
AI and Mathematical Technologies for Decision Support in Public Health

(Org: Qi Deng, Seyed Moghades and/et Jianhong Wu (York University))

CHRIS BAUCH, University of Waterloo

Tipping points in epidemiological systems

Dynamical transitions in complex systems continue to garner attention from mathematicians, on account of both their fascinating behaviour as well as their applications to public health. Many of these systems can be characterized as coupled behaviour-disease systems, where there is a two way feedback between some nonlinear transmission dynamics and a nonlinear human system. A familiar example is the COVID-19 pandemic, where a pandemic wave could drive widespread adoption of infection control measures but, as case incidence dropped, the subsequent relaxation of these measures fostered conditions for the next pandemic wave. In this talk, I will provide an overview of some of my research on mathematical modelling of tipping points in epidemiological systems, including coupled social-epidemiological systems. Methods include dynamical systems models, bifurcation theory, evolutionary game theory, and data-driven dynamical systems approaches assisted by deep learning.

MONICA COJOCARU, University of Guelph

Expanding optimization ensemble model methods for forecasting seasonal influenza in the U.S.

Each year, the seasonal influenza epidemic sees significant variability in its evolution. Accurate forecasts of future influenza cases are important for planning public health responses. The United States Centers for Disease Control and Prevention (CDC) has annually organized the FluSight competition to solicit forecasts from teams over forecast horizons (0–3). Using these data, the CDC produces an ensemble forecast of all submitted forecasts. In this paper, we introduce a weight-based ensemble forecasting method to predict laboratory-confirmed influenza hospital admissions for the 2024–2025 season. The method consists of determining optimal weights that are updated week-by-week throughout the FluSight competition to minimize the mean squared error (MSE) of a blend of teams’ previous forecasts compared to the truth data. Using these weights over an expanding time window starting at the beginning of the season, we produce our own future forecasts; we call our method the *expanding window optimization* (EWO). To improve EWO’s performance vis-a-vis the CDC ensemble model, we further introduce the *Adjusted-Weights EWO* (Adw-EWO) method. This new forecast is obtained by adding a correction term to the original EWO forecast, controlled by a parameter $\pi \in (0, 1)$. The correction term is computed using only the forecast errors at horizon 0 and is then applied uniformly across all forecast horizons. Our results show that the Adw-EWO method consistently outperforms the EWO across all horizons. Moreover, the Adw-EWO outperforms the CDC ensemble model at horizons 0, 1, and 2, while at horizon 3, the Adw-EWO and the CDC ensemble were roughly comparable.

QI DENG,

A physics-informed learner for decoding societal mobilization in epidemic transmission

Epidemic dynamics depend not only on contact-driven transmission but also on when individuals become mobilized into the effective susceptible pool. We formalize this process with the susceptibility mobilization function (SMF), a single time-resolved curve learned directly from case data. Using a physics-informed, covariate-free neural learner embedded in an extended SIR framework, we estimate SMFs for 210 COVID-19 waves across 30 Chinese provinces from 2020 to 2022. Despite substantial geographic and variant heterogeneity, SMFs exhibit consistent morphological structure that can be summarized using functional principal components. A hierarchical Bayesian alignment model links these morphological modes to societal-context domains such as mobility, population structure, urban form, and economic capacity, and reveals strong period sensitivity. During Delta waves, high-context settings displayed later and broader mobilization. Translating morphology into intervention guidance, we identify a robust principle: advancing actions slightly before the SMF peak reduces epidemic size more reliably than increasing intervention strength later. We also develop a two-axis provincial scorecard that separates current performance (SRS) from longitudinal progress (LPI), allowing fairer comparison across structurally diverse settings. By transforming diverse social

determinants into an interpretable temporal function, the SMF provides a generalizable methodological tool for analyzing context-driven transmission and supporting adaptive epidemic response.

