Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal of Biomedical Informatics 129 (2022) 104028

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Novel informatics approaches to COVID-19 Research: From methods to applications

1. Introduction

Started in December 2019, the Coronavirus Disease 2019 (COVID-19) rapidly developed into a worldwide pandemic. With the spread of the more infectious Omicron variant, over 328 million cases and 5.5 million deaths were reported globally, as of January 15th 2022 [1]. During the pandemic, unprecedented effort has been devoted to research to quickly understand transmission patterns, discover underlying molecular mechanisms, develop diagnostic tests, vaccines, and treatments, and evaluate public health policies to control COVID-19. The data-driven nature of biomedical research has made Biomedical Informatics –a discipline of developing and applying information science to biomedical problems to derive information, knowledge, and wisdom from data— very critical when combating the COVID-19 pandemic [2].

To highlight the development and application of novel informatics approaches to collecting, integrating, harmonizing, and analyzing COVID-19 data for scientific discoveries and clinical practice, the Journal of Biomedical Informatics (JBI) organized a special issue on "Novel Informatics Approaches to COVID-19 Research". Most papers were received in October of 2020; thus, many works in this collection were done in 2020. Forty-five articles are presented in the special issue, including three Commentaries, two Methodological Reviews, twenty-seven Original Research articles, and thirteen Special Communications. In Table 1 we have clustered the papers in the special issue by broad themes. Original Research and Special Communication articles cover diverse themes of COVID-19 research, including methods and applications of using clinical and public health data to support clinical research or practice (13 articles), epidemic models for forecasting and other applications (9 articles), translational bioinformatics approaches for analyzing genomic data (4 articles), and text mining methods for literature, social media, and clinical trial documents (7 articles), as well as seven other articles covering imaging, mHealth, and visualization technologies. The diversity of research topics in COVID-19 indicates the need of developing different informatics approaches to tackle this complex disease as well as more broadly infectious disease outbreaks and the study of other diseases across healthcare systems. The ongoing COVID 19 pandemic shows a continued need for these types of tools and approaches which can also be applied in other pandemic situations (especially with respiratory transmission).

2. The role of informatics in the COVID-19 pandemic - commentaries and reviews

During the pandemic, information technologies play an important role in collecting, normalizing, analyzing, and disseminating data, information, and knowledge for quick and correct healthcare, social, and economic responses. In this section we discuss the broader role of informatics during the COVID-19 pandemic.

(a) Commentaries. Three commentaries focused on the roles of informatics in COVID-19 and provide important insights. First, Ćurković et al. [3] discussed the issues of truthfulness of digital content during pandemic, e.g., misinformation and disinformation and their potential harms. They suggested creating a reproducible, valid and truthful informational landscape, by considering ethical governance of digital data/information/knowledge curation and dissemination. Second, Sudat et al. [4] emphasized the importance of data quality in electronic health records (EHRs), as EHR are a main source for real-time reporting of cases and outcomes to inform public health responses. They discussed lessons learned from a large integrated health system in northern California, including some limitations of EHR data and the need to incorporate data and analytics in the early stages of clinical crisis planning. Third, a group of biomedical informaticians at Stanford provided their insights from a different aspect – how to repurpose biomedical informaticians for COVID-19. [5] They described several COVID-19 projects by graduate students and postdoctoral fellows in Biomedical Informatics, such as artificial intelligence (AI) for vaccine development, mining knowledge from biomedical literature, and developing contact tracing methods and tools. Then they summarized lessons learned from those projects, such as how to make such collaborations more efficient for scientific progress. These insights are important for ongoing informatics efforts around the current pandemic and provide lessons for systems to monitor and intervene in infectious disease outbreaks more broadly.

(b) Methodological reviews Two methodological reviews also provided overviews of the use and development of digital technologies and AI methods in the pandemic. First, Thomas et al. [6] conducted a scoping review of a broad range of digital technologies used in COVID-19, including mobile applications, robots, drones, social media platforms, television, and associated technologies. The discussion about adoption status and associated issues of those digital technologies outlined in the review will be helpful for governments and policymakers to better understand utilities of such modern technologies in a pandemic. Second, another review by Shahid and team [7] focused on machine learning and deep learning methods used for COVID-19 screening, forecasting and tracking, and drug/vaccine development. It provides a comprehensive survey of how machine learning technologies have been used to fight against COVID-19. Again, these reviews provide guidance for fighting the current pandemic and provide insights for future similar work.

https://doi.org/10.1016/j.jbi.2022.104028
Received 7 February 2022; Accepted 10 February 2022
Available online 16 February 2022
1532-0464/© 2022 Elsevier Inc. This article is made available under the Elsevier license (http://www.elsevier.com/open-access/userlicense/1.0/).

Table 1
Articles in the COVOD-19 Special Issue.

