apply natural language processing to carry out a comprehensive textual analysis on COVID-19 publications. We infer keywords, topics, countries, and research articles, etc., which yield insights into the present challenges of COVID research from a social, economic, demographic, psychological, epidemiological, clinical, and medical standpoint. A schematic diagram of our workflow is illustrated in Figure 1. The contributions of the research are summarized as follows:

- Using the concept of *coefficient of variation*, we report the significant words present in the abstracts of the articles on three diseases apart from COVID-19, namely, MERS, Ebola, and Influenza. We identify the common, under-expressed, and over-expressed words in the abstracts of the COVID-19 articles with respect to those of Influenza, MERS, and Ebola.
- We quantify the similarity between each pair of diseases (in terms of *mean squared error* as a measure) and discuss the possible context behind the significant words
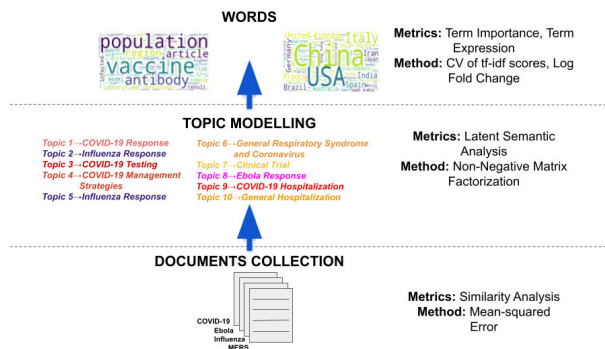
FIGURE 1. Summary of contributions: Each document consists of the abstracts of scientific literature of each disease used to analyze significant over- and under-expressed words, topics, countries and research articles.

identified in this study. We identify relevant countries, topics, and research articles that show the timeliness of COVID-19 research as well as their limitations.

- Our analysis throws up keywords such as *healthcare*, *clinical*, *risk*, *morbidity* that suggest that the scientific community are responding to the immediate existential threats related to transmissibility, health risks, and clinical care and are not yet invested on the long-term immunization and drug-related solutions. While research on medications related to COVID-19 emerges as a popular subject of interest, our study underscores the importance of diagnostics, containment, and short-term treatment plans and the need to broaden the scope of exploration.
- We report the names of the top countries, like China, USA, Italy, UK, and India, that come up in COVID-related research abstracts and show how mentions of such country names in the scientific literature have evolved over time.
- We also report the top topics in the research abstracts on COVID-19, MERS, Ebola, and Influenza based on their association with the top keywords.

II. APPROACH

In this section, we outline the different statistical and natural language processing methods applied during the textual analysis of the scientific literature on virus-borne diseases. We depict in Figure 1 that our analysis involves the identification of significant words in COVID-19 literature using metrics like coefficient of variation (CV), fold change, etc. We then apply latent semantic analysis to identify topics comprising these words. Finally, we study the similarities in the scientific literature of MERS, Ebola, Influenza and COVID-19.

A. DATA COLLECTION

The data was collected from Pubmed [45] which is an open-access search engine by the National Center for Biotechnology Information (NCBI) to retrieve the MEDLINE database of abstracts and citations on biological and

biomedical topics. For each of the four virus-borne diseases like COVID-19, Influenza, MERS, and Ebola, we collected the abstract data of the research papers in PubMed under ten different topics that address the genetic, epidemiological, clinical, psychological, socioeconomic, and public policy related questions about the corresponding disease. For example, in the case of COVID-19, the ten topics under which we have collected the abstract data are (1) COVID-19, (2) COVID-19 policy, (3) COVID-19 vaccine, (4) COVID-19 epidemiology, (5) COVID-19 genomics, (6) COVID-19 comorbidity, (7) COVID-19 outcome, (8) COVID-19 economy, (9) COVID-19 factors and (10) COVID-19 society. We restricted ourselves to the top $10,000$ abstracts for each of the topics. For topics not meeting this threshold, all the available instances were considered. In order to facilitate the comparative analysis, we created an integrated dataset (1) constituting the literature of the diseases and (2) COVID vs non-COVID abstract data.

B. DATA PREPROCESSING

The textual PubMed abstract data from these diseases contain a lot of noise consisting of bibliographic details. Irrelevant information like journal identification number, author information, and non-English words were directly removed from the text using regular expressions. The stopwords were eliminated using the gensim platform [46] in Python. Finally, *lemmatization* was performed to group inflected forms of the same words so they could be analyzed as a single item.

C. OVER- AND UNDER-EXPRESSED WORDS

We estimated the over- and under-expressed words (defined in Sec. II-C) in the COVID-19 abstracts in comparison to the research abstracts of the other diseases. Note that the expression of a word is a relative measure of its occurrence. This analysis was achieved in the following four steps:

- **Important words in each disease document:** The 250 most frequent words were selected for each disease from the preprocessed data on each disease. The 250 most frequent words were chosen because they provide a comprehensive representation of the popular words in the dataset. Frequency distribution of the top 400 words in the research abstracts of the four diseases, is shown in Figure 2 and depicts a very steep decline in the frequencies till the top 50 words and then, a gradual decline in the frequencies till the top 150 words. After that, the decline in frequency is negligible and the curve starts to flatten. Hence, we went ahead with the top 250 words as it allowed us to consider 100 more words after the curve starts to flatten; after 250 words, the curve straightens further. The range of the frequencies of the top 250 words in each document is in a similar range for COVID-19 and Influenza, and MERS to some extent too but in the case of Ebola, the range of the frequencies is quite smaller. The frequency of the words varied roughly between more than 100K and a few thousand
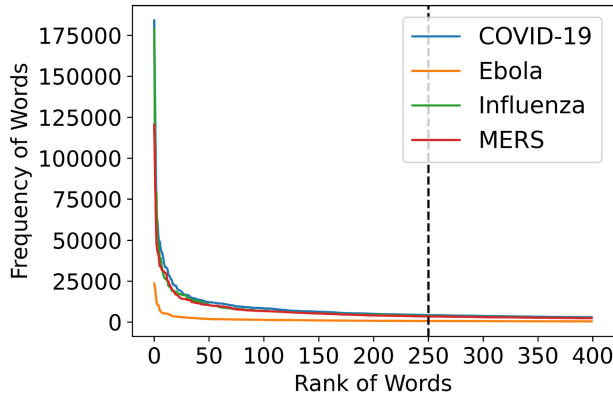
FIGURE 2. Frequency of the top 400 words in the Research Abstracts.

for COVID-19, Influenza, and MERS whereas it varied between a little more than 23K and a few hundreds in case of Ebola.

- **Significant common Words:**
  We use latent semantic analysis to identify significant words across all diseases, generated from the top 250 most frequent words of each disease. The relative importance of common words is measured in terms of the *tf-idf* score. Specifically, term frequency-inverse document frequency (*tf-idf*) is a statistical measure for calculating the importance of a term to a document in a collection of documents [47]. Term frequency (*tf*) of a word t is generally calculated by taking the ratio of the number of times a word appears in a document d (denoted by $f_{t,d}$) to the sum of the frequency of all the terms in d. Inverse document frequency (*idf*) quantifies how rare a word is in the corpus of documents D, by calculating the logarithm of the ratio of the total number of the documents and the number of documents that contain the word. The *tf-idf* score is calculated as:

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \epsilon d} f_{t',d}} \quad (1)$$

$$idf(t, d, D) = log \frac{N}{|d \epsilon D : t \epsilon d|} \quad (2)$$

$$tf - idf(t, d, D) = tf(t, d) \times idf(t, D) \quad (3)$$

Following this, we measure the coefficient of variation (CV) of each word as the ratio of the standard deviation (σ) to the mean (μ) of the *tf-idf* scores of these words across all documents (i.e., $\frac{\sigma}{\mu}$). Thus, low CV indicates a lower standard deviation and higher mean. In other words, it is a combination of a higher mean occurrence of a word and a low deviation from this mean across documents. Note that there exist words with equal *tf-idf* across the four diseases. Although mathematically their CV is infinite, we have considered these words to have $CV = 0$. Since CV cannot be negative as both mean and standard deviation are positive values, the lowest value of CV is 0. In our context, the overall importance of a term is inversely proportional to its CV i.e., the lower the CV, the higher is the importance of a word.

- **Similarity across diseases:** To understand how the scientific literature vary across any pair of diseases i and j, we calculate the mean squared error $MSE(L_i, L_j)$. Here L_i and L_j are the vectors containing *tf-idf* scores of the common words (refer Sec. II-C) for diseases i and j, respectively, arranged in the lexicographical order.

- **Over-expressed and under-expressed words:** The concept of log fold change is used to measure the quantitative change in an observed phenomenon in a given scenario with respect to a control scenario [48]. For any word w, $LR(w) > 0$ and $LR(w) < 0$ represent over- and under-expressed words, capturing the extent of relative increase or decrease in the occurrence of a word in COVID-19 literature vis-à-vis all the four diseases considered in this study. Given the frequency of word w, $h_d(w)$ in disease type d, it is measured as:

$$LR(w) = \log_2 \frac{h_{covid}(w) + 1}{\frac{1}{d} \times \sum_d (h_d(w) + 1)} \quad (4)$$

D. RELEVANT TOPICS

Let us consider a matrix of n documents and m significant common words $[X]_{n*m}$. We apply non-negative matrix factorization to identify two sub-matrices $[W]_{n*d}$ and $[H]_{d*m}$, i.e., $[X]_{n*m} \sim [W]_{n*l} \times [H]_{l*m}$, where l denotes a latent factor. Non-negative matrix factorization has been a prominent technique for information retrieval with huge applications in computational neuroscience, multidimensional data analysis, etc. [49]. The W and H matrices acquired by this decomposition allow us to utilize the latent factors (as topics) to capture the relationship between any pair of documents as well as the words. Also, we infer the semantic significance of a topic by analyzing the words with the highest weights in that topic.

E. TOP RESEARCH COUNTRIES AND ARTICLES

We termed the countries registering the highest occurrence across COVID-19 abstracts as top countries featuring in COVID-19-related research. Identification of top countries involved pruning the COVID-19 abstracts with variants in the country names. For instance, we encountered *United Kingdom*, *UK*, *England* – all of which refer to England. Similarly, the significance of a research article in the COVID-19 document is calculated in terms of the presence of the most important common words in the title or contents of the abstracts.

III. RESULTS

The results follow the same organization as the schematic depicted in Figure 1. We first analyze significant words (including countries) in terms of over- or under-expression in COVID-19 literature. Following this, we identify key topics and analyze the textual similarity in the literature of the different diseases.

