University of Saskatchewan

Forecasting

Agenda

- Introduction to Forecasting in Epidemiology
- Data Sources and Collection
- Types of Forecasting Models
- Case Study: Forecasting COVID-19
- Challenges and Opportunities
- Conclusion and Q&A

Introduction

- **Definition**
  - Forecasting in epidemiology involves predicting the trends, spread, and impact of diseases and health conditions within populations over time and/or space.
- **Key Objectives**
  - Predicting disease trends to prepare and respond effectively.
  - Informing public health decision-making and resource allocation.
  - Enhancing the understanding of disease dynamics and risk factors.

Introduction

- **Importance**
  - Aids in early detection and management of disease outbreaks.
  - Supports policy development and public health interventions.
  - Helps in reducing morbidity and mortality rates.
- **Methodologies**
  - Utilizes statistical, mathematical, and computational models.
  - Incorporates diverse data sources including historical, clinical, and environmental data.
  - Often employs machine learning and artificial intelligence for enhanced prediction accuracy.

UNIVERSITY OF SASKATCHEWAN

Qual versus Quant

- Qualitative forecasting techniques
    - Subjective based on the opinion and judgment of experts
    - Appropriate when past data are not available
    - They are usually applied to intermediate- or long-range decisions
    - Examples of qualitative forecasting methods
        - Informed pinion and judgment
        - Delphi method
        - Market research

Qual versus Quant

- Quantitative forecasting models
  - Used to forecast future data as a function of past data
  - Appropriate to use when past numerical data is available
  - When it is reasonable to assume that some of the patterns in the data are expected to continue into the future
  - Applied to short- or intermediate-range decisions
  - Examples
    - Last period demand
    - Simple and weighted N-Period moving averages
    - Simple exponential smoothing

UNIVERSITY OF SASKATCHEWAN

Data Sources and Collection

- Data quality is one of the most important drivers of model performance. If data are inconsistent or do not reflect reality, models have no reliable ground truth from which to learn or be evaluated.
- Invest in infrastructure for data collection
  - Prioritise collection of timely high temporal and spatial resolution data
  - Standardise reporting of data across jurisdictions (eg, US states)
  - Pursue high-quality data that captures risk-reduction behaviours
  - Expand genomic surveillance

https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00167-4/fulltext

Data Sources and Collection

- Historical Data
  - Previous disease outbreaks and trends
- Real-Time Data
  - Current disease spread, reported cases, and deaths
- Environmental Data
  - Factors like climate, population density, etc.
- Visual
  - Infographics or icons representing different data source

Generic Models

- SIR

Susceptible (S) → Infectious (I) → Recovered (R)

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\beta I S \tag{1}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I S - \gamma I \tag{2}$$