Reference	Title	First Author	Type of Article
Commentary and review (5)			
[3]	Epistemic responsibilities in the COVID-19 pandemic: Is a digital infosphere a friend or a foe?	Curkovic, M.	Commentary
[4]	Mind the Clinical-Analytic Gap: Electronic Health Records and COVID-19 Pandemic Response	Sudat,S.E.K.	Commentary
[5]	Repurposing biomedical informaticians for COVID-19	Sosa, D. N.	Commentary
[6]	Can technological advancements help to alleviate COVID-19 pandemic? a review	Thomas, M. J.	Methodological Review
[7]	Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance	Shahid, O.	Methodological Review
Clinical research and practice (13)			
[8]	Making science computable: Developing code systems for statistics, study design, and risk of bias	Alper, B. S.	Special Communication
[9]	Obtaining EHR-derived datasets for COVID-19 research within a short time: a flexible methodology based on Detailed Clinical Models	Pedrera-Jimenez, M.	Original Research
[10]	Developing a sampling method and preliminary taxonomy for classifying COVID-19 public health guidance for healthcare organizations and the general public	Taber, P.	Original Research
[11]	Visual comprehension and orientation into the COVID-19 CIDO ontology	Zheng, L.	Original Research
[12]	Phenotyping coronavirus disease 2019 during a global health pandemic: Lessons learned from the characterization of an early cohort	DeLozier, S.	Special Communication
[13]	Extracting COVID-19 diagnoses and symptoms from clinical text: A new annotated corpus and neural event extraction framework	Lybarger, K.	Original Research
[14]	ELII: A novel inverted index for fast temporal query, with application to a large Covid-19 EHR dataset	Huang, Y.	Original Research
[15]	Creating and implementing a COVID-19 recruitment Data Mart	Helmer, T. T.	Special Communication
[16]	Critical carE Database for Advanced Research (CEDAR): An automated method to support intensive care units with electronic health record data	Schenck, E. J.	Special Communication
[17]	ConceptWAS: A high-throughput method for early identification of COVID-19 presenting symptoms and characteristics from clinical notes	Zhao, J.	Special Communication
[18]	A retrospective approach to evaluating potential adverse outcomes associated with delay of procedures for cardiovascular and cancer-related diagnoses in the context of COVID-19	Zheng, N. S.	Original Research
[19]	Learning from past respiratory failure patients to triage COVID-19 patient ventilator needs: A multi-institutional study	Carmichael, H.	Special Communication
[20]	A predictive model of clinical deterioration among hospitalized COVID-19 patients by harnessing hospital course trajectories	Mauer, E.	Original Research
Forecasting and epidemic modelling (9)			
[21]	Implementation of stacking based ARIMA model for prediction of Covid-19 cases in India	Swaraj, A.	Original Research
[22]	A fusion of data science and feed-forward neural network-based modelling of COVID-19 outbreak forecasting in IRAQ	Aljaaf, A. J.	Original Research
[23]	Comparative study of machine learning methods for COVID-19 transmission forecasting	Dairi, A.	Original Research
[24]	A novel deep interval type-2 fuzzy LSTM (DIT2FLSTM) model applied to COVID-19 pandemic time-series prediction	Safari, A.	Original Research
[25]	COVID-19 modelling by time-varying transmission rate associated with mobility trend of driving via Apple Maps	JING, M.	Original Research
[26]	EPIsembleVis: A geo-visual analysis and comparison of the prediction ensembles of multiple COVID-19 models	Xu, H.	Special Communication
[27]	Population stratification enables modeling effects of reopening policies on mortality and hospitalization rates	Huang T.	Original Research
[28]	An integrated framework for modelling quantitative effects of entry restrictions and travel quarantine on importation risk of COVID-19	Chen, T.	Original Research
[29]	Dynamic graph and polynomial chaos based models for contact tracing data analysis and optimal testing prescription	Ubaru, S.	Original Research
Text mining of literature, social media, and trial documents (7)			
[30]	Searching for scientific evidence in a pandemic: An overview of TREC-COVID	Roberts, K.	Special Communication
[31]	A comparative analysis of system features used in the TREC-COVID information retrieval challenge	Chen, J.	Original Research
[32]	Drug repurposing for COVID-19 via knowledge graph completion	Zhang, R.	Original Research
[33]	Pulse of the pandemic: Iterative topic filtering for clinical information extraction from social media	Wu, J.	Original Research
[34]	Defining facets of social distancing during the COVID-19 pandemic: Twitter analysis	Kwon, J.	Original Research
[35]	Tracking and analysis of discourse dynamics and polarity during the early Corona pandemic in Iran	Jafarinejad, F.	Special Communication
[36]	Building an OMOP common data model-compliant	Sun, Y.	Original Research

(*continued on next page*)

Table 1 (*continued*)

Reference	Title	First Author	Type of Article
	annotated corpus for COVID-19 clinical trials		
Translational bioinformatics (4)			
[37]	A scheme for inferring viral-host associations based on codon usage patterns identifies the most affected signaling pathways during COVID-19	Das, J. K.	Original Research
[38]	Causal discovery using compression-complexity measures	Sy, P.	Original Research
[39]	Machine learning enabled identification of potential SARS-CoV-2 3CLpro inhibitors based on fixed molecular fingerprints and Graph-CNN neural representations	Haneczok, J.	Special Communication
[40]	Reservoir hosts prediction for COVID-19 by hybrid transfer learning model	Yang, Y.	Original Research
Others (7)			
[41]	Dynamic deformable attention network (DDANet) for COVID-19 lesions semantic segmentation	Rajamani, K. T.	Original Research
[42]	An aberration detection-based approach for sentinel syndromic surveillance of COVID-19 and other novel influenza-like illnesses	Wen, A.	Original Research
[43]	IoT-based GPS assisted surveillance system with inter-WBAN geographic routing for pandemic situations	Savasci Sen, S.	Special Communication
[44]	Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread	Sethi, S.	Original Research
[45]	Visual analytics of COVID-19 dissemination in Sao Paulo state, Brazil	Marcilio-Jr, W. E.	Original Research
[46]	COnVIDa: COVID-19 multidisciplinary data collection and dashboard	Martinez Beltran, E. T.	Special Communication
[47]	Factors influencing mHealth adoption and its impact on mental well-being during COVID-19 pandemic: A SEM-ANN approach	Alam, M. M. D.	Special Communication