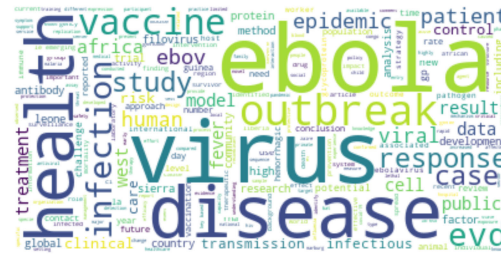**FIGURE 3.** Most frequent words in COVID-19 research.

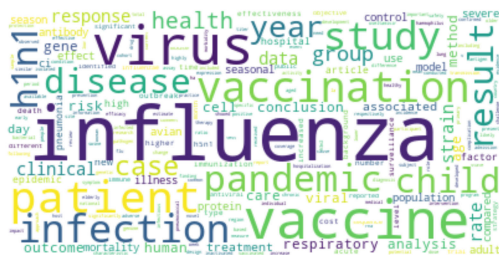**FIGURE 4.** Most frequent words in Ebola research.

A. IMPORTANT WORDS

The data from abstracts in COVID-19 and other viral disease-related research is preprocessed to filter out the noise and focus on the important words. The top 250 words are chosen from the abstracts of each disease category. These important words are analyzed and discussed as follows:

- **COVID-19:** Figure 3 shows the top COVID-19 words, where large font size is commensurate with higher term frequency. The virus responsible for the outbreak, *SARS-CoV-2*, is predictably a top word [6], [50]–[52]. Similarly, *Coronavirus*, *respiratory*, *severe* and *acute* are also top-ranked words [53], [54]. Other words like *pandemic*, *transmission*, *outbreak, etc* have often been mentioned [7], [21], [53], [55]. In addition, the city and country of origin of the virus – *Wuhan* and *China* – have come up extensively [51], [53]. Interestingly, studies have been carried out to understand the association between *cancer* and COVID-19 [56]. Such studies deal with cancer in children [57], and breast cancer [58], among others. Apart from cancer, comorbidities like *diabetes* and *cardiovascular* diseases as well as fever as a symptom of COVID have been mentioned [59], [60]. Studies have been also carried out on COVID-19 *vaccine*s and people's reaction to it [15]. Mental health is another popular topic, since COVID-19 has had serious psychological consequences. This is well reflected in some recent research works that addressed the issue of *mental* healthcare and short- and long-term psychological impacts of the pandemic [17], [18], [61]. Interesting topics like mental health issues in the elderly population or in people with comorbidities [20], [61], individual mental health [16], mental health effects due to COVID-19 media coverage [19], etc. have also been addressed recently.

  RNA and *protein* have come up while describing the single-stranded RNA-based COVID-19 virus and the effects of different proteins on it [53]. Studies on pneumonia or COPD (Chronic obstructive pulmonary disease) have been mentioned, making words like *pneumonia*, and *lung* highly relevant [62]. In addition, there are other works assessing the expression patterns and genetic polymorphism of Angiotensin Converting Enzyme 2 or *ACE2* [59], [63] – an enzyme present in the cell membranes in the lungs, arteries, heart, kidney, and

intestines, that is responsible for reducing blood pressure. *ACE2* has been mentioned with regard to a drug to treat cardiovascular diseases and serves as an entry point for many coronaviruses including COVID-19.
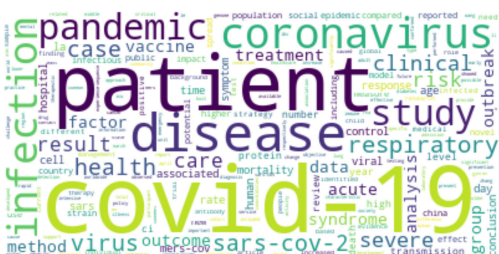
- **Ebola:** *Ebola* is understandably the most significant term, followed by *virus*, and *disease* [64]–[66]. Terms like *vaccine* [67] and *epidemic* [68] are next in line. Interestingly, the term *pandemic* [69] is less used in case of Ebola than other virus-spread diseases. The words like *West*, *Africa*, *Seirra Leone*, *Guinea*, *Liberia* are mentioned in the context of the first community transmission of the Ebola epidemic in West African countries. Similarly, *Congo*, a place that witnessed Ebola outbreak in a pre-vaccine period, has also been mentioned [67].

  Terms *RNA*, *protein*, *VP35* (viral protein 35) and *VP40* (viral protein 40) occur in the abstracts of Ebola research articles, where types of Ebola virus and their medical pathogenesis are discussed [70]–[72]. The symptoms of the disease is widely discussed through words like *hemorrhagic*, *fever* and *bats* [66], [68]. *Ebola* and *Marburg* are a type of *filovirus* and hence these terms appear on Ebola research articles where filovirus were mentioned [65], [73], [74]. Modified Vaccinia virus Ankara (MVA) is mentioned in the context of vaccine development to combat diseases like *Influenza*, *COVID-19*, *HIV*, *malaria*, *Influenza*, and Ebola [69], [75]. Research on SARS-CoV and SARS-CoV-2 also discuss their severity, as indicated through terms like *ventilation* [76] and Intensive Care Unit (*ICU*) [77], [78]. Most frequent words for Ebola related research abstracts are illustrated in Figure 4.

- **Influenza:** Apart from *Influenza*, the terms *virus*, *vaccine* and *patient* get frequent mentions [79], [80]; the virus that is responsible for causing influenza, *H1N1*, as well as a subtype of Influenza A virus, *H3N2* appears often [81]. There are references to genotyping [82], health implications [83] (such as pregnancy [84]) and the *avian* Influenza A virus i.e *H5N1* [85], [86]. The drug used to treat flu or Influenza is known as *oseltamivir*; this antiviral drug often comes up in research abstracts on Influenza virus [87], [88]. As an after-effect of Influenza, the term *pneumonia* [89] is mentioned, while *Haemophilus* becomes relevant as ample research discusses Haemophilus influenzae [90]. *Haemaglutinin* is

FIGURE 5. Most frequent words in Influenza research.

a membrane glycoprotein on the Influenza virus and quite a few research works talk about this membrane in connection with Influenza [91]. The most frequent words in the Influenza related research abstracts are also depicted in Figure 5.

- **MERS:** Unlike other diseases, the top word in case of MERS is *COVID-19*. Note that the words next in order are *coronavirus* and *pandemic* [92], [93]. Likewise, words like *SARS* and *SARS-CoV* have risen into prominence in the context of MERS research [94]–[96]. *Diabetes* is enlisted among meaningful words, since many studies talk about the effect of diabetes on a coronavirus (SARS, MERS or COVID-19) infected person [97].

  To our surprise, the terms like *MERS* and *MERS-CoV* appear after COVID-19 and SARS. This is because several research articles on MERS have also dealt with other coronaviruses [98], [99], whereas only a few are solely dedicated to MERS [100]. *Cancer* once again appears frequently due to the ample research on cancer patients suffering from the disease [101]. The term *IBV*, short for Infectious Bronchitis Virus [102], [103] come up, while a lot of research on Porcine Epidemic Diarrhea Virus (*PEDV*) [104], [105] is present in the MERS research abstracts. Once again, the most frequent words in the MERS research abstracts are shown in Figure 6.

Overall, there is a greater emphasis on the course of transmissibility, and health implications for patients with preexisting conditions in COVID-19 literature. For the other diseases, we observe higher mentions of the treatment measures and their evolution in terms of geographical epidemiology. This shows that we are still in the nascent stages of COVID-19

research with a limited understanding of treatment and mitigation strategies.

B. COMPARISON OF TERM EXPRESSION ACROSS DISEASES

In Section II-C, we define over- and under-expressed words as those that exhibit a relative increase and decrease in COVID literature as compared to all the four diseases considered in this study. Figures 8 and 9 visualize the top 25 words showing the highest variation in expression (i.e., log fold change measuring under-expression). It is worth noting the occurrences of Influenza viruses like *H3N2*, *H5N1*, *H1N1*, which have been discussed widely in the context of all virus borne diseases. Similarly, we find the mentions of places like Guinea, Liberia, and Sierra Leone that have suffered large-scale outbreaks of Ebola commonly manifested in flu-like symptoms. Considering the early stages of research and distribution on immunization, the term *vaccination* is relatively underrepresented in COVID literature.

Similarly, Figures 10 and 11 show the over-expressed words in COVID-19 abstracts, and the distribution of the first 25 over-expressed words by *LR* score. Understandably, the most over-expressed words (also flagged as important in Sec. III-G) are *ACE2*, *SARS-CoV-2* and *lockdown*. The relationship between susceptibility to COVID-19 and preexisting conditions is evident where terms such as *cancer*, *diabetes*, *cardiovascular* and *comorbidities* are over-expressed. Observe a more even distribution of the top over-expressed words as compared to the under-represented counterparts. Similarly, we have the list of under-expressed words in Fig. 9. We discuss them in more detail in Section III-G.

FIGURE 7. Significant common words across all diseases.

FIGURE 6. Most frequent words in MERS research.

FIGURE 8. Top under-expressed words based on log fold change.
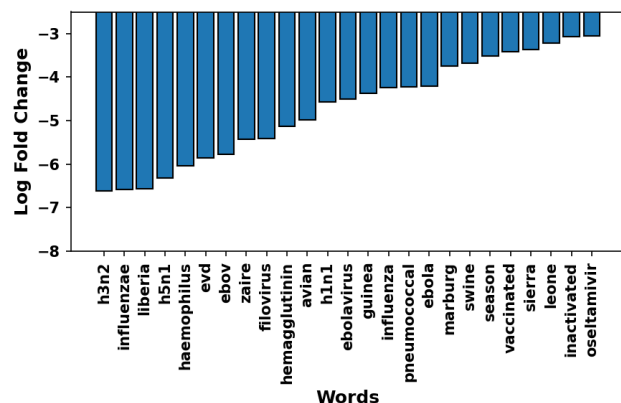
FIGURE 9. Distribution of the first 25 under-expressed words by the log ratio (*LR*).

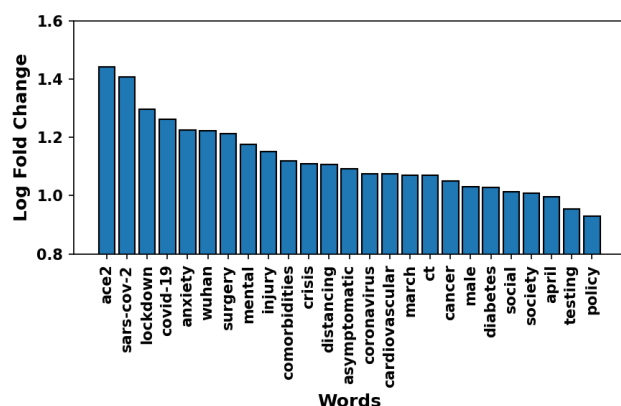**FIGURE 10.** Top over-expressed words based on log fold change.

