

Editorial

Biomedical and health informatics approaches remain essential for addressing the COVID-19 pandemic

Suzanne Bakken 

School of Nursing, Department of Biomedical Informatics, Data Science Institute Columbia University, New York, New York, USA

***Corresponding Author:** Suzanne Bakken, PhD, RN, FAAN, FACMI, FIAHSI, School of Nursing, Department of Biomedical Informatics, Data Science Institute Columbia University, 630 W. 168th Street, New York, USA (sbh22@cumc.columbia.edu)

Received 13 January 2021; Editorial Decision 13 January 2021; Accepted 14 January 2021

As this March issue of the Journal of the American Medical Informatics Association (JAMIA) is published, we have experienced a year of sheltering-in-place, wearing masks, frequent handwashing, and COVID-19 testing. For some of us, this year also included COVID-19 infection and loss of family, friends, and colleagues—and more recently, COVID-19 vaccination. I accepted JAMIA's first COVID-related paper on March 19, 2020 less than 24 hours after its submission,¹ and the accepted version was available online within days. To date, JAMIA has received almost 400 COVID-related submissions and published 69 contributions in issues since June 2020.

The March issue includes 7 papers and correspondence related to COVID-19 including a call from World Health Organization authors to strengthen data in response to COVID-19 and beyond,² a report on virtual care expansion in the Veterans Health Administration,³ and correspondence about telemedicine, privacy, and information security in the age of COVID-19.⁴ In this editorial, I highlight 5 papers that reflect the breadth of biomedical and health informatics approaches to addressing the COVID-19 pandemic.

Haendel and colleagues provide an overview of the rationale, design, infrastructure, and deployment of the National COVID Cohort Collaborative (N3C).⁵ N3C (covid.cd2h.org), an open science community focused on analyzing individual-level data from many centers, was developed by the Clinical and Translational Award Program, the National Center for Translational Sciences, and the scientific community to enable rapid collaboration among clinicians, researchers, and data scientists to identify treatments and specialized care and subsequently reduce the immediate and long-term consequences of COVID-19. To overcome technical, regulatory, policy, and governance barriers to sharing and harmonizing individual-level clinical data, N3C developed (a) legal agreements and governance for organizations and researchers; (b) data extraction scripts to iden-

tify and ingest positive, negative, and possible COVID-19 cases; (c) a data quality assurance and harmonization pipeline to create a single harmonized dataset; (d) a secure data enclave with data, machine learning, and statistical analytics tools; (e) dissemination mechanisms; and (f) a synthetic data pilot to democratize data access. There are 3 datasets for analysis: synthetic, deidentified, and limited. The Attribution and Publication Policy encompasses all N3C contributions as reflected in the author contribution statement for this paper. Analyses posted within the N3C enclave leverage the contributor attribution model to track the transitive credit of all upstream contributors. The N3C infrastructure is designed to be scalable and extensible to other topics.

Sun and colleagues designed COVID-19 Trial Finder, an open-source semantic search engine, to facilitate patient-centered search of COVID-19 trials.⁶ It is powered by a machine-readable dataset for all COVID-19 trials in the United States. COVID-19 Trial Finder also includes a web-based visualization of the geographic distribution of COVID-19 trials. The initial search is by location and radius distance from trial sites; this is refined through a set of dynamically generated medical questions to assess patient eligibility for nearby COVID-19 trials. They assessed the precision of COVID-19 Trial Finder for COVID-19 using 20 case reports from LitCOVID. Overall precision across the 20 cases was 79.76%, although it varied widely across cases. The major factor contributing to imprecision was the inability to generate relevant questions in some instances which prevented filtering out irrelevant trials. The system (<https://covidtrialx.dbmi.columbia.edu>) and its source code (<https://github.com/WengLab-InformaticsResearch/COVID19-TrialFinder>) are accessible online.

Hassandoust, Akhaghpour, and Johnson examine individual privacy concerns and intention to adopt contact tracing mobile applications through a situational privacy calculus model.⁷ Using structural equation modeling and a national sample (N = 853) of survey

respondents, they found that risk beliefs, perceived individual and societal benefits to public health, privacy concerns, privacy protection initiatives (legal and technical protection), and technology features (anonymity and use of less sensitive data) influenced intention to install a contact tracing mobile application. The relationship between trust in public health authorities and intention was indirect. Sex, education, and past invasion of privacy had no significant influence on intention. Study findings provide the foundation for actions to address the identified factors and, subsequently, increase adoption of contact tracing mobile applications.

Oiao and colleagues describe and evaluate the Focal Loss bAsed Neural Network EnsemblE (FLANNEL) approach for COVID-19 detection in chest x-ray images.⁸ They constructed a dataset from 2 publicly available sources comprising 5508 chest x-ray images for 2874 patients with 4 classes of diagnoses: normal, bacterial pneumonia, non-COVID-19 viral pneumonia, and COVID-19 pneumonia. The addition of focal loss to the neural network approach was designed to address class imbalance. FLANNEL consistently outperformed baseline models in COVID-19 identification achieving a precision of 0.78, recall of 0.86, and F1 of 0.82. These promising findings highlight the potential of the FLANNEL approach for differentiating COVID-19 from other types of pneumonias.

To estimate the hospitalization risk for people with comorbidities infected by SARS-CoV-2, Gao and Dong applied a Bayesian approach designed to overcome challenges with traditional biostatistical requirements for risk estimates.⁹ These include data about the number of infected people who were not hospitalized and their comorbidities. The former is particularly problematic for COVID-19, where many individuals are asymptomatic, and not all those infected have been tested. Using 2 large-scale datasets (2491 patients from COVID-NET and 5700 patients from New York hospitals), they estimated the posterior distribution of the risk ratio using the observed frequency of comorbidities in hospitalized COVID-19 patients and the prevalence of comorbidities in the general population. They found that cardiovascular diseases carried the highest hospitalization risk for COVID-19 patients, followed by diabetes, chronic respiratory disease, hypertension, and obesity. This innovative approach has the potential to assist with resource planning to manage the COVID-19 pandemic.

There is no doubt that the COVID-19 pandemic has highlighted the need for and visibility of biomedical and health informatics. For

JAMIA, this is reflected in a dramatic increase in Altmetric Attention Score which represents the mentions that papers are receiving in news outlets and social media. As always, I hope that you will consider JAMIA as a publication venue for informatics and data science research relevant to the challenging public health problems facing our world.

CONFLICT OF INTEREST

None declared.

REFERENCES

1. Reeves JJ, Hollandsworth HM, Torriani FJ, *et al.* Rapid response to COVID-19: health informatics support for outbreak management in an academic health system. *J Am Med Inform Assoc* 2020; 27 (6): 853–9.
2. Azzopardi-Muscat N, Kluge HHP, Asma S, *et al.* A call to strengthen data in response to COVID-19 and beyond. *J Am Med Inform Assoc* 2021; 28 (3): 638–9.
3. Ferguson JM, Jacobs J, Yefimova M, *et al.* Virtual care expansion in the Veterans Health Administration during the COVID-19 pandemic: clinical services and patient characteristics associated with utilization. *J Am Med Inform Assoc* 2021; 28 (3): 453–62.
4. Jalali MS, Landman A, Gordon WJ. Telemedicine, privacy, and information security in the age of COVID-19. *J Am Med Inform Assoc* 2021; 28 (3): 671–2.
5. Haendel MA, Chute CG, Bennett TD, *et al.* The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. *J Am Med Inform Assoc* 2021; 28 (3): 427–43.
6. Sun Y, Butler A, Lin F, *et al.* The COVID-19 Trial Finder. *J Am Med Inform Assoc* 2021; 28 (3): 616–21.
7. Hassandoust F, Akhaghpour S, Johnson AC. Individuals' privacy concerns and adoption of contact tracing mobile applications in a pandemic: a situational privacy calculus perspective. *J Am Med Inform Assoc* 2021; 28 (3): 463–71.
8. Oiao Z, Bae A, Glass LM, *et al.* FLANNEL (Focal Loss bAsed Neural Network EnsemblE) for COVID-19 detection. *J Am Med Inform Assoc* 2021; 28 (3): 444–52.
9. Gao X, Dong Q. A Bayesian framework for estimating the risk ratio of hospitalization for people with comorbidity infected by SARS-CoV-2 virus. *J Am Med Inform Assoc* 2021; 28 (3): 472–6.