3. Clinical research and practice using electronic health records

Clinical data of individual patients (e.g., stored in electronic health records - EHRs) or populations (e.g., from health departments) are critical data sources for studying disease comorbidities and comparing treatment effects. In this section we discuss four broad themes related to clinical research and practice in the context of the pandemic: (a) standards and information models, (b) methods, infrastructures, and resources to facilitate EHR-based clinical research, (c) methods for EHR data processing and analysis, and (d) predictive models.

(a) Standards and information models. As COVID-19 is a new disease, quickly defining standard representations (e.g., codes in medical terminologies) becomes the first step of informatics research. To address this gap, four papers present work on ontologies and information models related to COVID-19: (1) A group of researchers from the COVID-19 Knowledge Accelerator (COKA) initiative proposed the development of a code system for electronic data exchange for the identification, processing, and reporting of scientific findings of COVID-19 [8]; (2) To facilitate COVID-19 clinical research using EHRs data, Pedrera-Jimenez et al. [9] designed and implemented a flexible methodology based on detailed clinical models (DCM), which can quickly generate derived research datasets from ERHs without loss of meaning; (3) To better classify public health guidelines on COVID-19, Dr. Taber and team developed a taxonomy using a novel sampling method [10]; and (4) Zheng et al. [11] introduced an innovative visualization of the weighted aggregate taxonomy for better orientation and comprehension of CIDO (Coronavirus Infectious Disease Ontology), the largest COVID ontology. The lessons described above apply to both current and future infectious disease outbreaks. Some of the approaches (such as DCM) are generalizable to research on other diseases as well.

(b) Methods, infrastructures, and resources to facilitate EHR-based clinical research. It is non-trivial to extract and normalize EHR data to support COVID-19 research. Five papers have worked on developing novel methods, infrastructures, and resources to facilitate EHR-based clinical research. (1) DeLozier et al. [12] described their phenotyping methods and lessons learned from characterization of an early cohort of COVID-19 patients at Vanderbilt University Medical Center (VUMC). (2) Lybarger and team [13] developed a new annotated corpus and a neural event extraction framework for extracting COVID-19 diagnoses and symptoms from clinical documents, to facilitate in-depth analysis of COVID-19 symptoms. (3) As complex temporal queries may take a long time to execute on large EHR-databases, Huang, Li, and Zhang [14] developed a new tool called Event-level Inverted Index (ELII), which applies a novel inverted index for quickly searching COVID-19 cohorts. (4) On a related topic of finding patients, but for the purpose of trial recruitments, Helmer and the team at VUMC created and implemented a COVID-19 recruitment data mart, to facilitate time-sensitive trial opportunities such as COVID-19 [15]. (5) To support care at intensive care units (ICU), a group of researchers at New York-Presbyterian/Weill Cornell Medical Center has developed the Critical carE Database for Advanced Research (CEDAR), a method for extracting and transforming data from EHR systems for ICU uses [16]. Secondary use of EHRs data for clinical research remains challenging and the proposed methods here for COVID-19 data extraction and conversion could be generalizable to other diseases.

(c) Methods for EHR data processing and analysis. Novel informatics methods for EHR data processing and analysis have also been applied to support clinical studies. Zhao et al. [17] conducted a study and identified 68 COVID-19 symptom concepts by using a high-throughput method called ConceptWAS. Zheng et al. [18] performed a retrospective study to evaluate potential adverse outcomes associated with delay of procedures for cardiovascular and cancer-related diagnoses during COVID-19, enabling healthcare systems to communicate with patients to minimize adverse patient outcomes. Those studies demonstrate that novel informatics analysis methods can be applied to pandemic data and quickly generate clinically meaningful insights.

(d) Predictive models for EHR data. Predictive models based on EHRs data have also received great attention. Carmichael et al. [19] proposed an interesting approach to building predictive models for triage COVID-19 patient ventilator needs from past respiratory failure patients, as limited data was available at the beginning of the pandemic. They evaluated at multiple institutions and showed a reasonable performance. With the observation that some patients developed unanticipated deterioration after having relatively stable periods, Mauer et al. [20] built a predictive model of clinical deterioration for hospitalized COVID-19 patients, by incorporating hospital course trajectories. It is important to understand how predictive models can be quickly built for a new pandemic and these two studies provide some valuable insights.

4. Forecasting and epidemic modeling

In this special issue, the studies of (a) forecasting and (b) epidemic modelling either evaluated novel statistical and machine learning (ML) approaches for making short-term forecasts of epidemiological indicators, or they used epidemic models to evaluate public health and social measures, also known as non-pharmacological interventions, to control COVID-19.