FIGURE 11. Distribution of the first 25 over-expressed words by the log ratio (*LR*).
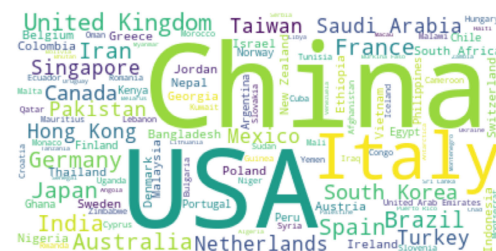

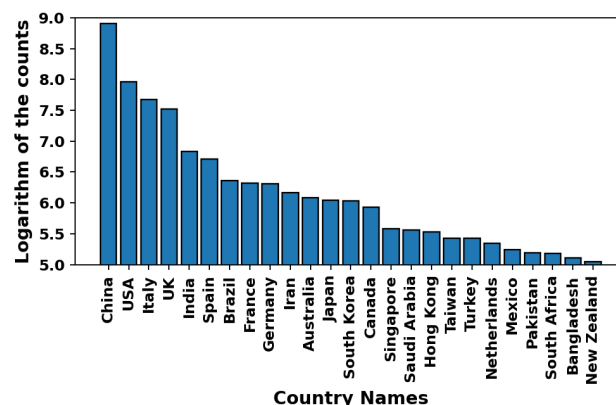**FIGURE 12.** Names of the most mentioned countries in COVID abstracts.

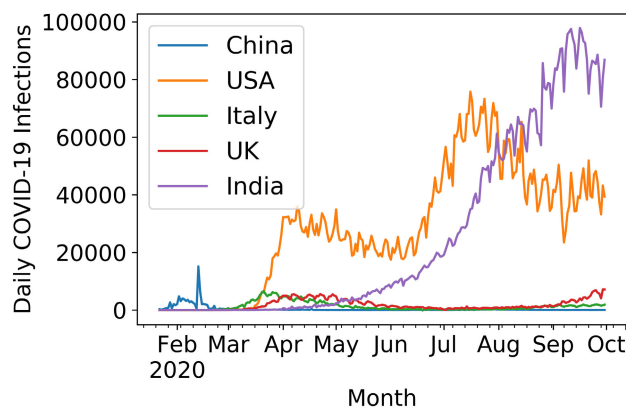**FIGURE 13.** Frequency of 25 most mentioned countries in COVID research.

FIGURE 14. Daily COVID-19 infections in the 5 countries mentioned highly in COVID-19 research abstracts.

C. SIGNIFICANT COUNTRIES

A significant country is one that appears (as a region where the scientific study on COVID-19 is focused) in abstracts with a high number of significant common words. Figures 12 and 13 show the country names pictorially and log frequency of the words associated with each country, where China, USA, Italy, UK, and India are ranked in the decreasing order of importance. Wuhan, China has been notoriously linked as the birthplace of and received attention in COVID literature [51], [53], [106]. We show in Figure 14 that China hit an early peak by February-March 2020, before

other nations (US, Italy, UK) were hard hit by the pandemic. We see a similar peak in the mention of China in COVID-19 literature (Figure 15). From early April 2020, we observe a disproportionately high amount of research within the USA and Italy, while India became a hotspot of COVID research from August onwards. Note that clinical and medical data from Italy and US have been reported in several research articles [107]–[111].

We next depict the relationship between the top countries and top words. Figure 16 shows the number of mentions of a country alongside the top words on a log scale. China, USA and Italy exhibit the highest presence for the top 5 words, followed by UK and India. From the most
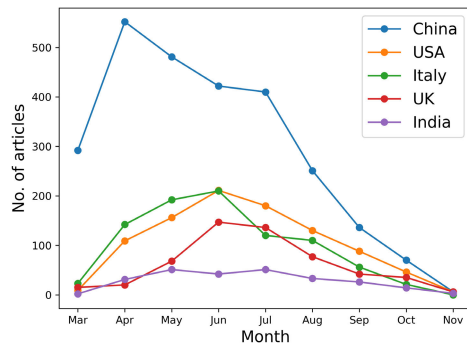
FIGURE 15. Evolution of the frequency of occurrence of the most mentioned countries in COVID-19 research abstracts.
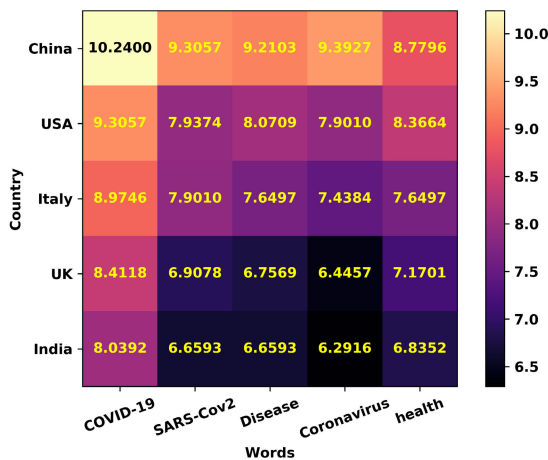

FIGURE 16. Counts (on a log scale) of the top 5 keywords co-occurring with the top 5 nations mentioned in COVID-19 research abstracts.

mentioned nations, we turn our attention to the least mentioned countries in scientific literature, namely, Andorra, Fiji, Papua New Guinea, New Caledonia, Mayotte, Belize, Bermuda, Kyrgyzstan, and the Northern Mariana Islands (see Figure 17). Andorra comes up rarely in October 2020 in the research [112] in the context of the patient characteristics, ICU mortality factors, and the clinical course of COVID-19 in Spain. Papua New Guinea, Fiji, Northern Mariana Islands, and New Caledonia show up together in March 2020 in [113], evaluating the risks of COVID-19 importation to the Pacific islands.

Mayotte [114] and Vanuatu [115] come up during September - November 2020 and September 2020 respectively, while Bermuda's cancer care in the COVID era was discussed in [116]. Kyrgyzstan has been discussed at length in genomic research on SARS-CoV-2 in August 2020. Figure 17 shows the daily infection count of the aforementioned nations. These daily COVID-19 infection numbers are exceptionally low (i.e., in the order of hundreds) compared to those from the countries receiving frequent mentions, explaining why they have received less attention from the scientific community.
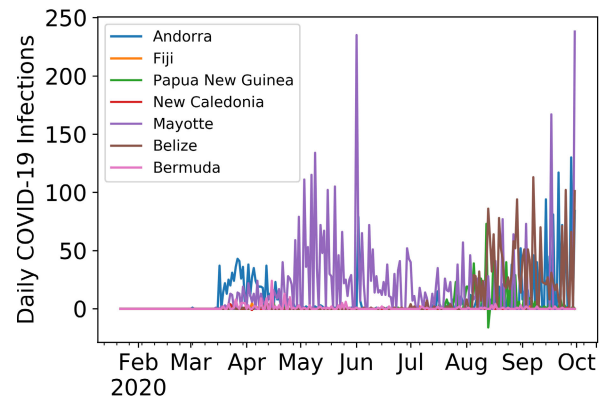
FIGURE 17. Evolution of daily COVID-19 infection cases in the least mentioned countries.

D. IDENTIFICATION OF RELEVANT TOPICS

We apply latent semantic analysis (see Sec. II-C) to identify $d = 10$ topics. We attempted to report some general topics from these four documents and hence, only enlist the top 10 topics; the subsequent topics beyond the top-10 did not represent meaningful and distinctly new directions. We explain in Sec. II-D that each topic is represented as a vector of weights corresponding to the common words. The broad semantic literal meaning can be inferred by analyzing the top 20 words in any given topic, as reported below:

- *Topic 1* → COVID-19 Response (top words: COVID-19, Respiratory, Coronavirus, Pandemic)
- *Topic 2* → Influenza Response (top words: Influenza, Pandemic, Vaccine)
- *Topic 3* → COVID-19 Testing (COVID-19, SARS-CoV-2, Syndrome, Acute, Severe, Positive)
- *Topic 4* → COVID-19 Management Strategies (top words: COVID-19, SARS-CoV-2, Wuhan, Management, Strategy)
- *Topic 5* →> Influenza Response (Influenza, H1N1, Vaccine, H5N1)
- *Topic 6* → General Respiratory Syndrome and Coronavirus (top words: Respiratory, MERS-CoV, SARS-CoV, Coronaviruses)
- *Topic 7* → Clinical Trial (top words: Admission, Intervention, Dose, Trial)
- *Topic 8* → Ebola Response (top words: Ebola, Epidemic, EboV, Africa)
- *Topic 9* → COVID-19 Hospitalization (top words: COVID-19, SARS-CoV-2, Surgery, Hospitalization, Admitted)
- *Topic 10* → General Hospitalization (top words: Admission, Mortality, Therapy)

Recall that this semantic analysis creates a matrix (W) with topic weights contributing to each document. We depict these weights in Figure 18. We observe that COVID-19 research is dominated by topics 1, 3, 4, and 9 dealing with COVID-19 Response, COVID-19 Testing, COVID-19 Management Strategies, and COVID-19 Hospitalization,
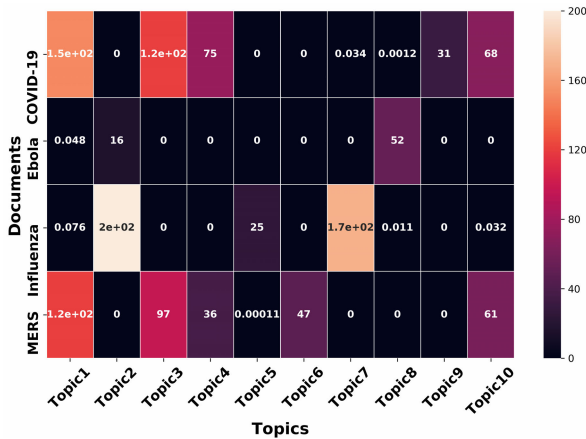
FIGURE 18. Relationship between disease documents and topics.

FIGURE 19. Similarity across disease literature based on mean squared error.

respectively. We report in Sec. III-E, that MERS shows significant overlap with COVID-19. MERS literature covers general coronavirus, general epidemics, or comparison between virus-borne diseases. On the other hand, topics 2 and 5 contribute heavily towards Influenza, and topic 8 features in Ebola-related publications. Finally topic 7, despite not being disease-specific, has a high weight in the Influenza document. To sum up, 7 out of the 10 topics are mostly seen in COVID-19, Influenza, and Ebola research, while the other 3 cover all four diseases.

E. SIMILARITY ACROSS RESEARCH ON VARIOUS DISEASES
The 250 keywords from the research abstracts on each disease are combined into a list of 403 unique words. Next, we apply the coefficient of variation on the tf-idf scores to rank the words in the integrated list. Recall from our discussion in Sec. II-C that the words with the least CV are the most significant (depicted in Figure 7). We observe that words like *population*, *risk*, *model*, *sample*, *case*, *region*, *vaccine*, *antibody*, *respiratory* and *mortality*, emerge as significant.