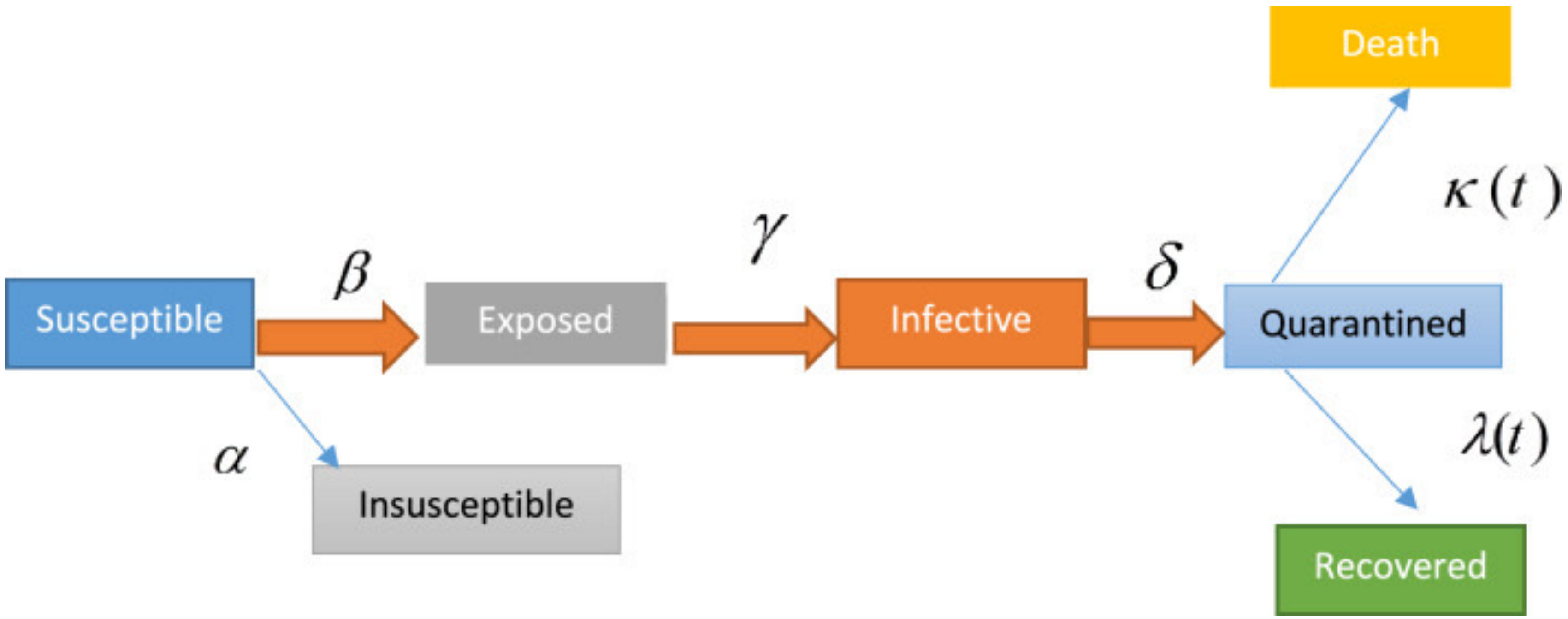
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I \tag{3}$$

where *S* is the number of individuals susceptible at time *t*; *I* is the number of infected individuals at time *t*; *R* is the number of recovered individuals at time t; and β and γ are the transmission rate and rate of recovery (removal), respectively.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861008/

Generic Models

- SEIR