(a) Forecasting. In short-term forecasting, the public health objective is to project metrics like daily new cases or hospitalizations 2–4 weeks into the future to inform operational decisions, such as managing health system response. Methods for short-term forecasting are usually data-driven, but they can also incorporate domain knowledge in the form of compartmental modelling (i.e., SIR-type models – see [48]), as a means of integrating and simultaneously forecasting in a coherent manner multiple epidemiological indicators, such as cases, hospitalizations, and deaths. The study by Swaraj et al. [21] uses a statistical data-driven approach, Aljaaf et al. [22], Dairi et al. [23], and Safari et al [24] use a machine learning data-driven approach, while Jing et al [25] uses compartmental modelling. Most of the forecasting studies were conducted as stand-alone evaluations or comparisons of methods by researchers who were not directly affiliated with public health authorities mandated to perform surveillance. In terms of epidemiological indicators being forecast, the studies focused on confirmed cases, hospitalizations and deaths, although the contribution of smartphone mapping data, was explored by Jing et al. [25] Methodologically, most studies used statistical and ML methods, either in comparison (Dairi et al [23] and Safari et al [24]), within an ensemble (Aljaaf et al [22] and Swaraj et al [21]), or as a part of a pipeline. Although results varied, findings were consistent with a recent review of methods for aberration detection in public health, with ML methods tending to outperform statistical methods for short-term forecasting. However, many of the studies in this special issue were based upon a relatively small amount of real-world data, so follow-up research to replicate and extend the findings as more data become available will be important. Another challenge with forecasting studies can be the comparison of different models and Xu et al [26] present a web-based tool for comparative visual analysis of multiple COVID-19 prediction models. In summary, the forecasting studies in this special issue provide evidence that statistical and ML models can provide accurate short-term forecasts of multiple epidemiological indicators of COVID-19. Future work should focus on extending these studies over longer intervals to encompass different epidemic dynamics and on demonstrating the connection between forecasting and decision-making.

(b) Epidemic modeling. One approach to evaluating the effect of previous and potential public health interventions to control COVID-19 is through epidemic modelling. In this application, longer-term projections (>1 month to years) are developed under different scenarios to inform the planning of public health and social measures to control and mitigate the effects of an epidemic. Given the need to evaluate specific actions and provide longer-term projections, these models usually incorporate mechanistic knowledge of epidemics through compartmental modelling. Huang et al [27] used a stratified compartmental model to evaluate the likely effect of different reopening policies on hospitalizations and deaths in Harris Country, Texas. In another study, Chen et al [28] developed an integrated framework around a compartmental model for evaluating entry restriction and travel quarantines. Finally, Ubaru et al [29] used a compartmental model with a dynamic graph model to evaluate approaches to contact tracing and testing. While the COVID-19 pandemic has stimulated extensive research and application in epidemic modelling, the studies in this special issue illustrate how these models can be used to address practical public health questions such as how best to relax public health controls and conduct contact tracing. This interface between epidemic modelling and public health decision making is a promising area for future research in biomedical informatics.

5. Text mining of literature, social media, and trial documents

As much important information of COVID-19 is embedded in narrative textual data sources such as clinical notes, biomedical literature, and social media, natural language processing (NLP) and text mining approaches have also been heavily investigated for COVID-19 research. In the previous section, we mentioned NLP approaches developed for clinical documents (e.g., COVID-19 diagnosis and symptom extraction [13]). Here we summarize NLP work for three public textual data sources: (a) biomedical literature, (b) social media, and (c) clinical trial documents.

(a) Biomedical literature. Two papers are related to the TREC-COVID Challenge, which is a shared task to evaluate information retrieval (IR) methods for finding COVID-19 related information from literature. First, Roberts et al. [30] provide an overview of the TREC-COVID Challenge, with details on data collection, corpus annotation, participating systems, and lessons learned from the challenge. Second, Chen and Hersh [31] conducted a more in-depth analysis of features used by the participating systems and identified some insights about how information retrieval techniques can help when building search systems for COVID-19. The third paper on literature mining is by Zhang et al. [32], who constructed a knowledge graph from PubMed and used it to identify drug candidates that may treat COVID-19, a novel literature-based drug repurposing approach. Such methods for finding relevant articles and automatically generating new hypotheses for scientific discoveries are fundamental research areas of biomedical informatics. As COVID-19 literature has been quickly accumulated, it is critical to develop such text mining approaches to keep researchers updated with newly available knowledge.

(b) Social media. Three papers have reported their methods and analysis on social media data such as Twitter. (1) Wu et al. [33] proposed an unsupervised approach that can efficiently mine clinically relevant information from a set of about 52 million COVID-19-related tweets, enabling real-time analysis of social media signals without much manual annotation. Two other papers have focused on detecting public perceptions on social distancing: (2) one was to detect facets of social distancing in a spatiotemporal context using US-based tweets [34] and (3) another one was to analyze social media data and news stories in Iran (Persian), to track discourse dynamics and analyze polarity from the community [35]. Methods proposed by those papers provide new ways to extract and analyze social media data during the COVID-19 pandemic, which can be applied to future pandemics, e.g., to understand public perception and to identify misinformation from social media.

(c) Clinical trial documents. Since the pandemic, many COVID-19 clinical trials have been conducted and it is important to organize information in trials to a structured format, so that they can be analyzed using computational approaches. Sun et al. [36] created an annotated corpus of eligibility criteria from 700 COVID-19 trials, to facilitate searching COVID-19 trials and developing machine learning-based information extraction systems. Such an approach can be applied to any disease to facilitate knowledge management of trials in a specific area of medicine.