We calculate the mean squared error (MSE) of the tf-idf vectors of 403 common words of the disease pairs and represent the variation among diseases in the form of a heatmap (see Figure 19). Note that MERS and COVID-19 articles are the most similar with MSE 9.1×10^{-7}, while MERS and Ebola are the most disparate with MSE of 1.9×10^{-5}. This similarity study reveals the mention of COVID-19 and MERS in the same studies. For example, [14] mentions MERS while talking about the vaccine development strategies for COVID-19 and [117] provides insights into COVID-19 or SARS-CoV-2 in the light of past coronavirus outbreaks like SARS and MERS. Another example of co-occurrence of disease names in the same study is in [67], that showed that H84T has demonstrated antiviral activity against Influenza A and B viruses and can be effective against Ebola. Note that the MSE values are small (of the order of 10^{-5}). This
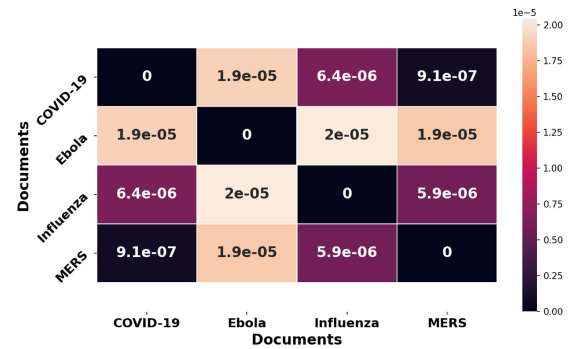
is because studies address multiple diseases at the same time [118], [119].

F. INFLUENTIAL RESEARCH ARTICLES
We have gathered the most influential research articles regarding COVID-19 (as discussed in Section II-E). We discuss the first 10 influential research articles as the total number of significant words in the top 10 research articles ranges between 70 and 80. In [120], the success and safety of oral administration of the kinase inhibitor drug was assessed on COVID-19 hospitalized patients. Similarly, the efficacy of antibody test in the diagnosis of COVID-19 in individuals with symptoms for over two weeks and not having an RT-PCR test or having a negative RT-PCR test was reported in [9]. The role of physical intervention in containing spread was analyzed in [121], while [12] performed a controlled trial to gauge the effectiveness of the antiviral activity of hydroxychloroquine against SARS-CoV-2. The ill-effects of tobacco on lung health and respiratory diseases were discussed in [122], and [13] studied the safety of convalescent plasma or hyperimmune immunoglobulin transfusion in the treatment of COVID-19 infected people. Next, an expository article was presented in [61] to help mental healthcare professionals understand the effect of COVID-19 on psychiatric patients, and [10] aimed to judge the diagnostic correctness of point-of-care antigen and molecular-based tests to ascertain whether a person in community, primary or secondary healthcare is COVID-19 infected or not. The evaluation and treatment of coronavirus disease like SARS-CoV, MERS-CoV, etc. was discussed in [123]. Finally, as in [13], [124] evaluated convalescent plasma or hyperimmune immunoglobulin transfusion in treating COVID-19 patients. Since we are still in the early phases of understanding SARS-CoV-2, the majority of the influential articles focus squarely on standard approaches of diagnostics, containment, and treatment.

G. IMPLICATIONS OF THE STUDY
As mentioned in Sec. I-B, the primary goal of our study is to find out the important long-term research topics on

COVID-19 in light of the studies on other virus-borne diseases. To assess this, we collected the abstract data on the *long-term impacts* of the respective diseases. Under this search technique, we found 52, 347 and 562 abstracts respectively pertaining to Ebola, Influenza and MERS. This data gives us an idea about the long-term research directions on the other virus-borne diseases. Such works may then imply the probable long-term research directions on COVID-19. We manually identified the dominant research directions from the top-10 topics (that were inferred using latent semantic analysis as before) pertaining to the abstract data on the other diseases (Ebola, Influenza and MERS) which are as follows:

- *Long-term Patient Surveys*
- *Future Disease Outbreak and its Long-term Effects*
- *Intervention Studies*
- *Mental and Other Disease Outcomes*
- *Long-term Vaccination Policies and its Effect on Children and Elderly*
- *Long-term Immunity and its Effect on Patients and Animals*

We next identified the top research directions from the data on the long-term implications of MERS alone as the COVID-19 abstracts were found to be most similar to the MERS abstracts. The broad semantic research areas manually mined from the MERS related abstracts are:

- *Long-term Therapies and Effect on Children*
- *Long-term Patient Surveys*
- *Intervention Studies*
- *Mental and Other Disease Outcomes*

Each of the four long-term research directions align well with our findings mentioned below considering the abstracts from general virus-borne diseases. Mental health, Long-term patient surveys and Intervention studies have emerged as long-term research directions in case of both MERS and the other three virus borne diseases.

First, this study highlights the need for global collaboration to create an international open-access repository that will guide the clinicians, health workers, patients and world leaders to achieve the highest level of patient care. Second, several infectious disease experts, virologists and immunologists agree that COVID-19 will linger as an endemic [125]. Our study shows the merits of designing strategies to curb future outbreaks. Third, our study reveals that the long-term psychological repercussions of COVID-19 surpasses other threats stemming from the virus [126] (see Sec. III-G), making research on COVID-related mental health an absolute imperative. Finally, Fig. 9 shows vaccination to be an under-expressed term in COVID abstracts. This implies that more research will need to go into designing new vaccines to combat new strains of the virus and protecting the elderly, children or patients with preexisting conditions as well as monitoring their effects on overall immunity.

IV. LIMITATIONS OF THE STUDY

Our study is based on the research article data collected from PubMed (see Sec. II-A for details). We restricted ourselves to the first 10,000 entries of each search topic allowed by PubMed. We compensated for this shortcoming by attempting multiple search terms, resulting in minor redundancy in search results. Some results were incomplete with respect to title, author details, publication dates, and contents however, their number was too limited to skew the trends in our findings. Finally, the fact that there is a notable similarity in some disease keywords added to the challenge of sifting through the welter of research articles. Specifically, our initial efforts on exploring SARS as a standalone document in the corpus was rendered impossible as it also picked up SARS-CoV-2 – the virus causing COVID-19.

Research Gaps in COVID-19: Our Topic Modeling framework can model some very broad topics related to these four diseases. We identified four topics related to COVID-19 pandemic which were too broad and were unable to identify any topic on social, cultural, economic and psychological impacts of COVID-19. The bulk of research on COVID-19 are on its viral transmission, different medications, clinical trials, etc. but the amount of research on COVID-19 diagnostics, therapeutics, vaccine and genomics are comparatively smaller. Since, the number of COVID-19 infections is still increasing around the globe, there is an urgent need for further research on COVID-19 vaccines, therapeutic measures and diagnostics. The community will also benefit significantly from additional investigations on probable future disease outbreaks. Intervention studies and patient surveys is another important future research topic on COVID-19. Although, there have been some research on mental illness with respect to COVID-19, most of these works fail to bring up the reasons behind this psychological effect.

V. CONCLUSION

We performed a comprehensive natural language processing based analysis on the existing scientific literature on COVID-19 to derive keywords that lend social, economic, demographic, psychological, epidemiological, clinical, and medical insights into our understanding of the disease. In light of three diseases, namely, MERS, Ebola, and Influenza, we identify over- and under-represented keywords in COVID-19 research, significant topics, countries, and research articles, as well as their implications on the future of these nascent fields of COVID-19 research.

We use the notion of coefficient of variation to find statistically and semantically important keywords and utilize it to pinpoint trends in the overall science of COVID-19. The references to *healthcare*, *clinical*, *risk*, *morbidity* suggest that the public at large and the scientific community are responding to the immediate threats posed by the virus, and the emphasis of COVID-19 research is squarely on the transmissibility, health implications, short-term treatment

plans and public policies (as opposed to long-term immunization and drug-related studies). MERS and COVID-19 exhibit a high co-occurrence in research articles. Topics like COVID-19 Response, COVID-19 Testing, COVID-19 Management Strategies, and COVID-19 Hospitalization forms the majority of the topics in the research abstracts on COVID-19, MERS, Ebola, and Influenza, followed by topics related to Influenza, and then Ebola. No dedicated topic related to MERS has been identified due to the high level of co-occurrence between COVID-19 and MERS-related articles. China, USA, Italy, UK, and India – countries hit hard by the pandemic – have contributed most to scientific studies. The majority of the high-impact articles address questions on diagnostics, containment, and immediate treatment. It is imperative that the researchers continue to leverage their substantial knowledge about SARS-CoV-2 to devise long-term vaccination and drug programs. Particularly, China, the US, UK, Italy, and India can spearhead international projects to create repositories on myriad branches of COVID-19 research.

Our work reports a wide spectrum of insights into the progression of COVID-19 and the course of the associated scientific studies. Comparative analysis between the abstract data on COVID-19 and that on other virus-borne epidemics have shown the need for future research on COVID-19 vaccine, genomics, therapeutics, etc. On top of that, further research on probable future outbreaks, patient surveys and intervention studies will definitely be helpful to the community. Also, we know that psychological problems are well associated with COVID-19 but there are not many works addressing its cause. Hence, it is important to identify whether lockdown or increased viral transmission or other factors are responsible for this illness. Time-dependent topic modeling can also be insightful to understand the temporal patterns in the topics, and hence is an avenue for potential future work.

REFERENCES

[1] D. Raoult, A. Zumla, F. Locatelli, G. Ippolito, and G. Kroemer, "Coronavirus infections: Epidemiological, clinical and immunological features and hypotheses," *Cell stress*, vol. 4, no. 4, p. 66, 2020, doi: 10.15698/cst2020.04.216.

[2] D. Isaacs, D. Flowers, J. R. Clarke, H. B. Valman, and M. R. MacNaughton, "Epidemiology of coronavirus respiratory infections," *Arch. Disease Childhood*, vol. 58, no. 7, pp. 500–503, Jul. 1983, doi: 10.1136/adc.58.7.500.

[3] D. X. Liu, J. Q. Liang, and T. S. Fung, "Human coronavirus-229E, -OC43, -NL63, and -HKU1 (coronaviridae)," in *Encyclopedia of Virology*, 4th ed., D. H. Bamford and M. Zuckerman, Eds. New York, NY, USA: Academic, 2021, pp. 428–440, doi: 10.1016/B978-0-12-809633-8.21501-X.

[4] D. Kagan, J. Moran-Gilad, and M. Fire, "Scientometric trends for coronaviruses and other emerging viral infections," *GigaScience*, vol. 9, no. 8, Aug. 2020, Art. no. giaa085, doi: 10.1093/gigascience/giaa085.

