2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Decision Support Component for the Localized Epidemiological Modelling of COVID-19

Philip Ciunkiewicz
Biometric Technologies Laboratory,
Department of Biomedical Engineering,
University of Calgary, Calgary, AB, Canada
Correspondence: philip.ciunkiewicz@ucalgary.ca

Svetlana Yanushkevich
Biometric Technologies Laboratory,
Department of Electrical and Software Engineering,
University of Calgary, Calgary, AB, Canada
Correspondence: syanshk@ucalgary.ca

***Abstract*—This study develops a decision support system for localized epidemiological modelling of infectious disease spread. We propose a Bayesian network topology for performing inference supplementary to an epidemiological simulation framework and a cohesive integration of this decision support system with the framework. The Bayesian network topology is structured with data defined as inputs, outputs, or derived features within the simulation framework. All features are motivated by their clinical relevance and utility for administrative policy guidance. Edges in the final network are quantitatively assessed using structural equation modelling to ensure strong causal connections. Various inference scenarios are demonstrated to provide proof of concept for real-world application and validation in future directions. The outcomes of this project contribute to a larger body of work for infectious disease risk mitigation and emergency management in generalized environments.**

***Index Terms*—Agent-based simulation, decision support, machine reasoning, risk, epidemiological model, COVID-19, causal network, Bayesian network, structural equation model**

I. Introduction

Infectious respiratory diseases, such as the SARS-CoV-2 virus causing COVID-19, can spread rapidly, leading to severe global outbreaks and millions of attributed deaths. To combat such diseases, previous studies have shown the efficacy of both non-pharmacological interventions (masking, physical distancing, capacity reduction, sanitation) and pharmacological interventions (vaccination) in mitigating spread and associated public health risks [1].

In a previous study, a robust epidemiological simulation framework which provides insights to facility administration and policy-makers was developed [2]. This simulation framework uses agent-based modelling techniques in highly localized environments such as long-term care facilities or university campuses to deliver specialized insights in each scenario. As illustrated in Figure 1, scenarios are broadly configurable and designed with generalizability as a core software engineering principle. This work expands on that previous study to provide decision support capabilities to the epidemiological simulation framework results.

Agent-based modelling techniques are broadly used and highly effective in epidemiological contexts due to their ability to explicitly model underlying mechanics of infectious disease transmission [3]–[6]. Such techniques translate well for modelling contact-driven and airborne disease spread, such as COVID-19, as motivated in previous work [2]. Agent-based techniques contrast against differential-equation-based compartmental models or fractional mathematical models used to represent broad populations [7]. Highly localized modelling further benefits from an agent-based approach, allowing for individual behaviour and interactions between agents to reveal emergent properties in the system.

The study focuses on developing a decision support system (DSS) that uses computational tools to aid human decision-making by supplying pertinent information and contextual analysis. The specific underlying model utilized for this work is the Bayesian network (BN), chosen for its proficiency in modelling various systems due to its probabilistic framework, functionality with limited or missing data, and clear handling of uncertainty [8]–[10]; through this model, the system is able to estimate probabilities and associated risk of different epidemiological outcomes.

This paper is structured as follows: the proposed methodology is described in Section II, experimental results, observations, and discussion are provided in Section III, and Section IV summarizes our findings and future directions.

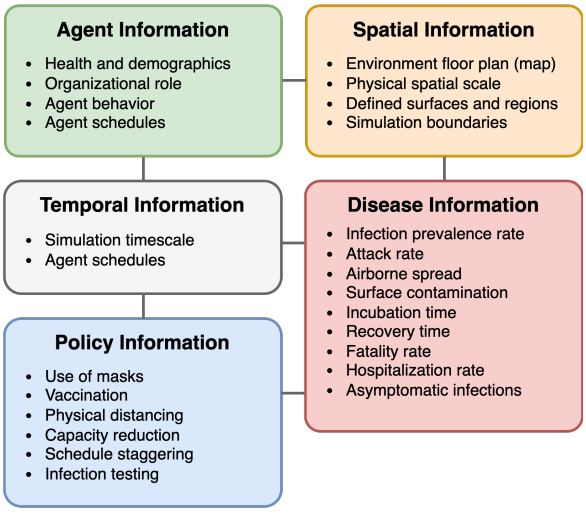

Fig. 1. Configurable parameters for the simulation in the published work [2]. Each labelled group of parameters is independently configurable; however, certain parameters are closely related and influence each other (such as temporal and agent information for scheduling).