6. Translational bioinformatics

Another important aspect for COVID-19 research is the exploration of biological data through translational bioinformatics approaches to understand its potential disease mechanism which can inform effective treatments. Four articles are included in the special issue: (1) Das et al. [37] proposed to infer viral-host associations based on codon usage patterns for identification of the most affected signaling pathways during COVID-19. A network analysis approach was proposed on top of the host-viral protein–protein interaction network constructed according to the similarity in codon usage patterns to understand the potential targeting mechanism of SARS-CoV-2 proteins. Many potential targets (host proteins) from the affected pathways were found to be associated with the various drug molecules that are either under clinical trial or in use during COVID-19, such as Arsenic trioxide, Dexamethasone, Hydroxychloroquine, Ritonavir, and Interferon beta. (2) Sy and Nagaraj [38] proposed a causal inference framework which relies on lossless compressors for inferring context-free grammars from sequence pairs and quantifies the extent to which the grammar inferred from one sequence compresses the other sequence. The authors applied their proposed

framework to assess the directional information exchange between genomic sequences of 16,619 SARS-CoV-2 virus isolates obtained from human samples, which presents novel opportunities for addressing key issues in sequence analysis. (3) Haneczok and Delijewski [39] investigated various ML approaches based on both shallow and deep graph neural network based molecule representations for molecular property prediction and illustrate their utility for identifying potential SARS-CoV-2 3CLpro inhibitors. (4) Yang et al. [40] proposed a hybrid transfer learning model to identify the reservoir hosts for COVID-19, which is crucial for controlling the outbreak of the epidemic. When predicting the hosts of newly discovered viruses, the proposed model utilizes the related virus domain as auxiliary domains to help build a robust model for the target virus domain, which is SARS-CoV-2 in the case of COVID-19. The authors have demonstrated the effectiveness of their proposed approach on publicly available benchmark and viral genome data. The authors have also predicted the potential animal reservoirs of COVID-19. These studies covered an important set of topics for machine learning in translational bioinformatics including deep learning, causal analysis, graph analysis, and knowledge transfer. This suggests that when facing emergencies like the COVID-19 pandemic, ML models with customized designs can be helpful in different scenarios. We should understand the advantage and applicability of these models and utilize them appropriately.

7. Other topics

There has also been research on other topics related to COVID-19 in this special issue, including (a) imaging data analysis, (b) surveillance methods and systems, (c) visualization technologies, and (d) mHealth. We summarize them below.

(a) Imaging data analysis. Rajamani et al. [41] proposed a new deep learning model called the Dynamic Deformable Attention Network (DDANet) to accurately segment the lesion regions from the lung CT scans of COVID-19 patients.

(b) Surveillance methods and systems. It is important to develop more efficient and accurate surveillance systems for pandemics. Three papers are related to this topic: (1) Wen et al. [42] explored sentinel syndromic surveillance based on deep learning for COVID-19 and other influenza-like illnesses. Their approach is based on aberration detection utilizing autoencoders that leverage symptom prevalence distributions to distinguish outbreaks of two ongoing diseases that share similar syndromes, even if they occur concurrently. (2) Savasci et al. [43] developed a surveillance system for the coronavirus pandemic with IoT applications and an inter-WBAN geographic routing algorithm, such that coronavirus symptoms can be monitored and the social distance with 'mask-wearing status' of persons can be displayed with proposed IoT software. Face masking is an effective public health policy to control the spread of COVID-19. (3) Sethi et al. [44] developed a computer vision system for real time identification of non-mask faces from public surveillance videos.

(c) Visualization technologies. Marcílio-Jr et al. [45] devised a visual analytics tool for monitoring COVID-19 dissemination for each city and its k-nearest neighbors, and validated their tool in the neighboring cities of São Paulo state, Brazil. Martinez Beltrán et al. [46] developed COnVIDa, a user-friendly tool which can easily gather multidisciplinary data related to the COVID-19 pandemic from different sources.

(d) mHealth. Alam et al. [47] examined the factors affecting the intention and actual usage behavior on mHealth adoption, investigated the effect of actual usage behavior of mHealth on mental well-being of the end-users, as well as the moderating role of self-quarantine on the intention-actual usage of mHealth under the coronavirus disease (COVID-19) pandemic situation.

These topics complement the other studies introduced in previous sections. With the ongoing pandemic, it is crucial to derive insights from various aspects so we can have a more comprehensive understanding on the situation and come up with better plans.

8. Summary

The papers collected in this special issue illustrate how informatics has played a critical role in developing computational methods to study the COVID-19 pandemic. Informatics researchers have addressed multiple aspects of COVID-19 control, including (1) exploiting and integrating diverse data types, including genomic, imaging, clinical, populational, social media, literature, and many more data sources; (2) using different technologies, including ontology, machine learning/deep learning, NLP, visualization, etc.; and (3) addressing broad applications, such as epidemic forecasting, clinical decision support, clinical research, mHealth, and drug development. Many challenges were also reported when applying digital technologies to COVID-19. Among them, trustfulness of digital assets (e.g., quality of data, information, and knowledge) and efficiency in coordinating many ongoing efforts were most prominent. These lessons have broader applicability beyond the current pandemic as many of the methods are applicable to other infectious disease outbreaks and some of the methods are even more generally applicable to the study of diseases across healthcare systems in general.