[5] T. Struyf, J. J. Deeks, J. Dinnes, Y. Takwoingi, C. Davenport, M. M. Leeflang, R. Spijker, L. Hooft, D. Emperador, S. Dittrich, J. Domen, S. R. A. Horn, and A. Van den Bruel, "Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease," *Cochrane Database Systematic Rev.*, no. 7, Jul. 2020, Art. no. CD013665, doi: 10.1002/14651858.CD013665.

[6] J. Li, K. C. Thoon, C. Y. Chong, M. Maiwald, K.-Q. Kam, K. Nadua, N. W. Tan, and C. F. Yung, "Comparative analysis of symptomatic and asymptomatic SARS-CoV-2 infection in children," *Ann. Acad. Med., Singap.*, vol. 49, no. 8, pp. 530–537, Aug. 2020.

[7] N. NeJhaddadgar, A. Ziapour, G. Zakkipour, J. Abbas, M. Abolfathi, and M. Shabani, "Effectiveness of telephone-based screening and triage during COVID-19 outbreak in the promoted primary healthcare system: A case study in Ardabil province, Iran," *J. Public Health*, vol 789, pp. 1–6, Nov. 2020, doi: 10.1007/s10389-020-01407-8.

[8] L. Rivett *et al.*, "Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission," *Elife*, vol. 9, May 2020, Art. no. se58728, doi: 10.7554/eLife.58728.

[9] J. J. Deeks, J. Dinnes, Y. Takwoingi, C. Davenport, R. Spijker, S. Taylor-Phillips, A. Adriano, S. Beese, J. Dretzke, L. F. di Ruffano, I. M. Harris, M. J. Price, S. Dittrich, D. Emperador, L. Hooft, M. M. Leeflang, and A. Van den Bruel, "Antibody tests for identification of current and past infection with SARS-CoV-2," *Cochrane Database Systematic Rev.*, no. 6, Jun. 2020, Art. no. CD013652, doi: 10.1002/14651858.CD013652.

[10] J. Dinnes, J. J. Deeks, A. Adriano, S. Berhane, C. Davenport, S. Dittrich, D. Emperador, Y. Takwoingi, J. Cunningham, S. Beese, J. Dretzke, L. F. di Ruffano, I. M. Harris, M. J. Price, S. Taylor-Phillips, L. Hooft, M. M. Leeflang, R. Spijker, and A. Van den Bruel, "Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection," *Cochrane Database Systematic Rev.*, vol. 3, no. 8, p. 412, Aug. 2020, doi: 10.1002/14651858.CD013705.

[11] S. Roy, R. Dutta, and P. Ghosh, "Towards dynamic lockdown strategies controlling pandemic spread under healthcare resource budget," *Appl. Netw. Sci.*, vol. 6, no. 1, pp. 1–15, Dec. 2021, doi: 10.1007/s41109-020-00349-0.

[12] T. Jefferson, R. Foxlee, C. D. Mar, L. Dooley, E. Ferroni, B. Hewak, A. Prabhala, S. Nair, and A. Rivetti, "Physical interventions to interrupt or reduce the spread of respiratory viruses: Systematic review," *Brit. Med. J.*, vol. 336, no. 7635, pp. 77–80, Jan. 2008, doi: 10.1136/bmj.39393.510347.BE.

[13] S. J. Valk, V. Piechotta, K. L. Chai, C. Doree, I. Monsef, E. M. Wood, A. Lamikanra, C. Kimber, Z. McQuilten, C. So-Osman, L. J. Estcourt, and N. Skoetz, "Convalescent plasma or hyperimmune immunoglobulin for people with COVID-19: A rapid review," *Cochrane Database Systematic Rev.*, no. 5, May 2020, Art. no. CD013600, doi: 10.1002/14651858.CD013600.

[14] L. Yang, D. Tian, and W. Liu, "Strategies for vaccine development of COVID-19," *Shengwu Gongcheng Xuebao/Chin. J. Biotechnol.*, vol. 36, no. 4, Apr. 2020, doi: 10.13345/j.cjb.200094.

[15] Z. Su, J. Wen, J. Abbas, D. McDonnell, A. Cheshmehzangi, X. Li, J. Ahmad, S. Šegalo, D. Maestro, and Y. Cai, "A race for a better understanding of COVID-19 vaccine non-adopters," *Brain, Behav., Immunity Health*, vol. 9, Dec. 2020, Art. no. 100159, doi: 10.1016/j.bbih.2020.100159.

[16] J. Abbas, "The impact of coronavirus (SARS-CoV2) epidemic on individuals mental health: The protective measures of Pakistan in managing and sustaining transmissible disease," *Psychiatria Danubina*, vol. 32, nos. 3–4, pp. 472–477, Dec. 2020, doi: 10.24869/psyd.2020.472.

[17] A. Maqsood, J. Abbas, G. Rehman, and R. Mubeen, "The paradigm shift for educational system continuance in the advent of COVID-19 pandemic: Mental health challenges and reflections," *Current Res. Behav. Sci.*, vol. 2, Nov. 2021, Art. no. 100011, doi: 10.1016/j.crbeha.2020.100011.

[18] J. Abbas, "Crisis management, transnational healthcare challenges and opportunities: The intersection of COVID-19 pandemic and global mental health," *Res. Globalization*, vol. 2021, Feb. 2021, Art. no. 100037, doi: 10.1016/j.resglo.2021.100037.

[19] Z. Su, D. McDonnell, J. Wen, M. Kozak, J. Abbas, S. Šegalo, X. Li, J. Ahmad, A. Cheshmehzangi, Y. Cai, L. Yang, and Y.-T. Xiang, "Mental health consequences of COVID-19 media coverage: The need for effective crisis communication practices," *Globalization Health*, vol. 17, no. 1, pp. 1–8, Dec. 2021, doi: 10.1186/s12992-020-00654-4.

[20] K. H. Shuja, M. Aqeel, E. A. Khan, and J. Abbas, "Letter to highlight the effects of isolation on elderly during COVID-19 outbreak," *Int. J. Geriatric Psychiatry*, vol. 35, no. 12, pp. 1–2, 2020, doi: 10.1002/gps.5423.

[21] J. Y. Lebni, J. Abbas, F. Moradi, M. R. Salahshoor, F. Chaboksavar, S. F. Irandoost, N. Nezhaddadgar, and A. Ziapour, "How the COVID-19 pandemic effected economic, social, political, and cultural factors: A lesson from iran," *Int. J. Social Psychiatry*, vol. 2020, Jul. 2020, Art. no. 0020764020939984, doi: 10.1177/0020764020939984.

[22] J. Abbas, R. Mubeen, P. T. Iorember, S. Raza, and G. Mamirkulova, "Exploring the impact of COVID-19 on tourism: Transformational potential and implications for a sustainable recovery of the travel and leisure industry," *Current Res. Behav. Sci.*, vol. 2, Nov. 2021, Art. no. 100033, doi: 10.1016/j.crbeha.2021.100033.

[23] S. Roy and P. Ghosh, "Factors affecting COVID-19 infected and death rates inform lockdown-related policymaking," *PLOS ONE*, vol. 15, no. 10, pp. 1–18, 2020, doi: 10.1371/journal.pone.0241165.

[24] S. Roy, R. Dutta, and P. Ghosh, "Recreational and philanthropic sectors are the worst-hit US industries in the COVID-19 aftermath," *Social Sci. Hum. Open*, vol. 3, no. 1, Jan. 2021, Art. no. 100098, doi: 10.1016/j.ssaho.2020.100098.

[25] S. Roy, A. Cherevko, S. Chakraborty, N. Ghosh, and P. Ghosh, "Leveraging network science for social distancing to curb pandemic spread," *IEEE Access*, vol. 9, pp. 26196–26207, 2021, doi: 10.1109/ACCESS.2021.3058206.

[26] S. Roy, R. Dutta, and P. Ghosh, "Optimal time-varying vaccine allocation amid pandemics with uncertain immunity ratios," *IEEE Access*, vol. 9, pp. 15110–15121, 2021, doi: 10.1109/ACCESS.2021.3053268.

[27] J. Rasheed, A. Jamil, A. A. Hameed, U. Aftab, J. Aftab, S. A. Shah, and D. Draheim, "A survey on artificial intelligence approaches in supporting frontline workers and decision makers for the COVID-19 pandemic," *Chaos, Solitons Fractals*, vol. 141, Dec. 2020, Art. no. 110337, doi: 10.1016/j.chaos.2020.110337.

[28] *Noisy COVID-19 X-ray Dataset*, Mendeley, London, U.K., 2020, doi: 10.17632/fjg5cbzffh.3.

[29] A. H. Osman, H. M. Aljahdali, S. M. Altarrazi, and A. Ahmed, "SOM-LWL method for identification of COVID-19 on chest X-rays," *PLoS ONE*, vol. 16, no. 2, pp. 1–26, 02 2021, doi: 10.1371/journal.pone.0247176.

[30] K. Haque and A. Abdelgawad, "A deep learning approach to detect COVID-19 patients from chest X-ray images," *Artif. Intell.*, vol. 1, no. 3, pp. 418–435, Sep. 2020, doi: 10.3390/ai1030027.

[31] A. Sedik, A. M. Iliyasu, B. A. El-Rahiem, M. E. A. Samea, A. Abdel-Raheem, M. Hammad, J. Peng, F. E. A. El-Samie, and A. A. A. El-Latif, "Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections," *Viruses*, vol. 12, no. 7, p. 769, Jul. 2020, doi: 10.3390/v12070769.

[32] *Combined COVID-19 Dataset*, Mendeley, London, U.K., 2020, doi: 10.17632/3pxjb8knp7.3.

[33] M. Haghani, M. C. J. Bliemer, F. Goerlandt, and J. Li, "The scientific literature on coronaviruses, COVID-19 and its associated safety-related research dimensions: A scientometric analysis and scoping review," *Saf. Sci.*, vol. 129, Sep. 2020, Art. no. 104806, doi: 10.1016/j.ssci.2020.104806.

[34] M. M. Hossain, "Current status of global research on novel coronavirus disease (COVID-19): A bibliometric analysis and knowledge mapping [version 1; peer review: 2 approved with reservations]," *F1000Research*, vol. 9, p. 374, 2020, doi: 10.12688/f1000research.23690.1.

[35] S. Verma and A. Gustafsson, "Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach," *J. Bus. Res.*, vol. 118, pp. 253–261, Sep. 2020, doi: 10.1016/j.jbusres.2020.06.057.