More math than you need. See the paper if you want.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861008/

Models

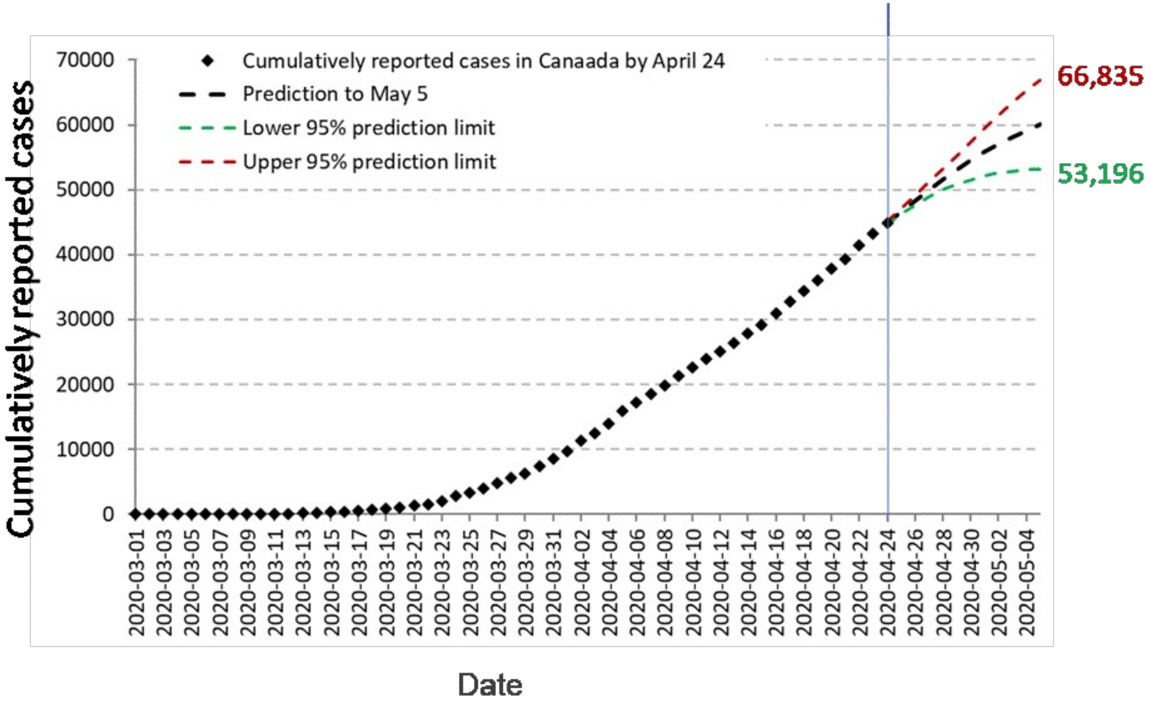
- Statistical forecast models
  - Short-range forecast of expected cases given recent incidence
- Long-range forecast models
  - Dynamic compartment model adapted to project near-future given recent incidence and scenarios for control/release/variants of concern
- Models to explore scenarios
  - More complex models
  - Agent-based model