Acknowledgement

The editors of this special issue would like to express our thanks to the reviewers for their time, expertise, and effort in commenting on the submitted articles, thereby providing feedback that resulted in improved versions of the papers that were ultimately selected for inclusion.

References

[1] Coronavirus Resource Center [https://coronavirus.jhu.edu], Accessed on 1/15/2022.
[2] S. Bakken, Informatics is a critical strategy in combating the COVID-19 pandemic, J Am Med Inform Assoc 27 (6) (2020) 843–844.
[3] M. Ćurković, A. Košec, M. Roje Bedeković, V. Bedeković, Epistemic responsibilities in the COVID-19 pandemic: Is a digital infosphere a friend or a foe? J Biomed Inform 115 (2021) 103709, https://doi.org/10.1016/j.jbi.2021.103709.
[4] S.E.K. Sudat, S.C. Robinson, S. Mudiganti, A. Mani, A.R. Pressman, Mind the clinical-analytic gap: Electronic health records and COVID-19 pandemic response, J Biomed Inform 116 (2021) 103715, https://doi.org/10.1016/j.jbi.2021.103715.
[5] D.N. Sosa, B. Chen, A. Kaushal, A. Lavertu, J. Lever, S. Rensi, R. Altman, Repurposing biomedical informaticians for COVID-19, J Biomed Inform 115 (2021) 103673, https://doi.org/10.1016/j.jbi.2021.103673.
[6] M.J. Thomas, V. Lal, A.K. Baby, M. Rabeeh VP, A. James, A.K. Raj, Can technological advancements help to alleviate COVID-19 pandemic? a review, J Biomed Inform 117 (2021) 103787, https://doi.org/10.1016/j.jbi.2021.103787.
[7] O. Shahid, M. Nasajpour, S. Pouriyeh, R.M. Parizi, M. Han, M. Valero, F. Li, M. Aledhari, Q.Z. Sheng, Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance, J Biomed Inform 117 (2021) 103751, https://doi.org/10.1016/j.jbi.2021.103751.
[8] B.S. Alper, J. Dehnbostel, M. Afzal, V. Subbian, A. Soares, I. Kunnamo, K. Shahin, R.C. McClure, Initiative C-KA: Making science computable: Developing code systems for statistics, study design, and risk of bias, J Biomed Inform 115 (2021) 103685, https://doi.org/10.1016/j.jbi.2021.103685.
[9] M. Pedrera-Jiménez, N. García-Barrio, J. Cruz-Rojo, A.I. Terriza-Torres, E.A. López-Jiménez, F. Calvo-Boyero, M.J. Jiménez-Cerezo, A.J. Blanco-Martínez, G. Roig-Domínguez, J.L. Cruz-Bermúdez, J.L. Bernal-Sobrino, P. Serrano-Balazote, A. Muñoz-Carrero, Obtaining EHR-derived datasets for COVID-19 research within a short time: a flexible methodology based on Detailed Clinical Models, J Biomed Inform 115 (2021) 103697, https://doi.org/10.1016/j.jbi.2021.103697.
[10] P. Taber, C.J. Staes, S. Phengphoo, E. Rocha, A. Lam, G. Del Fiol, S.M. Maviglia, R. A. Rocha, Developing a sampling method and preliminary taxonomy for classifying COVID-19 public health guidance for healthcare organizations and the general public, J Biomed Inform 120 (2021) 103852, https://doi.org/10.1016/j.jbi.2021.103852.
[11] L. Zheng, Y. Perl, Y. He, C. Ochs, J. Geller, H. Liu, V.K. Keloth, Visual comprehension and orientation into the COVID-19 CIDO ontology, J Biomed Inform 120 (2021) 103861, https://doi.org/10.1016/j.jbi.2021.103861.
[12] S. DeLozier, S. Bland, M. McPheeters, Q. Wells, E. Farber-Eger, C.A. Bejan, D. Fabbri, T. Rosenbloom, D. Roden, K.B. Johnson, W.-Q. Wei, J. Peterson, L. Bastarache, Phenotyping coronavirus disease 2019 during a global health pandemic: Lessons learned from the characterization of an early cohort, J Biomed Inform 117 (2021) 103777, https://doi.org/10.1016/j.jbi.2021.103777.