ABBAS GHASEMI, Toronto Metropolitan University (TMU)

From Flow Instability to Airborne Transmission of Respiratory Diseases: A Computational Fluid Dynamics Approach

Mathematical modeling and computational fluid dynamics (CFD) play an essential role in advancing our understanding of complex flow phenomena such as shear-layer instability, vortex dynamics, and the emergence of chaotic flow behaviour. Among a wide range of engineering and environmental systems, these phenomena often arise in the airborne transmission of respiratory pathogens, governing their spatio-temporal dispersion. Because they frequently occur at high Reynolds numbers and involve strong nonlinear interactions between exhaled droplets, fluid inertia, viscosity, and turbulence, analytical solutions are rarely possible. Mathematical modelling, therefore, provides the theoretical framework needed to describe these instabilities, while CFD simulations offer a powerful tool for resolving the detailed spatial and temporal evolution of the flowfield. From exhalation jets to the wakes formed behind walking individuals, shear layer instabilities (e.g., Kelvin–Helmholtz instability) are responsible for the formation, growth, and breakdown of coherent vortical structures, consequently governing the pathogen dispersion. Integrating mathematical modelling with modern CFD techniques—such as large eddy simulation (LES), direct numerical simulation (DNS) enables us to sufficiently resolve the length/time-scales emerging in the turbulence energy cascade and ultimately extract the vortex dynamics and chaotic flow structures critical to the dispersion of infectious pathogens.

MICHAEL Y. LI, University of Alberta

Modeling for a purpose: influenza outbreak in a boarding school revisited.

Assessment of models should depend on the modeling objectives. Models that incorporate more realistic mechanisms are more suitable for providing insights. To produce reliable and accurate predictions and inform public health decision making, parsimonious models are more appropriate and the modeling needs to respect the data. Most of all, model calibration results should be validated by data that is independent of the calibration data, before scenario analysis is made to inform policy. As a case study, we revisit the classical example of a 1978 influenza outbreak in a boarding school in England. We demonstrate that a parsimonious SIR model with data informed time-dependent parameters can produce both accurate fitting to the time series data and validation by the final size of the epidemic. Furthermore, modeling results also provide evidence of the likely epidemic control measures implemented at the school.

JUNLING MA, University of Victoria

Detecting the change of the exponential growth rate during an early stage of an epidemic

The exponential growth rate is a key indicator of transmission intensity during the early stages of an epidemic and is closely linked to the basic reproduction number through the serial interval. Changes in control measures, transmission patterns, or the emergence of new variants can alter this rate, making rapid and reliable detection of such shifts essential for informing public health responses. We first derive proper likelihood functions and then employ a hidden Markov model (HMM) to robustly estimate the exponential growth rate. Based on the model, we develop new statistical tools for real-time detection of changes in this rate. We retrospectively apply our framework to the early phase of the COVID-19 pandemic in BC, Canada, demonstrating both the effectiveness and the limitations of our approach.

CAROLYN MCGREGOR, Ontario Tech University

BOUCHRA NASRI, Université de Montréal

Infectious disease surveillance using deep learning models

Social media data has become widely used to monitor infectious diseases. This presentation will showcase case studies of social media data related to COVID-19 and Lyme disease, demonstrating the utility of such data in predicting cases and other disease-related characteristics.

NATHANIEL OSGOOD, University of Saskatchewan

Social Media-Based Respiratory Disease Surveillance: Multi-Assessor Labelling and Cross-Model Accuracy Assessment

Timely health surveillance reporting is of foremost importance within the context of respiratory infection outbreaks. Traditional public health surveillance is often marred by reporting delays and low ascertainment rates. Recognizing the surveillance potential of Twitter reports of health symptoms, our Computational Epidemiology and Public Health Informatics Laboratory collected a repository of hundreds of millions of Canadian tweets during 2016-2022.

As a central part of their COVID-19 pandemic work under contract with SHA, PHAC, and FNIHB, the applicants employed Particle Filtering and PMCMC of compartmental models with diverse empirical data for regular reporting and projections. That work further investigated the potential for augmenting such data with time series gathered from tweets automatically classified as plausible COVID-19 or Influenza cases. Substantially expanding on the results shared in Tian et al. 2025, we describe here an end-to-end project using multi-assessor tweet labeling, training and testing of language embedding with 24 diverse machine learning models while navigating class imbalances, and comparisons of model accuracy across multiple accuracy measures. 16 models and model variants were assessed for accuracy in identifying tweets reporting plausible COVID-19 cases, with an additional 8 models being assessed for classifying plausible influenza cases. Models were compared using recall, F1, area under the ROC Curve and accuracy. Among models, transformer-based tweets performed most favourably, followed by an ensemble method involving a diverse set of classical techniques. In closing, we discuss our efforts to shift data collection platforms, and to augment our methods to detect spread of health disinformation via sheaf-based deep learning.

AFFAN SHOUKAT, University of Regina

Physics Informed Neural Networks for Fractional Logistic Growth Models

We present fractional physics-informed neural networks (fPINNs) for solving generalized logistic growth models governed by the Atangana-Baleanu fractional derivative and proportional delay. By combining automatic differentiation with numerical quadrature for the nonlocal Mittag-Leffler kernel, fPINNs can accurately capture memory-dependent and time-lagged dynamics. We show an application in epidemiology, where fPINNs can be used to infer hidden growth rates and memory effects from limited time-series data of infectious disease outbreaks.

EDWARD THOMMES, Sanofi/University of Guelph

Long-range forecasting of seasonal influenza vaccine uptake using web search data

The population-level burden of an influenza season depends strongly on the proportion of people vaccinated beforehand. Being able to predict whether a given season is on track to have low, high or average uptake would provide a critical piece of highly actionable intelligence for all stakeholders involved: Governments, public health authorities, healthcare providers, vaccines manufacturers, and not least, the population itself. In particular, sufficiently early advance warning provides the opportunity for interventions—such as increased investment in awareness programs—to proactively boost participation and avert a projected shortfall. We present an ensemble forecast model which utilizes a panel of Google web search queries to make meaningful predictions about US national-level total seasonal vaccine uptake as early as the beginning of the year in which the season begins.

FUNDING AND DISCLOSURES: This work was supported by an NSERC Alliance grant co-funded by Sanofi. ET is an employee of Sanofi and may hold shares and/or stock options

WOLDEGEBRIEL ASSEFA WOLDEGERIMA, York University

Toy Introduction to Epidemiology-Informed Neural Networks (EINNs) with Application

The integration of mechanistic modeling and machine learning has the potential to revolutionize the way we understand complex biological systems. Particularly, the development of Informed Neural Networks (INNs), a class of hybrid models that embed domain-specific knowledge, such as differential equations, into the neural network architecture, has attracted many researchers recently. Epidemiology-Informed Neural Networks (EINNs) incorporate domain-specific knowledge from disease dynamics (e.g., differential equation models, compartmental models like SIR, vector-host interactions, etc.) into their architecture, loss functions, or training process. The loss function that is minimized during training is the combined loss of the data and the DE residuals. This method helps to improve learning, prediction accuracy, interpretability, and parameter estimation, particularly in scenarios where data is sparse or noisy. In this talk, I will quickly introduce the foundations of EINNs. I will then present some results from our study that we trained an EINN on synthetic data derived from an SI-SIR model designed for Avian influenza and show the model's accuracy in predicting extinction and persistence conditions. In the method, a twelve-layer hidden model was constructed with sixty-four neurons per layer, and the ReLU activation function was used. The network is trained to predict the time evolution of five state variables for birds and humans over 50,000 epochs. The overall loss is minimized to 0.000006, characterized by a combination of data and physics losses, enabling the EINN to follow the differential equations describing the disease progression.

SICHENG ZHAO, McMaster University

Improving Infectious Disease Prevalence Estimation and Parameter Inference Using Number of Tests and Positivity Data

Prevalence of infection is a critical variable for modeling infectious dynamics and public health decision-making. However, estimating true prevalence from surveillance data remains challenging. Here we discuss three novel attempts to model testing mechanism using the beta-distribution, hazard rates and odds ratios respectively. These methods aim to link prevalence with the number of tests, test positivity, and some test characteristics at each data point in a robust, flexible, and theoretically justified manner. We further present a data-fitting framework based on the odds ratio approach and demonstrate its performance using simulated datasets as a proof of concept.