https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/epidemiological-economic-research-data/mathematical-modelling.html

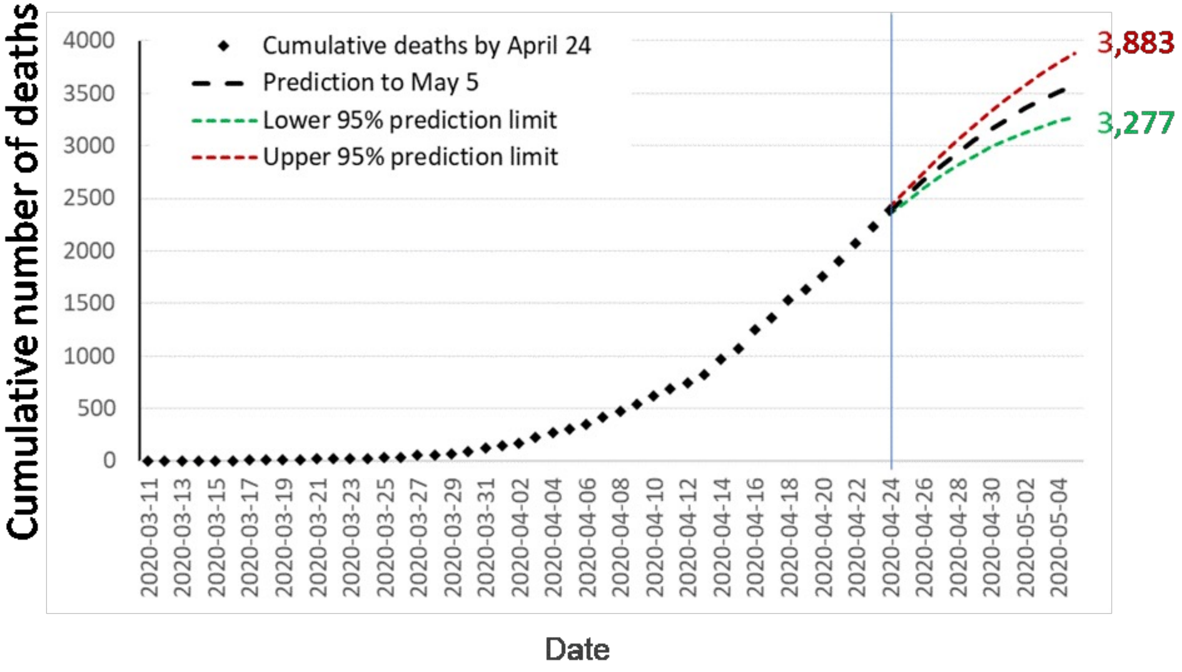
University of Saskatchewan

Models – Short Range

53,196 to 66,835 cases by May 5



3,277 to 3,883 deaths by May 5

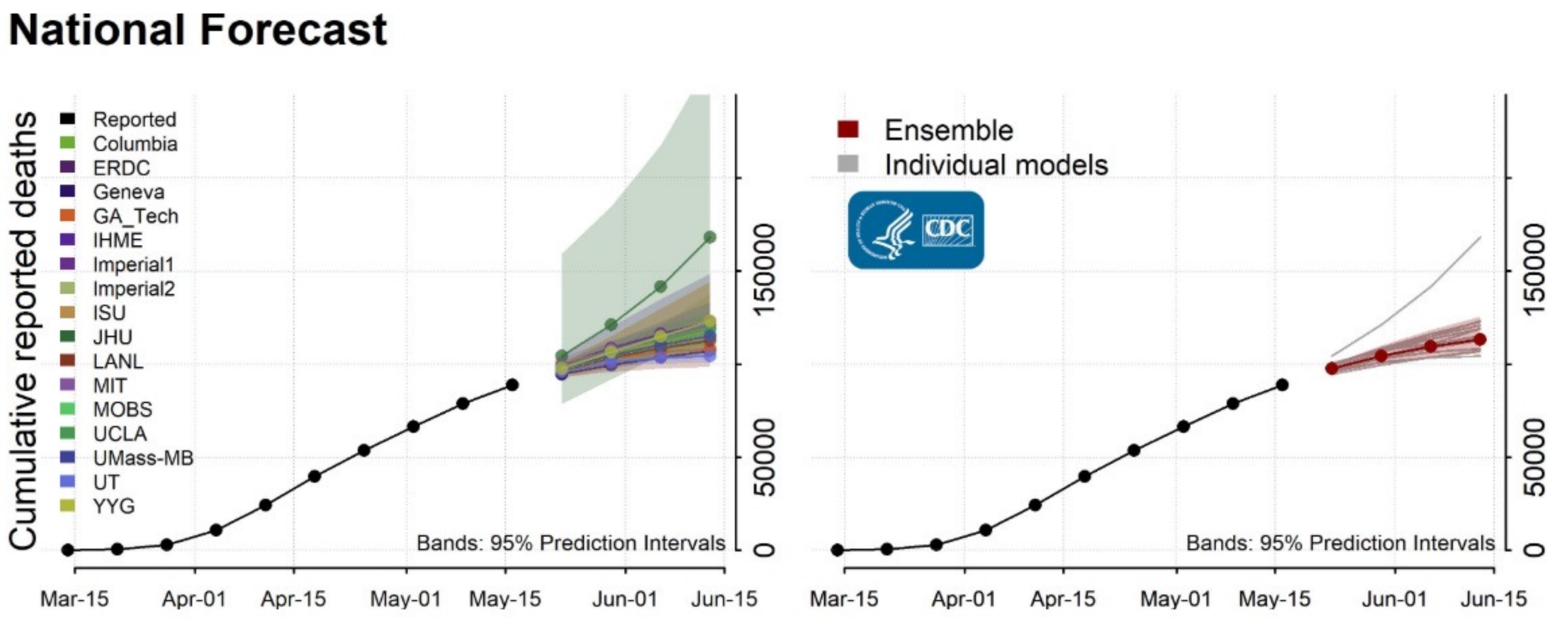


https://www.canada.ca/en/public-health/services/publications/diseases-conditions/covid-19-using-data-modelling-inform-public-health-action-april-28-2020.html

University of Saskatchewan

Models – Long Range

National Forecast

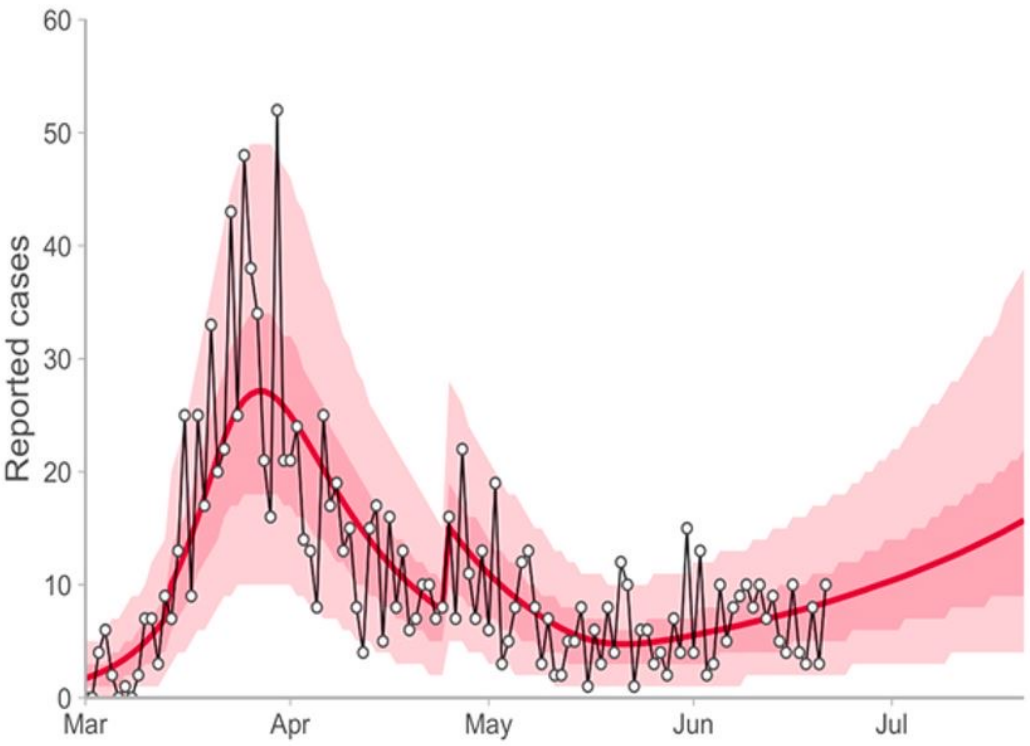


https://media.defense.gov/2020/May/27/2002307574/-1/-1/0/200526-A-A1403-001.JPG

Models

Dynamic Compartmental Modelling: Projections

- Our model suggests a slight increase in new cases during June and the possibility of continued growth in new cases during the summer.
- Recent contact rates are estimated to be at roughly 65% of normal.



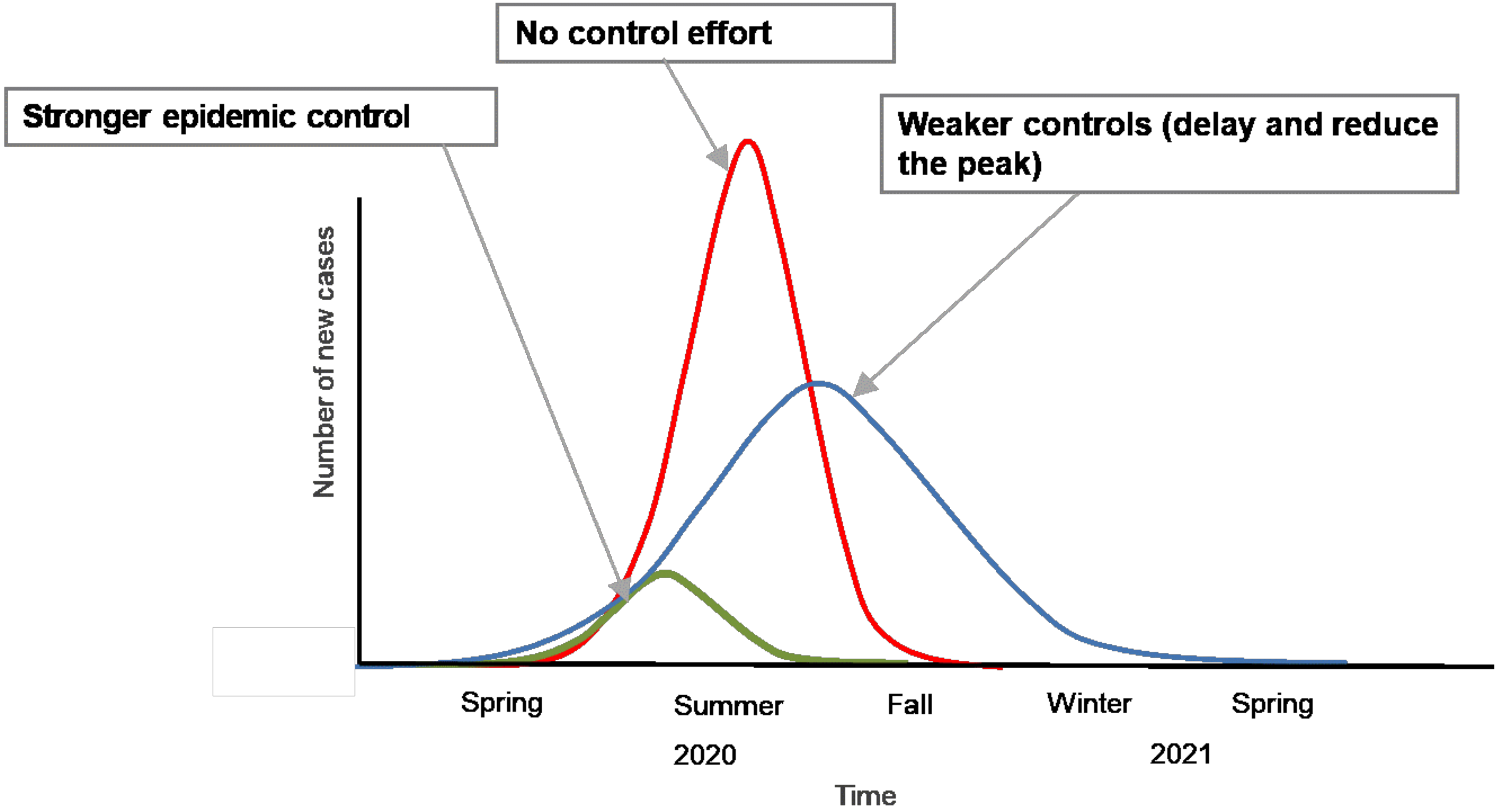
Solid line: mean; shaded bands: 50% and 90% credible intervals; Open circles: reported cases. Cases used for model fitting exclude those attributed to outbreak clusters.

COVID-19 IN BC

https://www.cheknews.ca/bc-health-officials-to-release-latest-covid-19-modelling-data-tuesday-677838/

UNIVERSITY OF SASKATCHEWAN

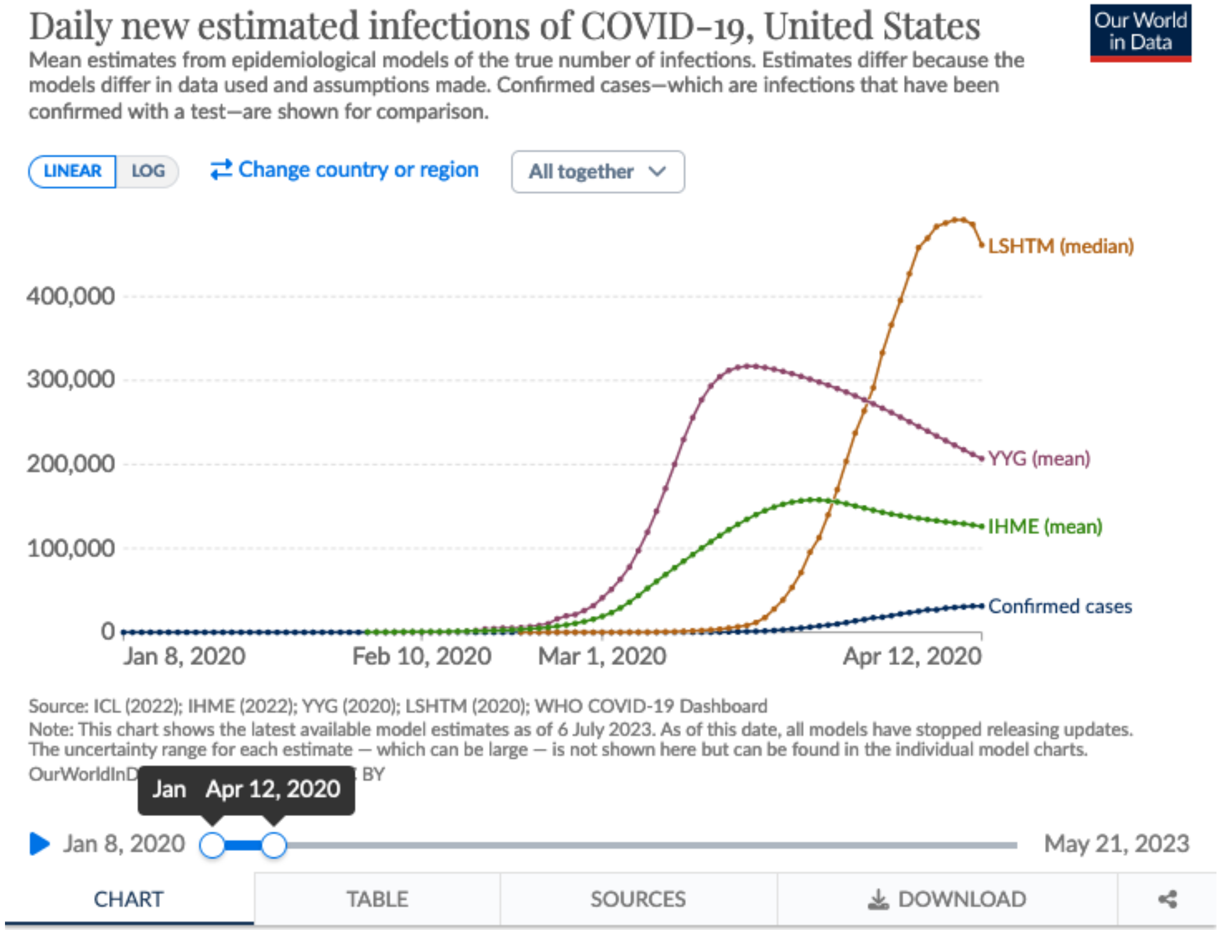
Models - Scenario



https://www.canada.ca/en/public-health/services/publications/diseases-conditions/covid-19-using-data-modelling-inform-public-health-action-april-28-2020.html

Model Assumptions

- Model parameter assumptions play a critical role in any forecast
  - IFR: infection fatality rate
  - Rt: effective reproduction number



https://ourworldindata.org/covid-models

Types of Models

- **Autoregressive model**
  - The autoregressive time-series model is known as a useful tool to model dependent data and has been applied to various real-world problems
  - **Simple moving average**
    - A simple moving average (SMA) is defined as the unweighted mean of an equal number of data on either side of a central value (7 day moving average).
  - **Autoregressive integrated moving average (ARIMA)**
    - An autoregressive integrated moving average model is a generalized form of the autoregressive moving average model.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861008/

Types of Models

- **Exponential models**
  - Exponential models are suitable in the modeling of several phenomena, such as populations, interest rates, and infectious diseases
    - **Logistic functions**
    - **Deep learning**
    - **Regression methods**
    - **Prophet algorithm**
    - **Genetic programming**
    - **SIR**
    - **SEIR**
    - **SIRD**

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861008/

Case Study

- Examine the differences between the models
    - **https://ourworldindata.org/covid-models**
- Answer the following questions
    - What type of model is used?
    - What type of analysis method is used?
    - What are the strengths of the model?
    - What are the limitations of the model?
    - Do you trust these models?

https://ourworldindata.org/covid-models

University of Saskatchewan

Challenges and Opportunities

- Challenges
  - Data accuracy, timely reporting, model limitations, etc.
- Opportunities
  - Improved data collection, advanced models, global collaboration
- Visual
  - Infographic of challenges and opportunities or a balance scale

Challenges and Opportunities

- Technology
  - The role of AI and machine learning in future forecasting versus traditional models
  - How much AI are you going to have to know
- Collaboration
  - Provincial and global data sharing and collaborative forecasting efforts
- Visual
  - Infographic or icons indicating future trends