[13] K. Lybarger, M. Ostendorf, M. Thompson, M. Yetisgen, Extracting COVID-19 diagnoses and symptoms from clinical text: A new annotated corpus and neural event extraction framework, J Biomed Inform 117 (2021) 103761, https://doi.org/10.1016/j.jbi.2021.103761.
[14] Y. Huang, X. Li, G.-Q. Zhang, ELII: A novel inverted index for fast temporal query, with application to a large Covid-19 EHR dataset, J Biomed Inform 117 (2021) 103744, https://doi.org/10.1016/j.jbi.2021.103744.
[15] T.T. Helmer, A.A. Lewis, M. McEver, F. Delacqua, C.L. Pastern, N. Kennedy, T. L. Edwards, B.O. Woodward, P.A. Harris, Creating and implementing a COVID-19 recruitment Data Mart, J Biomed Inform 117 (2021) 103765, https://doi.org/10.1016/j.jbi.2021.103765.
[16] E.J. Schenck, K.L. Hoffman, M. Cusick, J. Kabariti, E.T. Sholle, T.R. Campion, Critical carE Database for Advanced Research (CEDAR): An automated method to support intensive care units with electronic health record data, J Biomed Inform 118 (2021) 103789, https://doi.org/10.1016/j.jbi.2021.103789.
[17] J. Zhao, M.E. Grabowska, V.E. Kerchberger, J.C. Smith, H.N. Eken, QiPing Feng, J. F. Peterson, S. Trent Rosenbloom, K.B. Johnson, W.-Q. Wei, ConceptWAS: A high-throughput method for early identification of COVID-19 presenting symptoms and characteristics from clinical notes, J Biomed Inform 117 (2021) 103748, https://doi.org/10.1016/j.jbi.2021.103748.
[18] N.S. Zheng, J.L. Warner, T.J. Osterman, Q.S. Wells, X.-O. Shu, S.A. Deppen, S. J. Karp, S. Dwyer, QiPing Feng, N.J. Cox, J.F. Peterson, C.M. Stein, D.M. Roden, K. B. Johnson, W.-Q. Wei, A retrospective approach to evaluating potential adverse outcomes associated with delay of procedures for cardiovascular and cancer-related diagnoses in the context of COVID-19, J Biomed Inform 113 (2021) 103657, https://doi.org/10.1016/j.jbi.2020.103657.
[19] H. Carmichael, J. Coquet, R. Sun, S. Sang, D. Groat, S.M. Asch, J. Bledsoe, I. D. Peltan, J.R. Jacobs, T. Hernandez-Boussard, Learning from past respiratory failure patients to triage COVID-19 patient ventilator needs: A multi-institutional study, J Biomed Inform 119 (2021) 103802, https://doi.org/10.1016/j.jbi.2021.103802.
[20] E. Mauer, J. Lee, J. Choi, H. Zhang, K.L. Hoffman, I.J. Easthausen, M. Rajan, M. G. Weiner, R. Kaushal, M.M. Safford, P.A.D. Steel, S. Banerjee, A predictive model of clinical deterioration among hospitalized COVID-19 patients by harnessing hospital course trajectories, J Biomed Inform 118 (2021) 103794, https://doi.org/10.1016/j.jbi.2021.103794.
[21] A. Swaraj, K. Verma, A. Kaur, G. Singh, A. Kumar, L. Melo de Sales, Melo de Sales L: Implementation of stacking based ARIMA model for prediction of Covid-19 cases in India, J Biomed Inform 121 (2021) 103887, https://doi.org/10.1016/j.jbi.2021.103887.
[22] A.J. Aljaaf, T.M. Mohsin, D. Al-Jumeily, M. Alloghani, A fusion of data science and feed-forward neural network-based modelling of COVID-19 outbreak forecasting in IRAQ, J Biomed Inform 118 (2021) 103766, https://doi.org/10.1016/j.jbi.2021.103766.
[23] A. Dairi, F. Harrou, A. Zeroual, M.M. Hittawe, Y. Sun, Comparative study of machine learning methods for COVID-19 transmission forecasting, J Biomed Inform 118 (2021) 103791, https://doi.org/10.1016/j.jbi.2021.103791.
[24] A. Safari, R. Hosseini, M. Mazinani, A novel deep interval type-2 fuzzy LSTM (DIT2FLSTM) model applied to COVID-19 pandemic time-series prediction, J Biomed Inform 123 (2021) 103920, https://doi.org/10.1016/j.jbi.2021.103920.
[25] M. Jing, K.Y. Ng, B.M. Namee, P. Biglarbeigi, R. Brisk, R. Bond, D. Finlay, J. McLaughlin, COVID-19 modelling by time-varying transmission rate associated with mobility trend of driving via Apple Maps, J Biomed Inform 122 (2021) 103905, https://doi.org/10.1016/j.jbi.2021.103905.
[26] H. Xu, A. Berres, G. Thakur, J. Sanyal, S. Chinthavali, EPIsembleVis: A geo-visual analysis and comparison of the prediction ensembles of multiple COVID-19 models, J Biomed Inform 124 (2021) 103941, https://doi.org/10.1016/j.jbi.2021.103941.
[27] T. Huang, Y. Chu, S. Shams, Y. Kim, A.V. Annapragada, D. Subramanian, I. Kakadiaris, A. Gottlieb, X. Jiang, Population stratification enables modeling effects of reopening policies on mortality and hospitalization rates, J Biomed Inform 119 (2021) 103818, https://doi.org/10.1016/j.jbi.2021.103818.
[28] T. Chen, S. Huang, G. Li, Y. Zhang, Y.e. Li, J. Zhu, X. Shi, X. Li, G. Xie, L. Zhang, An integrated framework for modelling quantitative effects of entry restrictions and travel quarantine on importation risk of COVID-19, J Biomed Inform 118 (2021) 103800, https://doi.org/10.1016/j.jbi.2021.103800.
[29] S. Ubaru, L. Horesh, G. Cohen, Dynamic graph and polynomial chaos based models for contact tracing data analysis and optimal testing prescription, J Biomed Inform 122 (2021) 103901, https://doi.org/10.1016/j.jbi.2021.103901.