[36] C. Aguado-Cortés and V. M. Castaño, "Translational knowledge map of COVID-19," 2020, *arXiv:2003.10434*. [Online]. Available: https://arxiv.org/abs/2003.10434

[37] C. Ordun, S. Purushotham, and E. Raff, "Exploratory analysis of COVID-19 Tweets using topic modeling, UMAP, and DiGraphs," 2020, *arXiv:2005.03082*. [Online]. Available: https://arxiv.org/abs/2005.03082

[38] A. Nasir, K. Shaukat, I. A. Hameed, S. Luo, T. M. Alam, and F. Iqbal, "A bibliometric analysis of corona pandemic in social sciences: A review of influential aspects and conceptual structure," *IEEE Access*, vol. 8, pp. 133377–133402, 2020, doi: 10.1109/ACCESS.2020.3008733.

[39] M. Haghani and M. C. J. Bliemer, "COVID-19 pandemic and the unprecedented mobilisation of scholarly efforts prompted by a health crisis: Scientometric comparisons across SARS, MERS and 2019-nCoV literature," *Scientometrics*, vol. 125, no. 3, pp. 2695–2726, Dec. 2020, doi: 10.1007/s11192-020-03706-z.

[40] X. Cheng, Q. Cao, and S. S. Liao, "An overview of literature on COVID-19, MERS and SARS: Using text mining and latent Dirichlet allocation," *J. Inf. Sci.*, Aug. 2020, Art. no. 0165551520954674, doi: 10.1177/0165551520954674.

[41] M. Dong, X. Cao, M. Liang, L. Li, G. Liu, and H. Liang, "Understand research hotspots surrounding COVID-19 and other coronavirus infections using topic modeling," *Medrxiv*, 2020, doi: 10.1101/2020.03.26.20044164.

[42] H. Norman, N. Nordin, R. Din, M. Ally, and H. Dogan, "Exploring the roles of social participation in mobile social media learning: A social network analysis," *Int. Rev. Res. Open Distrib. Learn.*, vol. 16, no. 4, pp. 205–224, Nov. 2015, doi: 10.19173/irrodl.v16i4.2124.

[43] C. Evans, "Twitter for teaching: Can social media be used to enhance the process of learning?" *Brit. J. Educ. Technol.*, vol. 45, no. 5, pp. 902–915, Sep. 2014, doi: 10.1111/bjet.12099.

[44] Y. Liu, "Social media tools as a learning resource," *J. Educ. Technol. Develop. Exchange*, vol. 3, no. 1, pp. 101–114, Jun. 2010, doi: 10.18785/jetde.0301.08.

[45] K. Canese and S. Weis, "PubMed: The bibliographic database," in *The NCBI Handbook*, 2nd ed. Bethesda, MA, USA: National Center for Biotechnology Information (US), 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK153385/

[46] R. Řehůřek and P. Sojka, "Gensim—Statistical semantics in Python," EuroScipy, Paris, France, 2011. [Online]. Available: https://www.euroscipy.org/conference/euroscipy2011

[47] A. Aizawa, "An information-theoretic perspective of tf–idf measures," *Inf. Process. Manage.*, vol. 39, no. 1, pp. 45–65, 2003, doi: 10.1016/S0306-4573(02)00021-3.

[48] D. Mutch, A. Berger, R. Mansourian, A. Rytz, and M. Roberts, "The limit fold change model: A practical approach for selecting differentially expressed genes from microarray data," *BMC Bioinf.*, vol. 3, no. 1, pp. 1–11, 2002, doi: 10.1186/1471-2105-3-17.

[49] A. Cichocki and A.-H. Phan, "Fast local algorithms for large scale nonnegative matrix and tensor factorizations," *IEICE Trans. Fundamentals Electron., Commun. Comput. Sci.*, vol. 92, no. 3, pp. 708–721, 2009, doi: 10.1587/transfun.E92.A.708.

[50] V. Shah, T. Ko Ko, M. Zuckerman, J. Vidler, S. Sharif, V. Mehra, S. Gandhi, A. Kuhnl, D. Yallop, D. Avenoso, and C. Rice, "Poor outcome and prolonged persistence of SARS-CoV-2 RNA in COVID-19 patients with haematological malignancies; king's college hospital experience," *Brit. J. Haematol.*, vol. 190, no. 5, pp. 279–e282, Sep. 2020, doi: 10.1111/bjh.16935.

[51] C. Vellas, P. Delobel, P. De Souto Barreto, and J. Izopet, "COVID-19, virology and geroscience: A perspective," *J. Nutrition, Health Aging*, vol. 24, no. 7, pp. 685–691, Jul. 2020, doi: 10.1007/s12603-020-1416-2.

[52] C. M. Poh, G. Carissimo, B. Wang, S. N. Amrun, C. Y.-P. Lee, R. S.-L. Chee, S.-W. Fong, N. K.-W. Yeo, W.-H. Lee, A. Torres-Ruesta, and Y. S. Leo, "Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients," *Nature Commun.*, vol. 11, no. 1, p. 2806, Dec. 2020, doi: 10.1038/s41467-020-16638-2.

[53] S. Kannan, P. S. S. Ali, A. Sheeza, and K. Hemalatha, "COVID-19 (novel coronavirus 2019)-recent trends," *Eur. Rev. Med. Pharmacol. Sci*, vol. 24, no. 4, pp. 2006–2011, 2020, doi: 10.26355/eurrev_202002_20378.

[54] B. Shanmugaraj, K. Siriwattananon, K. Wangkanont, and W. Phoolcharoen, "Perspectives on monoclonal antibody therapy as potential therapeutic intervention for coronavirus disease-19 (COVID-19)," *Asian Pacific J. Allergy Immunol.*, vol. 38, no. 1, pp. 10–18, 2020, doi: 10.12932/AP-200220-0773.

[55] M.-Y. Zhou, X.-L. Xie, Y.-G. Peng, M.-J. Wu, X.-Z. Deng, Y. Wu, L.-J. Xiong, and L.-H. Shang, "From SARS to COVID-19: What we have learned about children infected with COVID-19," *Int. J. Infectious Diseases*, vol. 96, pp. 710–714, Jul. 2020, doi: 10.1016/j.ijid.2020.04.090.

[56] C. Bailey, J. R. M. Black, and C. Swanton, "Cancer research: The lessons to learn from COVID-19," *Cancer Discovery*, vol. 10, no. 9, pp. 1263–1266, Sep. 2020, doi: 10.1158/2159-8290.CD-20-0823.

[57] M. Sullivan *et al.*, "The COVID-19 pandemic: A rapid global response for children with cancer from SIOP, COG, SIOP-E, SIOP-PODC, IPSO, PROS, CCI, and St Jude Global," *Pediatric Blood Cancer*, vol. 67, no. 7, Jul. 2020, Art. no. e28409, doi: 10.1002/pbc.28409.

[58] F. P. Cavalcante, G. G. Novita, E. C. Millen, F. P. Zerwes, V. M. de Oliveira, A. L. L. Sousa, and R. F. Junior, "Management of early breast cancer during the COVID-19 pandemic in Brazil," *Breast Cancer Res. Treatment*, vol. 184, no. 2, pp. 637–647, Nov. 2020, doi: 10.1007/s10549-020-05877-y.

[59] Y. Hou, J. Zhao, W. Martin, A. Kallianpur, M. K. Chung, L. Jehi, N. Sharifi, S. Erzurum, C. Eng, and F. Cheng, "New insights into genetic susceptibility of COVID-19: An ACE2 and TMPRSS2 polymorphism analysis," *BMC Med.*, vol. 18, no. 1, pp. 216–223, Dec. 2020, doi: 10.1186/s12916-020-01673-z.

[60] J. Vallamkondu, A. John, W. Y. Wani, S. P. Ramadevi, K. K. Jella, P. H. Reddy, and R. Kandimalla, "SARS-CoV-2 pathophysiology and assessment of coronaviruses in CNS diseases with a focus on therapeutic targets," *Biochimica Biophysica Acta*, vol. 1866, no. 10, Oct. 2020, Art. no. 165889, doi: 10.1016/j.bbadis.2020.165889.

[61] S. H. Preskorn, "The 5% of the population at high risk for severe COVID-19 infection is identifiable and needs to be taken into account when reopening the economy," *J. Psychiatric Pract.*, vol. 26, no. 3, pp. 219–227, May 2020, doi: 10.1097/PRA.0000000000000475.

[62] D.-G. Ahn, H.-J. Shin, M.-H. Kim, S. Lee, H.-S. Kim, J. Myoung, B.-T. Kim, and S.-J. Kim, "Current status of epidemiology, diagnosis, therapeutics, and vaccines for novel coronavirus disease 2019 (COVID-19)," *J. Microbiol. Biotechnol.*, vol. 30, no. 3, pp. 313–324, Mar. 2020, doi: 10.4014/jmb.2003.03011.

[63] G. Li, X. He, L. Zhang, Q. Ran, J. Wang, A. Xiong, D. Wu, F. Chen, J. Sun, and C. Chang, "Assessing ACE2 expression patterns in lung tissues in the pathogenesis of COVID-19," *J. Autoimmunity*, vol. 112, Aug. 2020, Art. no. 102463, doi: 10.1016/j.jaut.2020.102463.

[64] B. M. Zawilińska and M. Kosz-Vnenchak, "General introduction into the Ebola virus biology and disease," *Folia Medica Cracoviensia*, vol. 54, no. 3, pp. 57–65, 2014. [Online]. Available: http://europepmc.org/abstract/MED/25694096

[65] A. Rivera and I. Messaoudi, "Molecular mechanisms of Ebola pathogenesis," *J. Leukocyte Biol.*, vol. 100, no. 5, pp. 889–904, Nov. 2016, doi: 10.1189/jlb.4RI0316-099RR.

[66] R. K. Singh, K. Dhama, Y. S. Malik, M. A. Ramakrishnan, K. Karthik, R. Khandia, R. Tiwari, A. Munjal, M. Saminathan, S. Sachan, P. A. Desingu, J. J. Kattoor, H. M. N. Iqbal, and S. K. Joshi, "Ebola virus—Epidemiology, diagnosis, and control: Threat to humans, lessons learnt, and preparedness plans—An update on its 40 year's journey," *Veterinary Quart.*, vol. 37, no. 1, pp. 98–135, Jan. 2017, doi: 10.1080/01652176.2017.1309474.

[67] E. M. Coves-Datson, J. Dyall, L. E. DeWald, S. R. King, D. Dube, M. Legendre, E. Nelson, K. C. Drews, R. Gross, D. M. Gerhardt, L. Torzewski, E. Postnikova, J. Y. Liang, B. Ban, J. Shetty, L. E. Hensley, P. B. Jahrling, G. G. Olinger, J. M. White, and D. M. Markovitz, "Inhibition of Ebola virus by a molecularly engineered banana lectin," *PLOS Neglected Tropical Diseases*, vol. 13, no. 7, pp. 1–20, 2019, doi: 10.1371/journal.pntd.0007595.

[68] A. Marí Saéz *et al.*, "Investigating the zoonotic origin of the West African Ebola epidemic," *EMBO Mol. Med.*, vol. 7, no. 1, pp. 17–23, Jan. 2015, doi: 10.15252/emmm.201404792.

[69] V. K. Shah, P. Firmal, A. Alam, D. Ganguly, and S. Chattopadhyay, "Overview of immune response during SARS-CoV-2 infection: Lessons from the past," *Frontiers Immunol.*, vol. 11, p. 1949, Aug. 2020, doi: 10.3389/fimmu.2020.01949.

[70] Y. Sugita, H. Matsunami, Y. Kawaoka, T. Noda, and M. Wolf, "Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 å resolution," *Nature*, vol. 563, no. 7729, pp. 137–140, Nov. 2018, doi: 10.1038/s41586-018-0630-0.

[71] L. Zhu, T. Gao, W. Yang, Y. Liu, X. Liu, Y. Hu, Y. Jin, P. Li, K. Xu, G. Zou, L. Zhao, R. Cao, W. Zhong, X. Xia, and C. Cao, "Ebola virus replication is regulated by the phosphorylation of viral protein VP35," *Biochem. Biophys. Res. Commun.*, vol. 521, no. 3, pp. 687–692, Jan. 2020, doi: 10.1016/j.bbrc.2019.10.147.

[72] M. L. Pleet, J. Erickson, C. DeMarino, R. A. Barclay, M. Cowen, B. Lepene, J. Liang, J. H. Kuhn, L. Prugar, S. W. Stonier, J. M. Dye, W. Zhou, L. A. Liotta, M. J. Aman, and F. Kashanchi, "Ebola virus VP40 modulates cell cycle and biogenesis of extracellular vesicles," *J. Infectious Diseases*, vol. 218, no. 5, pp. S365–S387, Nov. 2018, doi: 10.1093/infdis/jiy472.

[73] L. Wu, D. Jin, D. Wang, X. Jing, P. Gong, Y. Qin, and M. Chen, "The two-stage interaction of Ebola virus VP40 with nucleoprotein results in a switch from viral RNA synthesis to virion assembly/budding," *Protein Cell*, vol. 1210, pp. 1–21, Nov. 2020, doi: 10.1007/s13238-020-00764-0.

[74] M. Kerper and Y. Puckett. (Jan. 2021). *Filovirus.* StatPearls Treasure Island (FL): StatPearls Publishing. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK544258/

[75] N. K. Alharbi, A. J. Spencer, A. M. Salman, C. M. Tully, S. K. Chinnakannan, T. Lambe, Y. Yamaguchi, S. J. Morris, T. Orubu, S. J. Draper, A. V. S. Hill, and S. C. Gilbert, "Enhancing cellular immunogenicity of MVA-vectored vaccines by utilizing the F11L endogenous promoter," *Vaccine*, vol. 34, no. 1, pp. 49–55, Jan. 2016, doi: 10.1016/j.vaccine.2015.11.028.

[76] C.-H. Hsiao, M.-Z. Wu, S.-W. Hsieh, L.-C. Chien, K.-C. Hwang, and I.-J. Su, "Clinicopathology of severe acute respiratory syndrome: An autopsy case report," *J. Formosan Med. Assoc.= Taiwan yi zhi*, vol. 103, no. 10, pp. 787–792, Oct. 2004. [Online]. Available: https://europepmc.org/abstract/MED/15490031

[77] F. F. Yuan, J. Tanner, P. K. S. Chan, S. Biffin, W. B. Dyer, A. F. Geczy, J. W. Tang, D. S. C. Hui, J. J. Y. Sung, and J. S. Sullivan, "Influence of FcgammaRIIA and MBL polymorphisms on severe acute respiratory syndrome," *Tissue Antigens*, vol. 66, no. 4, pp. 291–296, Oct. 2005, doi: 10.1111/j.1399-0039.2005.00476.x.

[78] D. Escudero, J. A. Boga, J. Fernández, L. Forcelledo, S. Balboa, R. Albillos, I. Astola, E. García-Prieto, M. E. Álvarez-Argüelles, G. Martín, J. Jiménez, and F. Vázquez, "SARS-CoV-2 analysis on environmental surfaces collected in an intensive care unit: Keeping ernest Shackleton's spirit," *Intensive Care Med. Exp.*, vol. 8, no. 1, pp. 68–75, Dec. 2020, doi: 10.1186/s40635-020-00349-5.

[79] E. Petersen, M. Koopmans, U. Go, D. H. Hamer, N. Petrosillo, F. Castelli, M. Storgaard, S. Al Khalili, and L. Simonsen, "Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics," *Lancet Infectious Diseases*, vol. 20, no. 9, pp. e238–e244, 2020, doi: 10.1016/S1473-3099(20)30484-9.

[80] S. Esposito and N. Principi, "Influenza vaccination and prevention of antimicrobial resistance," *Expert Rev. Vaccines*, vol. 17, no. 10, pp. 881–888, Oct. 2018, doi: 10.1080/14760584.2018.1525298.

[81] S. A. Rasmussen and D. J. Jamieson, "Influenza and pregnancy: No time for complacency," *Obstetrics Gynecol.*, vol. 133, no. 1, pp. 23–26, Jan. 2019, doi: 10.1097/AOG.0000000000003040.

[82] B. Simon, M. Pichon, M. Valette, G. Burfin, M. Richard, B. Lina, and L. Josset, "Whole genome sequencing of A(H3N2) influenza viruses reveals variants associated with severity during the 2016–2017 season," *Viruses*, vol. 11, no. 2, p. 108, Jan. 2019, doi: 10.3390/v11020108.

[83] B. F. Johnson, L. E. Wilson, J. Ellis, A. J. Elliot, W. S. Barclay, R. G. Pebody, J. McMenamin, D. M. Fleming, and M. C. Zambon, "Fatal cases of influenza a in childhood," *PLoS ONE*, vol. 4, no. 10, pp. 1–6, 2009, doi: 10.1371/journal.pone.0007671.

[84] N. E. MacDonald and J. C. McDonald, "The benefits of influenza vaccine in pregnancy for the fetus and the infant younger than six months of age," *Paediatrics Child Health*, vol. 19, no. 9, pp. e121–e122, Nov. 2014, doi: 10.1093/pch/19.9.e121.

[85] L. C. Jennings, A. S. Monto, P. K. Chan, T. D. Szucs, and K. G. Nicholson, "Stockpiling prepandemic influenza vaccines: A new cornerstone of pandemic preparedness plans," *Lancet Infectious Diseases*, vol. 8, no. 10, pp. 650–658, 2008, doi: 10.1016/S1473-3099(08)70232-9.

[86] L. Zhou, R. Ren, L. Yang, C. Bao, J. Wu, D. Wang, C. Li, N. Xiang, Y. Wang, D. Li, and H. Sui, "Sudden increase in human infection with avian influenza A(H7N9) virus in China, September–December 2016," *Western Pacific Surveill. Response J.*, vol. 8, no. 1, pp. 6–14, Apr. 2017, doi: 10.5365/WPSAR.2017.8.1.001.

[87] N. J. Dharan, L. V. Gubareva, J. J. Meyer, M. Okomo-Adhiambo, R. C. McClinton, S. A. Marshall, K. S. George, S. Epperson, L. Brammer, A. I. Klimov, and J. S. Bresee, "Infections with oseltamivir-resistant influenza A (H1N1) virus in the United States," *J. Amer. Med. Assoc.*, vol. 301, no. 10, pp. 1034–1041, 2009, doi: 10.1001/jama.2009.294.

[88] K. Nicholson, F. Aoki, A. Osterhaus, S. Trottier, O. Carewicz, C. Mercier, A. Rode, N. Kinnersley, and P. Ward, "Efficacy and safety of oseltamivir in treatment of acute influenza: A randomised controlled trial," *Lancet*, vol. 355, no. 9218, pp. 1845–1850, 2000, doi: 10.1016/S0140-6736(00)02288-1.

[89] V. Demicheli, T. Jefferson, C. Di Pietrantonj, E. Ferroni, S. Thorning, R. E. Thomas, and A. Rivetti, "Vaccines for preventing influenza in the elderly," *Cochrane Database Systematic Rev.*, vol. 2, Feb. 2018, doi: 10.1002/14651858.CD004876.pub4.

[90] M. S. Niederman, "Review of treatment guidelines for community-acquired pneumonia," *Amer. J. Med. Supplements*, vol. 117, no. 3, pp. 51–57, Aug. 2004, doi: 10.1016/j.amjmed.2004.07.008.

[91] R. Nishioka, A. Satomura, J. Yamada, K. Kuroda, and M. Ueda, "Rapid preparation of mutated influenza hemagglutinins for influenza virus pandemic prevention," *AMB Exp.*, vol. 6, no. 1, pp. 1–8, Dec. 2016, doi: 10.1186/s13568-016-0179-y.

[92] M. A. Psotka *et al.*, "Conduct of clinical trials in the era of COVID-19," *J. Amer. College Cardiol.*, vol. 76, no. 20, pp. 2368–2378, Nov. 2020, doi: 10.1016/j.jacc.2020.09.544.

[93] G. Aggarwal, B. M. Henry, S. Aggarwal, and S. Bangalore, "Cardiovascular safety of potential drugs for the treatment of coronavirus disease 2019," *Amer. J. Cardiol.*, vol. 128, pp. 147–150, Aug. 2020, doi: 10.1016/j.amjcard.2020.04.054.

[94] C. M. Chu, Y. Y. Leung, J. Y. H. Hui, I. F. N. Hung, V. L. Chan, W. S. Leung, K. I. Law, C. S. Chan, K. S. Chan, and K. Y. Yuen, "Spontaneous pneumomediastinum in patients with severe acute respiratory syndrome," *Eur. Respiratory J.*, vol. 23, no. 6, pp. 802–804, Jun. 2004, doi: 10.1183/09031936.04.00096404.

[95] C. Tsai and C.-Y. Chi, "Clinical characteristics of fatal patients with severe acute respiratory syndrome in a medical center in Taipei," *J. Chin. Med. Assoc.*, vol. 66, no. 6, pp. 323–327, Jun. 2003. [Online]. Available: https://europepmc.org/abstract/MED/12889500

[96] K. Ksiezakowska, M. Laszczyk, J. Wilczynski, and D. Nowakowska, "SARS-CoV infection and pregnancy," *Ginekologia Polska*, vol. 79, no. 1, pp. 47–50, 2008. [Online]. Available: https://journals.viamedica.pl/ginekologia_polska/article/view/46767

[97] I. Pinedo-Torres, M. Flores-Fernandez, M. Yovera-Aldana, C. Gutierrez-Ortiz, P. Zegarra-Lizana, C. Intimayta-Escalante, C. Moran-Marinos, C. Alva-Diaz, and K. Pacheco-Barrios, "Prevalence of diabetes mellitus and its associated unfavorable outcomes in patients with acute respiratory syndromes due to coronaviruses infection: A systematic review and meta-analysis," *Clin. Med. Insights, Endocrinol. Diabetes*, vol. 13, Oct. 2020, Art. no. 1179551420962495, doi: 10.1177/1179551420962495.

[98] D. Di Mascio, A. Khalil, G. Saccone, G. Rizzo, D. Buca, M. Liberati, J. Vecchiet, L. Nappi, G. Scambia, V. Berghella, and F. D'Antonio, "Outcome of coronavirus spectrum infections (SARS, MERS, COVID-19) during pregnancy: A systematic review and meta-analysis," *Amer. J. Obstetrics Gynecol. MFM*, vol. 2, no. 2, May 2020, Art. no. 100107. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589933320300379

[99] V. Kumar, "Emerging human coronavirus infections (SARS, MERS, and COVID-19): Where they are leading us," *Int. Rev. Immunol.*, vol. 40, nos. 1–2, pp. 5–53, Mar. 2021, doi: 10.1080/08830185.2020.1800688.

[100] K. Schwarz and D. A. Groneberg, "Aktueller Überblick zum MERS-virus," *Zentralblatt Arbeitsmedizin, Arbeitsschutz Ergonomie*, vol. 65, no. 6, pp. 353–354, Nov. 2015, doi: 10.1007/s40664-015-0062-8.

[101] O. D. Lara, R. E. O'Cearbhaill, M. J. Smith, M. E. Sutter, A. Knisely, J. McEachron, L. R. Gabor, J. Jee, J. E. Fehniger, Y. Lee, S. S. Isani, J. D. Wright, and B. Pothuri, "COVID-19 outcomes of patients with gynecologic cancer in New York city," *Cancer*, vol. 126, no. 19, pp. 4294–4303, Oct. 2020, doi: 10.1002/cncr.33084.

[102] B. Jordan, "Vaccination against infectious bronchitis virus: A continuous challenge," *Veterinary Microbiol.*, vol. 206, pp. 137–143, Jul. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378113517300093

[103] R. A. Zegpi, C. Breedlove, V. L. van Santen, C. R. Rasmussen-Ivey, and H. Toro, "Kidney cell–adapted infectious bronchitis ArkDPI vaccine is stable and protective," *Avian Diseases*, vol. 61, no. 2, pp. 221–228, Jun. 2017, doi: 10.1637/11537-111416-Reg.1.

[104] Y. Hou and Q. Wang, "Emerging highly virulent porcine epidemic diarrhea virus: Molecular mechanisms of attenuation and rational design of live attenuated vaccines," *Int. J. Mol. Sci.*, vol. 20, no. 21, p. 5478, Nov. 2019, doi: 10.3390/ijms20215478.

[105] T. Sato, K. Oroku, Y. Ohshima, Y. Furuya, and C. Sasakawa, "Efficacy of genogroup 1 based porcine epidemic diarrhea live vaccine against genogroup 2 field strain in Japan," *Virol. J.*, vol. 15, no. 1, p. 28, Dec. 2018, doi: 10.1186/s12985-018-0940-8.

[106] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, and Y. Bi, "Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding," *Lancet*, vol. 395, no. 10224, pp. 565–574, 2020, doi: 10.1016/S0140-6736(20)30251-8.

[107] G. Bruno, S. Perelli, C. Fabrizio, and G. B. Buccoliero, "Short-term outcomes in individuals aged 75 or older with severe coronavirus disease (COVID-19): First observations from an infectious diseases unit in southern Italy," *J. Infection*, vol. 81, no. 2, pp. e86–e88, Aug. 2020, doi: 10.1016/j.jinf.2020.05.024.

[108] L. Bigiani, S. Bigiani, and A. Bigiani, "How to minimize the impact of pandemic events: Lessons from the COVID-19 crisis," *Int. J. Health Policy Manage.*, vol. 9, pp. 469–474, Jul. 2020, doi: 10.34172/ijhpm.2020.115.

[109] L. Simione and C. Gnagnarella, "Differences between health workers and general population in risk perception, behaviors, and psychological distress related to COVID-19 spread in Italy," *Frontiers Psychol.*, vol. 11, p. 2166, Sep. 2020, doi: 10.3389/fpsyg.2020.02166.

[110] R. Tirupathi, V. Muradova, R. Shekhar, S. A. Salim, J. A. Al-Tawfiq, and V. Palabindala, "COVID-19 disparity among racial and ethnic minorities in the US: A cross sectional analysis," *Travel Med. Infectious Disease*, vol. 38, Nov. 2020, Art. no. 101904, doi: 10.1016/j.tmaid.2020.101904.

[111] H. J. Krouse, "COVID-19 and the widening gap in health inequity," *Otolaryngol. Head Neck Surg.*, vol. 163, no. 1, pp. 65–66, Jul. 2020, doi: 10.1177/0194599820926463.

[112] C. Ferrando *et al.*, "Patient characteristics, clinical course and factors associated to ICU mortality in critically ill patients infected with SARS-CoV-2 in spain: A prospective, cohort, multicentre study," *Revista Española de Anestesiología y Reanimación*, vol. 67, no. 8, pp. 425–437, Oct. 2020, doi: 10.1016/j.redare.2020.07.001.

[113] A. T. Craig, A. E. Heywood, and J. Hall, "Risk of COVID-19 importation to the pacific islands through global air travel," *Epidemiol. Infection*, vol. 148, p. e71, Mar. 2020, doi: 10.1017/S0950268820000710.

[114] P. Bourée, "La COVID-19 à Mayotte : Toujours en ≪ rouge ≫," *Revue Francophone Laboratoires*, vol. 2020, no. 526, pp. 70–75, 2020, doi: 10.1016/S1773-035X(20)30316-6.

[115] D. J. Steenbergen, P. T. Neihapi, D. Koran, A. Sami, V. Malverus, R. Ephraim, and N. Andrew, "COVID-19 restrictions amidst cyclones and volcanoes: A rapid assessment of early impacts on livelihoods and food security in coastal communities in Vanuatu," *Mar. Policy*, vol. 121, Nov. 2020, Art. no. 104199, doi: 10.1016/j.marpol.2020.104199.

[116] C. Fosker, "COVID-19 and cancer care in Bermuda," *Lancet Oncol.*, vol. 21, no. 6, pp. 761–762, 2020, doi: 10.1016/S1470-2045(20)30272-2.

[117] A. A. T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I. K. Singh, A. Singh, S. M. Atif, G. Hariprasad, G. M. Hasan, and M. I. Hassan, "Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach," *Biochimica Biophysica Acta*, vol. 1866, no. 10, Oct. 2020, Art. no. 165878, doi: 10.1016/j.bbadis.2020.165878.

[118] T. Kustin and A. Stern, "Biased mutation and selection in RNA viruses," *Mol. Biol. Evol.*, vol. 38, no. 2, pp. 575–588, Jan. 2021, doi: 10.1093/molbev/msaa247.

[119] F. Narita, Z. Wang, H. Kurita, Z. Li, Y. Shi, Y. Jia, and C. Soutis, "A review of piezoelectric and magnetostrictive biosensor materials for detection of COVID-19 and other viruses," *Adv. Mater.*, vol. 33, no. 1, Jan. 2021, Art. no. 2005448, doi: 10.1002/adma.202005448.

[120] A. Emadi, J. V. Chua, R. Talwani, S. M. Bentzen, and J. Baddley, "Safety and efficacy of imatinib for hospitalized adults with COVID-19: A structured summary of a study protocol for a randomised controlled trial," *Trials*, vol. 21, no. 1, p. 897, Dec. 2020, doi: 10.1186/s13063-020-04819-9.

[121] O. Nanni, P. Viale, B. Vertogen, C. Lilli, C. Zingaretti, C. Donati, C. Masini, M. Monti, P. Serra, R. Vespignani, and V. Grossi, "PROTECT trial: A cluster-randomized study with hydroxychloroquine versus observational support for prevention or early-phase treatment of coronavirus disease (COVID-19): A structured summary of a study protocol for a randomized controlled trial," *Trials*, vol. 21, no. 1, p. 689, Dec. 2020, doi: 10.1186/s13063-020-04527-4.

[122] C. Vardavas and K. Nikitara, "COVID-19 and smoking: A systematic review of the evidence," *Tobacco Induced Diseases*, vol. 18, p. 20, Mar. 2020, doi: 10.18332/tid/119324.

[123] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Di Napoli. (2021). *Features, Evaluation and Treatment Coronavirus (COVID-19)*. Statpearls. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK554776/

[124] K. L. Chai, S. J. Valk, V. Piechotta, C. Kimber, I. Monsef, C. Doree, E. M. Wood, A. A. Lamikanra, D. J. Roberts, Z. McQuilten, C. So-Osman, L. J. Estcourt, and N. Skoetz, "Convalescent plasma or hyperimmune immunoglobulin for people with COVID-19: A living systematic review," *Cochrane Database Systematic Rev.*, vol. 10, Oct. 2020, Art. no. CD013600, doi: 10.1002/14651858.CD013600.pub3.

[125] N. Phillips, "The coronavirus is here to stay—Here's what that means," *Nature*, vol. 590, no. 7846, pp. 382–384, Feb. 2021, doi: 10.1038/d41586-021-00396-2.

[126] K. Goyal, P. Chauhan, K. Chhikara, P. Gupta, and M. P. Singh, "Fear of COVID 2019: First suicidal case in India!" *Asian J. Psychiatry*, vol. 49, Mar. 2020, Art. no. 101989, doi: 10.1016/j.ajp.2020.101989.

PRIYANKAR BOSE received the B.Tech. degree in electronics and telecommunication engineering from KIIT University, India, in 2018. He is currently pursuing the Ph.D. degree with the Biological Networks Laboratory, Department of Computer Science, School of Engineering, Virginia Commonwealth University. He works on a variety of problems in the field of biological data sciences. His research interests include data/text mining, machine learning, NLP, and biological data modeling and simulations.

SATYAKI ROY received the Ph.D. degree in computer science from the Missouri University of Science and Technology, USA, in 2019. He is currently a Postdoctoral Research Associate with the Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. His research interests include computational biology, network science and optimization, wireless sensor networks, epidemiology, machine learning, and parallel computing.

PREETAM GHOSH received the B.S. degree in computer science from Jadavpur University, Kolkata, India, and the M.S. and Ph.D. degrees in computer science and engineering from The University of Texas at Arlington. He is currently a Professor and directs the Biological Networks Laboratory, Department of Computer Science, Virginia Commonwealth University. His research interests include algorithms, stochastic modeling and simulation, network science and machine learning related approaches in systems biology and computational epidemiology, and mobile computing related issues in pervasive grids that has resulted in more than 170 conference and journal articles and several federally funded research projects from NSF, NIH, DoD, and US-VHA. He also serves as the Secretary/the Treasurer of ACM SIGBio.

• • •