II. Methodology

This paper implements a novel approach for epidemiological decision support in highly localized modelling contexts. We extend a previously developed epidemiological simulation framework by designing a causal network for risk assessment on top of simulated results [2]. Decision support capabilities are designed to provide facility administration and policy-makers with the prospective insights required to manage risk effectively in respective localized environments. Simulation inputs and outputs are used along with derived features to achieve this goal.

A. Summary of Previous Work

Full implementation methodology and open-source code are available in the previous publication [2]. Specific details are beneficial to the reader for contextualizing the work proposed in this paper.

1) Agents and Agent Behaviour: The atomic unit of information within the simulation is the agent. Agents represent individuals in a given environment and are assigned behavioural patterns, schedules, demographic, and epidemiological attributes. Figure 1 illustrates the components of agent information contributing to their operation within the simulation. Along with this data, agents also track their epidemiological status for modelling disease spread as shown in Figure 2. The current epidemiological status of an agent influences their behaviour and susceptibility to infection.

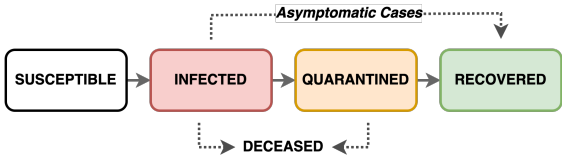

Fig. 2. Possible epidemiological statuses for agents, as defined in the previously published work [2]. Solid arrows represent typical disease progression; dotted arrows represent rare cases.

2) Infection Portrait: The infection portrait defines all aspects of the infectious disease required for effective simulation. This includes disease transmission mechanisms, population infection statistics, airborne viral particle mechanics, and disease progression. Disease lifecycle modelling and transmission are based on the Susceptible-Infected-Recovered (SIR) scheme defined in Equation (1) [11].

$$\frac{dS}{dt} = -\frac{\beta I S}{N}; \qquad \frac{dI}{dt} = \frac{\beta I S}{N} - \gamma I; \qquad \frac{dR}{dt} = \gamma I. \quad (1)$$

Parameters S, I, and R denote the susceptible, infected, and recovered populations. Parameter N represents the total population count. Coefficients β and γ are epidemiological parameters representing infection and recovery characteristics with the ratio $R_0 = \beta/\gamma$ defining the basic reproduction number. We extend this model to include an additional "Quarantined" state where the agent is temporarily removed from the model and a more general "Removed" state for agents who do not interact epidemiologically. "Removed" agents include recovered, hospitalized, and deceased agents.

Simulated disease progression is stochastic. State transitions are based on data observed in the literature and modelled by sampling log-normal distributions, defined for variable X in Equation (2) [4], [12]–[14].

$$X = e^{\mu + \sigma Z} \quad (2)$$

Parameter Z is a standard normal variable, and μ, σ are the respective mean and standard deviation of the natural logarithm of X instead of X itself. As a probability density function, $f_X(x)$ is rewritten as

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \quad (3)$$

State transitions are considered for asymptomatic cases, long-COVID cases, severe cases and hospitalizations, case fatalities, presymptomatic windows, and recovery windows.

B. Decision Support Features

The data for performing decision support in this study is constrained to the data provided by the previously developed simulation framework [2]. This limitation results from structuring our decision support system (DSS) to rely only on simulation data without needing real-world observations beyond configuring the simulation scenarios. Such a design allows for greater generalizability in arbitrary downstream applications and greatly simplified deployment. Furthermore, epidemiologically relevant information can be introduced to the simulation framework before use in decision support for benefit across both systems.

Features available for decision support can be categorized into three groups: simulation inputs, outputs, and derived features. Simulation inputs include parameters described in Figure 1, particularly elements of agent, disease, and policy information. Simulation outputs include infection statistics, hospitalizations, fatalities, and other reported epidemiological outcomes. Prominent factors in the simulation for determining epidemiological outcomes are itemized in Table I. Derived features are described in further detail.

1) Prevention Index: In previous work, the "Prevention Index" was defined. This derived feature is a compound measure quantifying how well an agent is protected against infectious disease transmission. A value of 0.0 corresponds to zero additional protection against infection, while a value of 1.0 corresponds to complete immunity. Factors such as agent masking habits, vaccination type, and vaccination doses contribute to the prevention index. The efficacy of the reported pharmacological and non-pharmacological intervention is described in Table II [2]. In this work, we expand the prevention index to include factors such as agent age, sex, and immunocompromisation [14].

2) Excess Risk of Infection: In previous work, the excess risk of infection was reported as a clinically relevant contextual reframing of the infection prevalence rate (IPR) for decision support. This quantity allows for a more accurate relative comparison of risk mitigation strategies simulated by the framework. Such relative comparison is important for decision support as it better isolates local policy effects from population statistics.

TABLE I
AVAILABLE SIMULATION FEATURES USED FOR DECISION SUPPORT. THE ALIAS COLUMN REPRESENTS FEATURE NAMES DEFINED IN PYTHON. ALIASES MARKED AS – DENOTE LATENT VARIABLES.

Category	Feature	Alias	Type	Description
Inputs	Immunocompromised	`immunocompromised`	Boolean	Agent immunocompromisation status
Inputs	Sex	`sex`	Categorical	Agent sex
Inputs	Age	`age`	Numeric	Agent age
Inputs	Organizational Role	`role`	Categorical	Agent role within organization
Inputs	Masking	`mask`	Categorical	Masking policy enforcement
Inputs	Vaccination	`vax`	Categorical	Vaccination policy enforcement
Inputs	Distancing	`distancing`	Numeric	Distancing policy enforcement
Inputs	Capacity	`capacity`	Numeric	Capacity policy enforcement
Outputs	Total Infections	–	Numeric	Total infections at scenario end
Outputs	Total Long-COVID cases	–	Numeric	Total long-COVID cases at scenario end
Outputs	Agent Long-COVID	`long_covid`	Boolean	Agent experiencing long-covid case
Outputs	Total Hospitalizations	–	Numeric	Total hospitalizations at scenario end
Outputs	Agent Hospitalized	`hospitalized`	Boolean	Agent hospitalized during infection
Outputs	Total Fatalities	–	Numeric	Total fatalities at scenario end
Outputs	Agent Deceased	`deceased`	Boolean	Agent deceased during infection
Derived	Prevention Index	`prevention_index`	Numeric	Agent composite prevention index
Derived	Excess Risk of Infection	`infected`	Boolean	Agent risk of infection in excess of population IPR

TABLE II
PREVENTION INDEX VALUES FOR MASKING AND VACCINATION.

Mask			Vaccine		
-			1 Dose	2 Dose	
Cloth	Surgical	N95	Any	AstraZeneca	Pfizer / Moderna
0.3	0.5	0.85	0.31	0.67	0.88

C. Generating Data

The dataset used for modelling is generated using the epidemiological simulation framework. The simulated scenarios all use the research lab environment established for proof of concept in the previous work [2]. Agent behaviour is designed to replicate lab researchers' schedules and patterns. Scenarios were run for ∼210 hours of contact (5-second simulated temporal resolution for 150,000 time-steps), or equivalently 2.5 weeks simulating 12-hour days between 7am-7pm. Full environmental disinfection occurs after 7pm each day while agents are absent from the environment. These scenario run parameters reflect those in previous work and are used as a basis for future comparison.

A total of 128 unique scenarios were simulated, representing a combinatorial product of the following parameters:

- Masking:
  - No masking ("nomask"),
  - Cloth masks ("cloth"),
  - Surgical masks ("surgical"),
  - N95 respirators ("n95").
- Vaccination:
  - No vaccination ("novax"),
  - One dose any vaccine ("1dose"),
  - Two doses Oxford AstraZeneca ("astra"),
  - Two doses Pfizer / Moderna MRNA ("mrna").
- Capacity Reduction: 2, 4, 6, and 8 agents
- Physical Distancing:
  - Enforced distancing ("True")
  - No distancing ("False")

The 128 unique scenarios were each simulated over an ensemble of 2,500 runs to generate meaningful statistics, resulting in 320,000 samples as input evidence for our BN model.

1) Feature Pre-Processing: Data processing is a necessary component of many modelling workflows. The nature of data processing will change depending on the specific algorithm(s) being used for modelling. This decision support application implements a BN, which requires data to be discretized to compute the underlying conditional probabilities effectively. We apply dynamic discretization techniques to iteratively optimize the discretized binning for each individual feature based on the underlying distribution [15]. This approach has been shown to produce data discretizations with significantly reduced relative entropy error E (discrete analog to Kullback–Leibler divergence), defined in Equation (4) [16].

$$E_j = \left[\frac{f_{max} - \bar{f}}{f_{max} - f_{min}} f_{min} \log \frac{f_{min}}{\bar{f}} + \frac{\bar{f} - f_{min}}{f_{max} - f_{min}} f_{max} \log \frac{f_{max}}{\bar{f}} \right] |w_j| \tag{4}$$

Here $|w_j|$ denotes the length of the discretized interval j, and f_{min}, f_{max}, $\bar{f}$ denote the value of the underlying distribution $f(x)$ at the interval endpoints and midpoint, respectively.

D. Bayesian Network Modelling

An initial expert-knowledge-driven BN topology was developed based on fundamental epidemiology concepts and the construction of the simulation framework. This initial topology was then iteratively refined using quantitative validation techniques. The final network topology is structured to retain the known underlying flow of information in the simulation, from inputs to derived features, and lastly to outputs. The input nodes of Sex, Masking, Vaccination, Age, and Immunocompromised are connected directly to the Prevention Index. The Prevention index and input features of

Organizational Role, Distancing, and Capacity are connected to the Excess Risk of Infection. The Excess Risk of Infection ultimately connects to the output features of Infections, Long-COVID Cases, Hospitalizations, and Fatalities. This topology is illustrated in Figure 3.

Various inference tasks can be performed using the final BN with computed conditional probabilities. These inference tasks represent the decision support capabilities of the model, and they provide end-users with epidemiological insights to guide policy choices.

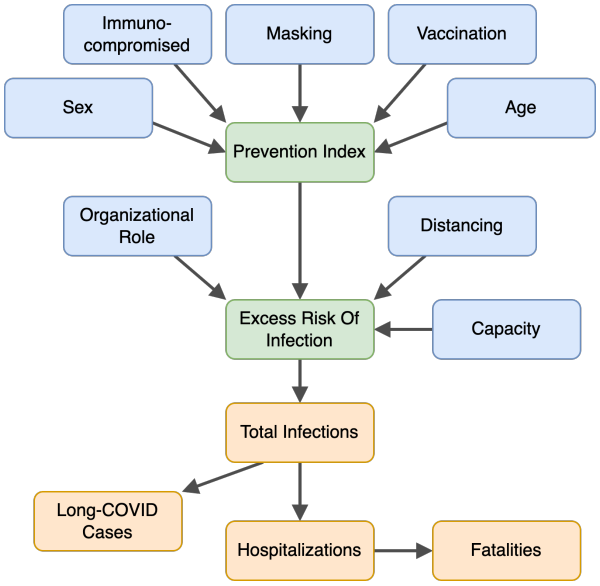

Fig. 3. Bayesian network topology for providing decision support to the epidemiological simulation framework. Blue nodes denote input features, green nodes denote derived features, and orange nodes denote output features.

E. Bayesian Network Topology Validation

Structural equation modelling (SEM) validates and motivates the network topology defined for decision support [17]–[19]. Causal relationships in an underlying system are modelled by relationships structured in a BN if and only if the BN structure of the system is equivalent to a valid SEM structure of the same underlying system [20]. SEM broadly describes techniques and approaches for estimating causal relationships between variables, combining elements from regression, path analysis, and factor analysis [19]. SEM techniques can be used as a means to evaluate causal relationships quantitatively. As with BNs, variables in SEM can be either latent or explicit, providing greater flexibility in modelling. Latent variables in SEM can be expressed as a combination of one or more explicit variables. In conjunction with our expert-knowledge-driven initial topology, the causal connections are iteratively updated using SEM.

In the literature, SEM provides a robust framework for quantifying and testing theoretical graphical models such as BN topologies. More directly, causal relationships are evaluated when observation data is available by comparing the observed covariance matrix to the expected covariance matrix from the hypothesized model. This null hypothesis H_0 is expressed mathematically by Equation (5) [18].

$$H_0 : \Sigma = \Sigma(\theta) \tag{5}$$

Here Σ denotes the observed covariance matrix computed from data, $\Sigma(\theta)$ represents the expected covariance matrix from the hypothesized model, and θ is a vector containing the model parameters. By comparing the covariance matrices in Equation (5) and failing to reject the null hypothesis, the observed data would support the theorized model and imply causality in the structure. Accepting the null hypothesis is the objective of applying SEM.

SEM mathematically operates by optimizing the model parameters such that the null hypothesis can be accepted. The model parameters are estimated using techniques such as maximum likelihood or least squares to minimize the error between the observed covariance matrix and the model-estimated covariance matrix. The estimated parameters for each edge in the SEM graph can be inspected along with their corresponding p-value, representing the probability of rejecting the null hypothesis [19].

III. Results

Results from this study include the formulation and optimization of a novel BN topology for providing decision support with epidemiological simulation outputs and the quantitative assessment of causal connections within this theorized model. We also demonstrate the utility of inference with such a network and how decision support is achieved in practical deployments.

A. Bayesian Network Inference

Figure 4 illustrates the final BN topology and calculated conditional probabilities at each node in the graph. These conditional probability tables (CPTs) include numeric variables `age` and `prevention_index`, which have been dynamically discretized for modelling. By construction, many of the CPTs, such as `mask`, `vax`, and `distancing`, are balanced as a result of simulation and scenario configuration. The `capacity` and `role` CPTs are also distributed based on the scenario configuration. CPTs including `age`, `sex`, and `immunocompromised` are based on census data and distributions described in the literature [14]. Before simulating any scenarios, all CPTs corresponding to these input features are known. CPTs for simulation outputs and derived features cannot be calculated before running the simulations, and thus are more variable in nature. Notably missing from the network are the "total" variables (infections, long-covid cases, hospitalizations, and fatalities), which we treat as latent variables due to the formulation of our data as agent-based, with each sample representing an individual.

From Figure 4, we can derive the excess risk of infection by observing the proportion of agents infected during simulation runtime. The overall probability of infection is 0.18%. Compared to the simulated population infection prevalence rate of 2.10%, the 0.18% probability of infection translates into a relative excess infection risk of infection of 8.57%.

When considering long-covid cases, we see that the overall proportion of positive results is 0.40%, higher than the overall probability of infection. This result is a consequence of the `infected` field exclusively tracking new infections which occur within the simulation runtime. Agents infected outside the simulation can still experience long-covid, however, they are intentionally omitted from the `infected` statistics. The same outcomes are true for hospitalizations, at 0.30% positive results, and fatalities, at 0.04% positive results. Each of these results represents the average prevalence across 128 distinct simulation scenarios, and while informative, these raw values are less useful for decision support than BN inference with provided evidence.

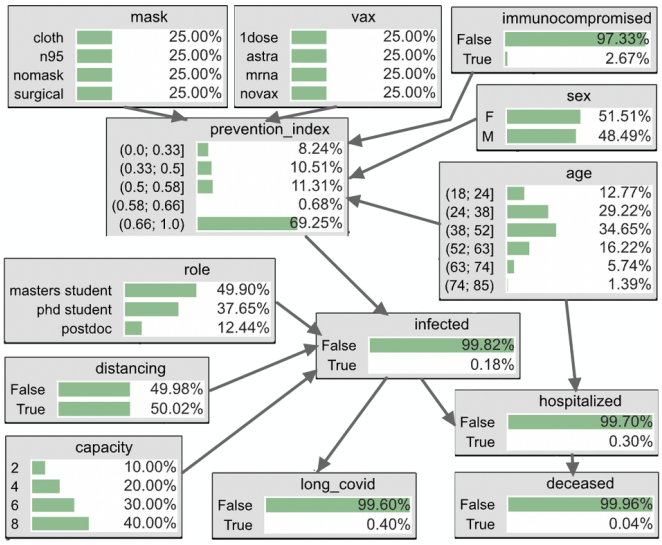

Fig. 4. Bayesian network conditional probability tables calculated from the simulated scenario data. Observations for parameter learning are aggregated from 320,000 scenario samples across 128 distinct scenario configurations.

Simulation outputs and derived features provide objectives for decision support, such as minimizing the excess risk of infection or minimizing hospitalizations and fatalities. Operational considerations such as facility capacity can provide challenging and contradictory objectives for mitigating epidemiological risk. These challenging objectives provide great value to administrators and extend the potential utility provided by the proposed DSS.

Performing BN inference by providing evidence allows for more targeted insights and stronger decision support. In this example, we provided evidence to the BN for the following features: `mask="nomask"`, `vax="novax"`, `age="(24;38]"`, and `capacity="8"`. The masking, vaccination, and capacity evidence represent an epidemiological worst-case scenario for infectious disease spread. While maintaining the largest occupancy, no preventative measures are enforced, resulting in a probability of infection of 1.20%. This probability corresponds to a relative excess risk of infection of 57.14%, dramatically higher than the overall average. With the higher risk of infection, downstream nodes such as long-covid cases, hospitalizations, and fatalities also increase in prevalence to 0.56%, 0.38%, and 0.04%, respectively. Facility administrators can directly apply this information to determine an appropriate risk exposure and compromise between policy enforcement and implementation difficulty or resource cost.

Numerous exercises can be performed similarly. Clear choices include varying the preventative measures and tracking the infected proportion, observing the effect of capacity and age on hospitalizations, and investigating the effect of organizational roles on epidemiological outcomes. Sensitivity analysis can be performed on such results to quantify the relationships between given parent and child nodes, further contributing to decision support capabilities.

B. Causality

Table III details the results of SEM analysis using the final BN topology as our hypothesized model. For each edge in the graph, the strength of the causal connection between the parent and child node is estimated by the SEM model, and this value is reported in the "Estimate" column. Causal relations with the highest strength include $vax \rightarrow prevention\_index$ at 0.69, $mask \rightarrow prevention\_index$ at 0.58, and $age \rightarrow prevention\_index$ at -0.31. Here, negative values represent an inverse relation; zero represents a complete lack of relation, and a magnitude of 1.0 represents a perfect correlation. Certain relations showed very low estimates below 0.01; however, only $role \rightarrow infected$ exhibited a correspondingly poor p-value.

The probability of rejecting the null hypothesis is also provided under the "P-Value" column. P-values lower than 0.05 are generally considered to be statistically significant [19]. Of the 13 causal relations in our topology, nine have a p-value smaller than 0.01, meaning that a majority of the causal connections in our network have a $> 99\%$ probability of accepting the null hypothesis. Only one causal connection, $role \rightarrow infected$, has a p-value larger than 0.05. Based on domain expertise and the known behaviour of the simulation framework, this connection remained part of the network topology despite not returning a statistically significant p-value.

TABLE III
SEMOPY MODEL INSPECTION OUTPUT.

Parent Node	Child Node	Estimate	P-Value
age	prevention_index	-0.31	<0.01
vax	prevention_index	0.69	<0.01
mask	prevention_index	0.58	<0.01
sex	prevention_index	<0.01	<0.01
immunocompromised	prevention_index	<0.01	0.02
prevention_index	infected	-0.05	<0.01
role	infected	<0.01	0.09
capacity	infected	0.03	<0.01
distancing	infected	-0.02	0.03
infected	hospitalized	0.14	<0.01
age	hospitalized	0.02	<0.01
infected	long_covid	0.10	<0.01
hospitalized	deceased	-0.01	0.05

IV. Conclusion and Future Work

In this paper, we extended an existing epidemiological simulation framework to provide a layer of decision support for facility administrators and policy-makers. The implemented DSS comprised a BN with SEM techniques applied to validate the network topology quantitatively. The final validated topology with learned CPTs provides general insights into the underlying distributions of data in the model. Performing inference with the BN by providing evidence extends these insights and allows for comprehensive decision support at the policy level. This decision support can be performed across varied simulation scenarios to generate diverse insights, as demonstrated in our inference examples. Our causality validation using SEM found that all but one of the causal relations in our final network have a statistically significant p-value lower than 0.05. This work proposes a novel BN topology to extend the localized epidemiological simulation of COVID-19 with decision support capabilities validated using SEM techniques.

Real-world validation is a critical component required for further development of this technology. We work closely with the Brenda Strafford Foundation (BSF), a non-profit organization specializing in long-term care services. The BSF operates multiple long-term care facilities and has ongoing research relationships with our institution. The aim for real-world validation is to partner with the BSF and collect the required information to perform localized simulations using the framework. These simulations would use the day-to-day operation of long-term care centres as the base scenarios, with facility policies and epidemiological outcomes tracked over time. Simulation results and decision support would be reviewed by researchers and BSF facility administrators, and the results would be published as part of a larger collaborative investigation. Data for this validation component has been collected, and ethics approval for the study has been granted. Our validation strategy includes benchmarking the simulation results against observed statistics for aggregate epidemiological measures such as the localized excess risk of infection and case counts. Temporal characteristics are also being investigated to evaluate the validity of outbreak incidence and progression.

From the generalizability of our simulation framework and DSS, studying infectious diseases beyond COVID-19 in arbitrary environments is planned following validation.

Acknowledgements

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) One Health Modelling Network for Emerging Infections, part of the NSERC-PHAC funded Emerging Infectious Diseases Modelling Initiative; the Brenda Strafford Foundation; the Alberta Innovates Graduate Student Scholarship Ph.D. program; and the NSERC Canada Graduate Scholarships - Ph.D program.

References

[1] M. Qian and J. Jiang, "COVID-19 and social distancing," *Journal of Public Health*, vol. 30, no. 1, pp. 259–261, Jan. 2022.

[2] P. Ciunkiewicz, W. Brooke, M. Rogers, and S. Yanushkevich, "Agent-based epidemiological modeling of COVID-19 in localized environments," *Computers in Biology and Medicine*, vol. 144, p. 105396, May 2022.

[3] E. Cuevas, "An agent-based model to evaluate the COVID-19 transmission risks in facilities," *Computers in Biology and Medicine*, vol. 121, p. 103827, Jun. 2020.

[4] L. K. N. Nguyen, S. Howick, D. McLafferty, G. H. Anderson, S. J. Pravinkumar, R. V. D. Meer *et al.*, "Evaluating intervention strategies in controlling coronavirus disease 2019 (COVID-19) spread in care homes: An agent-based model," *Infection Control & Hospital Epidemiology*, pp. 1–11, Dec. 2020.

[5] T. N. Vilches, S. Nourbakhsh, K. Zhang, L. Juden-Kelly, L. E. Cipriano, J. M. Langley *et al.*, "Multifaceted strategies for the control of COVID-19 outbreaks in long-term care facilities in Ontario, Canada," *Preventive Medicine*, vol. 148, p. 106564, Jul. 2021.

[6] N. Mahdizadeh Gharakhanlou and N. Hooshangi, "Spatio-temporal simulation of the novel coronavirus (COVID-19) outbreak using the agent-based modeling approach (case study: Urmia, Iran)," *Informatics in Medicine Unlocked*, vol. 20, p. 100403, Jan. 2020.

[7] P. Kumar, V. S. Erturk, K. S. Nisar, W. Jamshed, and M. S. Mohamed, "Fractional dynamics of 2019-nCOV in Spain at different transmission rate with an idea of optimal control problem formulation," *Alexandria Engineering Journal*, vol. 61, no. 3, pp. 2204–2219, Mar. 2022.

[8] M. Neil, N. Fenton, M. Osman, and S. McLachlan, "Bayesian network analysis of Covid-19 data reveals higher infection prevalence rates and lower fatality rates than widely reported," *Journal of Risk Research*, vol. 23, no. 7-8, pp. 866–879, Aug. 2020.

[9] P. Ciunkiewicz, M. Roumeliotis, K. Stenhouse, P. McGeachy, P. Grendarova, S. Quirk *et al.*, "Modelling Toxicity Risk and Uncertainty in Breast Radiotherapy with Bayesian Networks," in *Medical Physics*, vol. 48. Wiley, 2021.

[10] P. Ciunkiewicz, M. Roumeliotis, K. Stenhouse, P. McGeachy, S. Quirk, P. Grendarova *et al.*, "Assessment of tissue toxicity risk in breast radiotherapy using Bayesian networks," *Medical Physics*, vol. 49, no. 6, pp. 3585–3596, 2022.

[11] W. O. Kermack and A. G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.

[12] C. McAloon, A. Collins, K. Hunt, A. Barber, A. W. Byrne, F. Butler *et al.*, "Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research," *BMJ Open*, vol. 10, no. 8, p. e039652, Aug. 2020.

[13] O. Byambasuren, M. Cardona, K. Bell, J. Clark, M.-L. McLaws, and P. Glasziou, "Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: Systematic review and meta-analysis," *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada*, Dec. 2020.

[14] N. G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, and R. M. Eggo, "Age-dependent effects in the transmission and control of COVID-19 epidemics," *Nature Medicine*, vol. 26, no. 8, pp. 1205–1211, Aug. 2020.

[15] N. Fenton and M. Neil, *Risk Assessment and Decision Analysis with Bayesian Networks*, 1st ed. USA: CRC Press, Inc., 2012.

[16] P. Ciunkiewicz, S. Yanushkevich, M. Roumeliotis, K. Stenhouse, P. McGeachy, S. Quirk *et al.*, "Improved Design of Bayesian Networks for Modelling Toxicity Risk in Breast Radiotherapy using Dynamic Discretization," in *2022 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2022, pp. 01–08.

[17] S. Gupta and H. W. Kim, "Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities," *European Journal of Operational Research*, vol. 190, no. 3, pp. 818–833, Nov. 2008.

[18] X. Xu, J. Sun, H. Nie, D.-k. Yuan, and J.-h. Tao, "Linking structural equation modeling with Bayesian network and its application to coastal phytoplankton dynamics in the Bohai Bay," *China Ocean Engineering*, vol. 30, no. 5, pp. 733–748, Oct. 2016.

[19] A. Igolkina and G. Meshcheryakov, "semopy: A Python Package for Structural Equation Modeling," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 27, no. 6, pp. 952–963, Nov. 2020.

[20] M. J. Druzdzel and H. A. Simon, "Causality in Bayesian Belief Networks," in *Uncertainty in Artificial Intelligence*, D. Heckerman and A. Mamdani, Eds. Morgan Kaufmann, Jan. 1993, pp. 3–11.