[30] K. Roberts, T. Alam, S. Bedrick, D. Demner-Fushman, K. Lo, I. Soboroff, E. Voorhees, L.L. Wang, W.R. Hersh, Searching for scientific evidence in a pandemic: An overview of TREC-COVID, J Biomed Inform 121 (2021) 103865, https://doi.org/10.1016/j.jbi.2021.103865.
[31] J.S. Chen, W.R. Hersh, A comparative analysis of system features used in the TREC-COVID information retrieval challenge, J Biomed Inform 117 (2021) 103745, https://doi.org/10.1016/j.jbi.2021.103745.
[32] R. Zhang, D. Hristovski, D. Schutte, A. Kastrin, M. Fiszman, H. Kilicoglu, Drug repurposing for COVID-19 via knowledge graph completion, J Biomed Inform 115 (2021) 103696, https://doi.org/10.1016/j.jbi.2021.103696.
[33] J. Wu, V. Sivaraman, D. Kumar, J.M. Banda, D. Sontag, Pulse of the pandemic: Iterative topic filtering for clinical information extraction from social media, J Biomed Inform 120 (2021) 103844, https://doi.org/10.1016/j.jbi.2021.103844.
[34] J. Kwon, C. Grady, J.T. Feliciano, S.J. Fodeh, Defining facets of social distancing during the COVID-19 pandemic: Twitter analysis, J Biomed Inform 111 (2020) 103601, https://doi.org/10.1016/j.jbi.2020.103601.
[35] F. Jafarinejad, M. Rahimi, H. Mashayekhi, Tracking and analysis of discourse dynamics and polarity during the early Corona pandemic in Iran, J Biomed Inform 121 (2021) 103862, https://doi.org/10.1016/j.jbi.2021.103862.
[36] Y. Sun, A. Butler, L.A. Stewart, H. Liu, C. Yuan, C.T. Southard, J.H. Kim, C. Weng, Building an OMOP common data model-compliant annotated corpus for COVID-19 clinical trials, J Biomed Inform 118 (2021) 103790, https://doi.org/10.1016/j.jbi.2021.103790.
[37] J.K. Das, S. Chakraborty, S. Roy, A scheme for inferring viral-host associations based on codon usage patterns identifies the most affected signaling pathways during COVID-19, J Biomed Inform 118 (2021) 103801, https://doi.org/10.1016/j.jbi.2021.103801.
[38] P. Sy, N. Nagaraj, Causal discovery using compression-complexity measures, J Biomed Inform 117 (2021) 103724, https://doi.org/10.1016/j.jbi.2021.103724.
[39] J. Haneczok, M. Delijewski, Machine learning enabled identification of potential SARS-CoV-2 3CLpro inhibitors based on fixed molecular fingerprints and Graph-CNN neural representations, J Biomed Inform 119 (2021) 103821, https://doi.org/10.1016/j.jbi.2021.103821.
[40] Y. Yang, J. Guo, P. Wang, Y. Wang, M. Yu, X. Wang, P.o. Yang, L. Sun, Reservoir hosts prediction for COVID-19 by hybrid transfer learning model, J Biomed Inform 117 (2021) 103736, https://doi.org/10.1016/j.jbi.2021.103736.
[41] K.T. Rajamani, H. Siebert, M.P. Heinrich, Dynamic deformable attention network (DDANet) for COVID-19 lesions semantic segmentation, J Biomed Inform 119 (2021) 103816, https://doi.org/10.1016/j.jbi.2021.103816.
[42] A. Wen, L. Wang, H. He, S. Liu, S. Fu, S. Sohn, J.A. Kugel, V.C. Kaggal, M. Huang, Y. Wang, F. Shen, J. Fan, H. Liu, An aberration detection-based approach for sentinel syndromic surveillance of COVID-19 and other novel influenza-like illnesses, J Biomed Inform 113 (2021) 103660, https://doi.org/10.1016/j.jbi.2020.103660.
[43] S. Savaşcı Şen, M. Cicioğlu, A. Çalhan, IoT-based GPS assisted surveillance system with inter-WBAN geographic routing for pandemic situations, J Biomed Inform 116 (2021) 103731, https://doi.org/10.1016/j.jbi.2021.103731.
[44] S. Sethi, M. Kathuria, T. Kaushik, Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread, J Biomed Inform 120 (2021) 103848, https://doi.org/10.1016/j.jbi.2021.103848.
[45] W.E. Marcilio-Jr, D.M. Eler, R.E. Garcia, R.C.M. Correia, R.M.B. Rodrigues, Visual analytics of COVID-19 dissemination in Sao Paulo state, Brazil. *J Biomed Inform* 117 (2021), 103753.
[46] E.T. Martínez Beltrán, M. Quiles Pérez, J. Pastor-Galindo, P. Nespoli, F.J. García Clemente, F. Gómez Mármol, COnVIDa: COVID-19 multidisciplinary data collection and dashboard, J Biomed Inform 117 (2021) 103760, https://doi.org/10.1016/j.jbi.2021.103760.
[47] M.M.D. Alam, M.Z. Alam, S.A. Rahman, S.K. Taghizadeh, Factors influencing mHealth adoption and its impact on mental well-being during COVID-19 pandemic: A SEM-ANN approach, J Biomed Inform 116 (2021) 103722, https://doi.org/10.1016/j.jbi.2021.103722.
[48] W.M. Getz, R. Salter, O. Muellerklein, H.S. Yoon, K. Tallam, Modeling epidemics: A primer and Numerus Model Builder implementation, Epidemics 25 (2018) 9–19.

Hua Xu
School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

David L. Buckeridge
Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

Fei Wang
Department of Population Health Sciences, Cornell University, New York, NY, USA

Peter Tarczy-Hornoch
Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA