Disease Forecasting

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Dr. Jaspreet Kaur Dhanjal
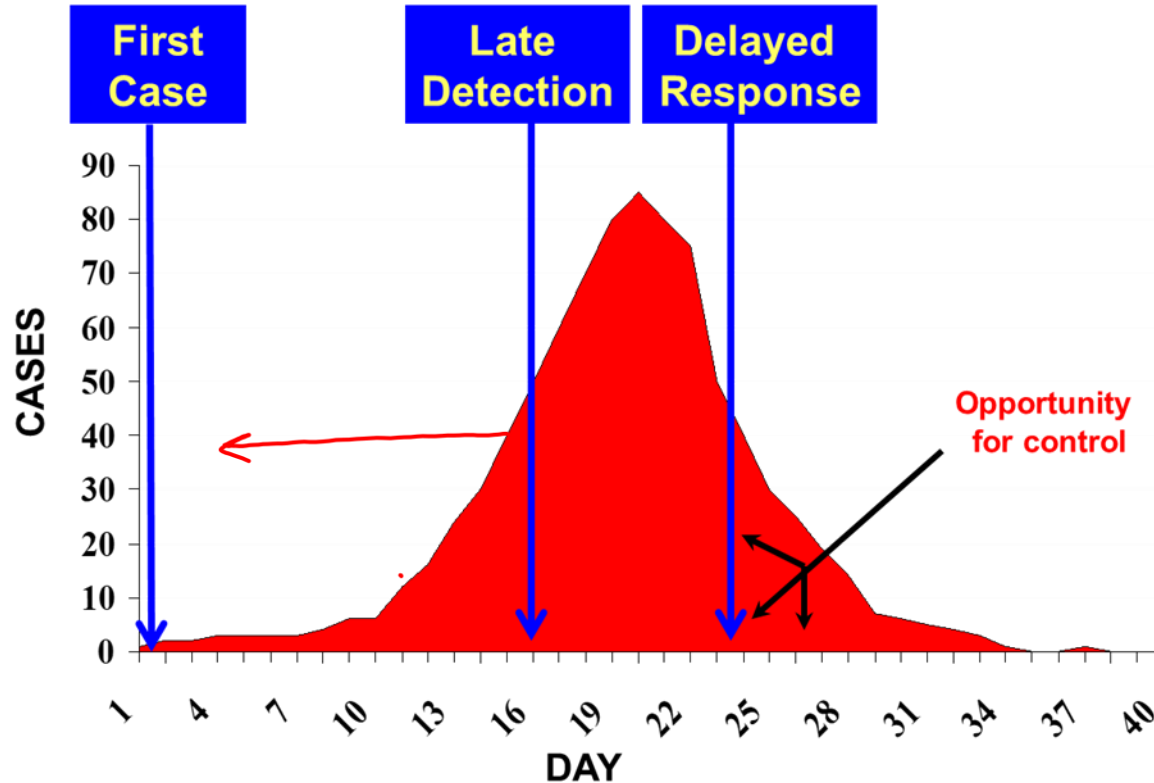**Assistant Professor, Department of Computational Biology**
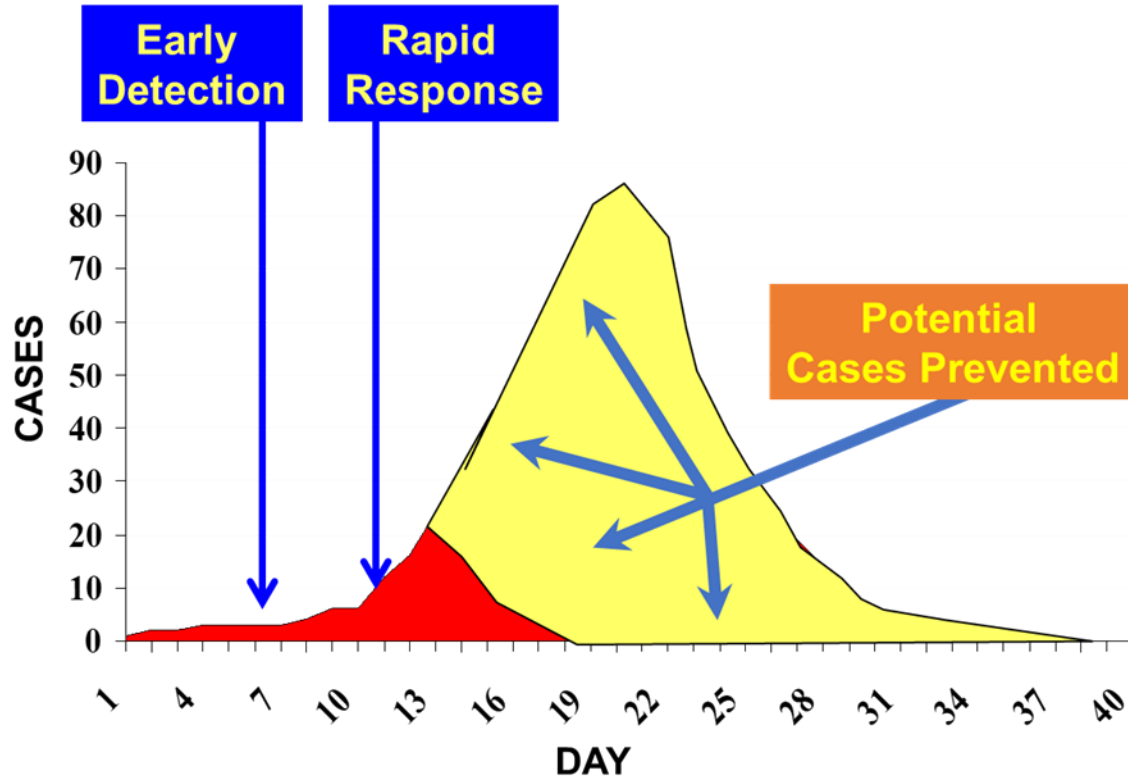**Email ID: jaspreet@iiitd.ac.in**

November 4, 2025

Why Disease Forecasting?

Outbreak detection and response without preparedness

Why Disease Forecasting?
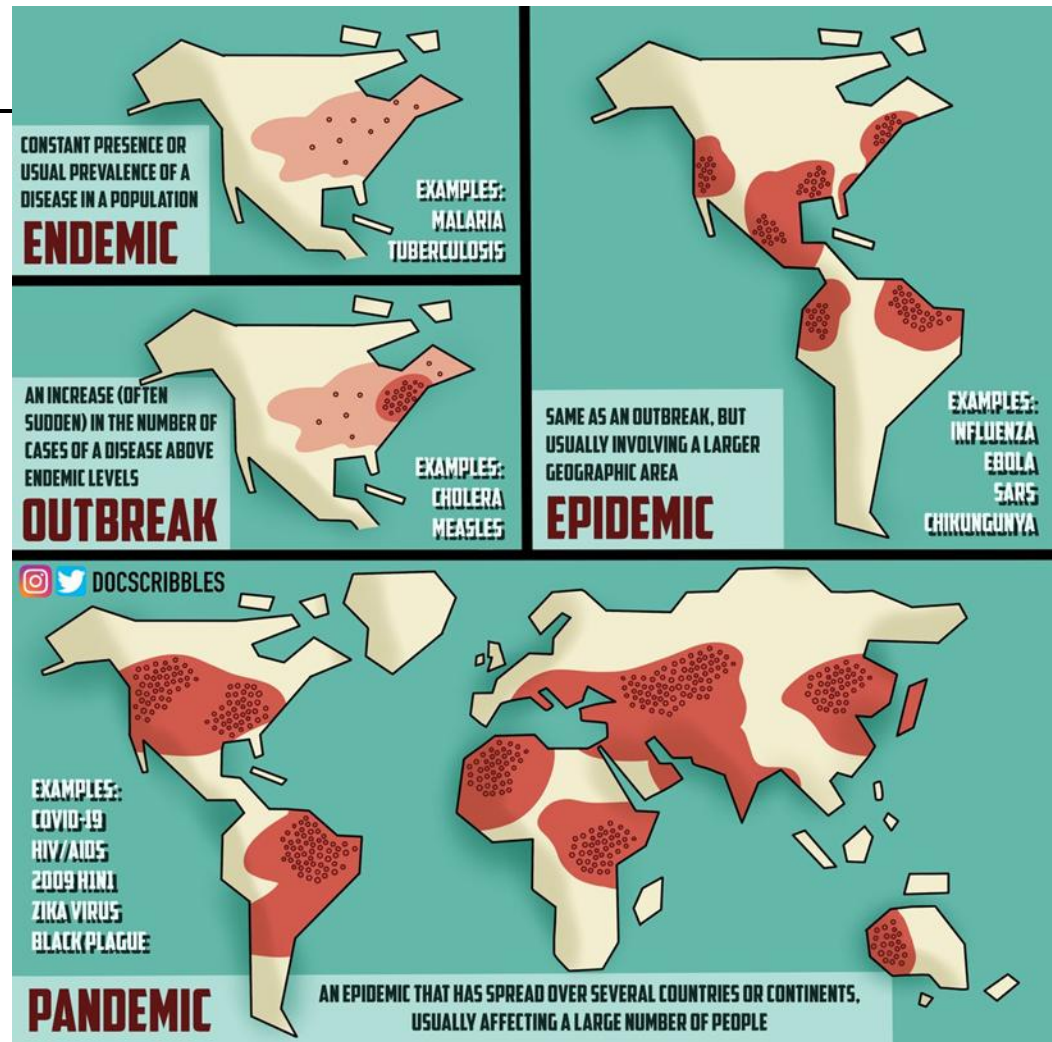
Outbreak detection and response with preparedness

What is Epidemiology?

- The science which investigates the causes and controls of epidemic diseases.

- The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems.

Infectious Diseases

- **Endemic**

- **Outbreak**

- **Epidemic**

- **Pandemic**

History of Epidemiology

Examples of epidemics include:

- 1918 flu pandemic (killed more people than World War I!)
- 2003 SARS
- 2009 Swine flu
- 2014 Ebola
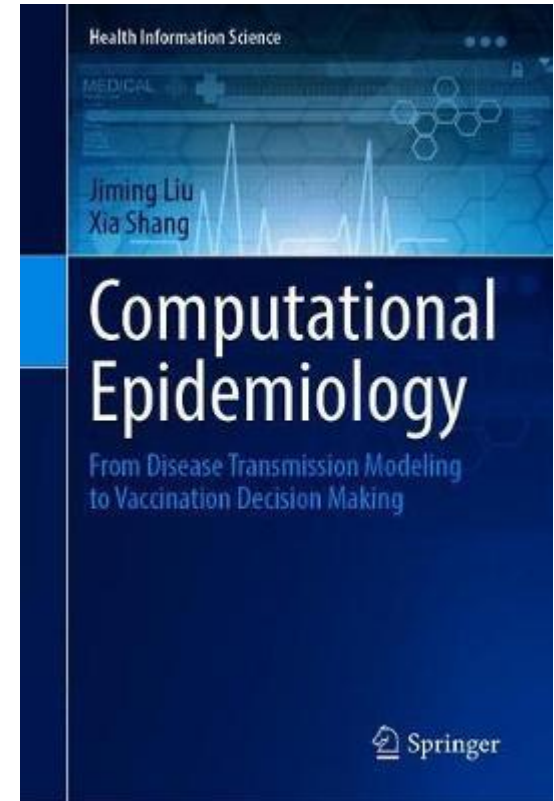- 2020 COVID-19 (ongoing)

Examples of epidemics further in the past include:

- Plagues in Roman times
- The Plague from the Middle Ages (killed 30-60% of Europe's population)
- A smallpox outbreak in the 1500s

Computational Epidemiology

- The development of computational and mathematical methods, tools, and techniques to support epidemiology.

- Data-based approaches to analyzing epidemiology problems can save a lot of time and resources compared to traditional laboratory/experimental approaches.

Health Information Science

Jiming Liu
Xia Shang

Computational Epidemiology

From Disease Transmission Modeling to Vaccination Decision Making

Springer

Smallpox: Bernoulli and Jenner (1726-1800)

- Jenner basically discovered vaccination; it turned out that infecting people with cowpox, a zoonotic (meaning it transfers from animals to humans) virus that is relatively mild in humans, provided immunity against smallpox.
- But Bernoulli was the first major example of someone using a mathematical argument to argue in favor of it.

DANIEL BERNOVLLIVS

8

Cholera: John Snow (1800-1900)

- Miasma theory- Bad air makes one sick.
- After observing the cholera outbreak, John Snow believed that cholera was being transmitted through water.

Supplier	Number of houses	Cholera deaths	Deaths per 10,000 houses
S&V	40,046	1,263	315
Lambeth	26,107	98	37
Rest of London	256,423	1,422	59

Figure : A table showing the data that John Snow collected about cholera counts from a region upstream of the Thames (Lambeth), a region downstream of the Thames (S&V), and the rest of London.
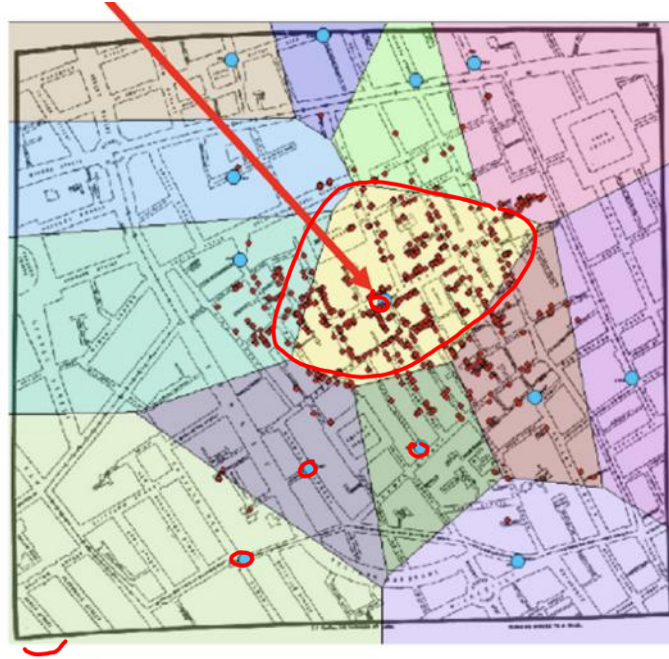
Cholera: John Snow (1800-1900)

Figure : A voronoi map created by John Snow showing cholera cases separated into groups by the closest water pump. The yellow region, which had the most cases, was the Broad Street region where Snow convinced the authorities to replace the pump to reduce cholera cases.

Malaria: Ross, McDonald, Lotka, McKendrick (1900-1960)

- Ross found the cause of malaria by dissecting an Anopheles mosquito.

- He then used a mathematical model to suggest that mosquito reduction could contain the spread of the disease.

Data science for Epidemiology

These are some of the questions data science can address in epidemiology:

1. When and where did the outbreak start? Who initially got infected?

2. What can we expect as the epidemic spreads? What kind of people are likely to be infected? When will the case number peak?

3. How to control the epidemic? What preventative measures can be taken?

Case study

nature

Explore Content ⌄   Journal Information ⌄   Publish With Us ⌄

nature > letters > article

Published: 19 February 2009

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg, Matthew H. Mohebbi ✉, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski & Larry Brilliant

Nature **457**, 1012–1014(2009) | Cite this article

16k Accesses | **2217** Citations | **548** Altmetric | Metrics

Case study

Introduction

- Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year.

- In addition to seasonal influenza, a new strain of influenza virus against which no previous immunity exists demonstrated that human-to-human transmission could result in a pandemic with millions of fatalities.

- Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza.

14

Case study

Methodology adopted

- Large numbers of Google search queries were analyzed to track influenza-like illness in a population.

- The relative frequency of certain queries could be highly correlated with the percentage of physician visits in which a patient presents with influenza-like symptoms.

- Using such information the level of weekly influenza activity in different geographical regions can be accurately predicted without much reporting lag.
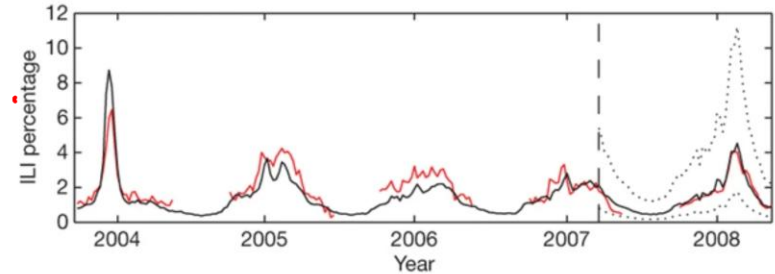
Case study

Results

Table | Topics found in search queries which were found to be most correlated with CDC ILI data

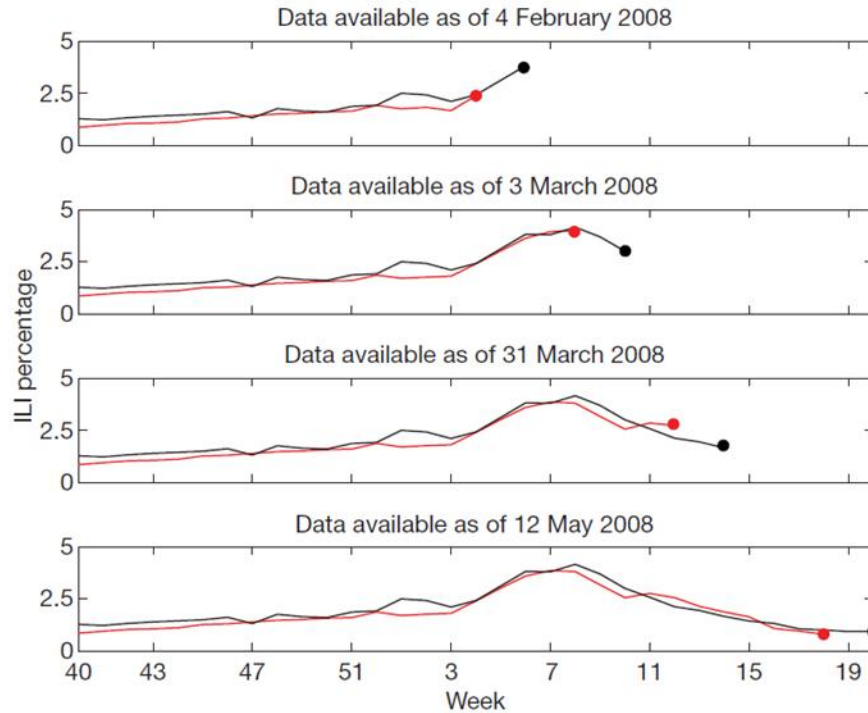| Search query topic	Top 45 queries	
	n	Weighted
Influenza complication	11	18.15
Cold/flu remedy	8	5.05
General influenza symptoms	5	2.60
Term for influenza	4	3.74
Specific influenza symptom	4	2.54
Symptoms of an influenza complication	4	2.21
Antibiotic medication	3	6.23
General influenza remedies	2	0.18
Symptoms of a related disease	2	1.66
Antiviral medication	1	0.39
Related disease	1	6.66
Unrelated to influenza	0	0.00
Total	45	49.40

Figure 2: A comparison of model estimates for the mid-Atlantic region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated.
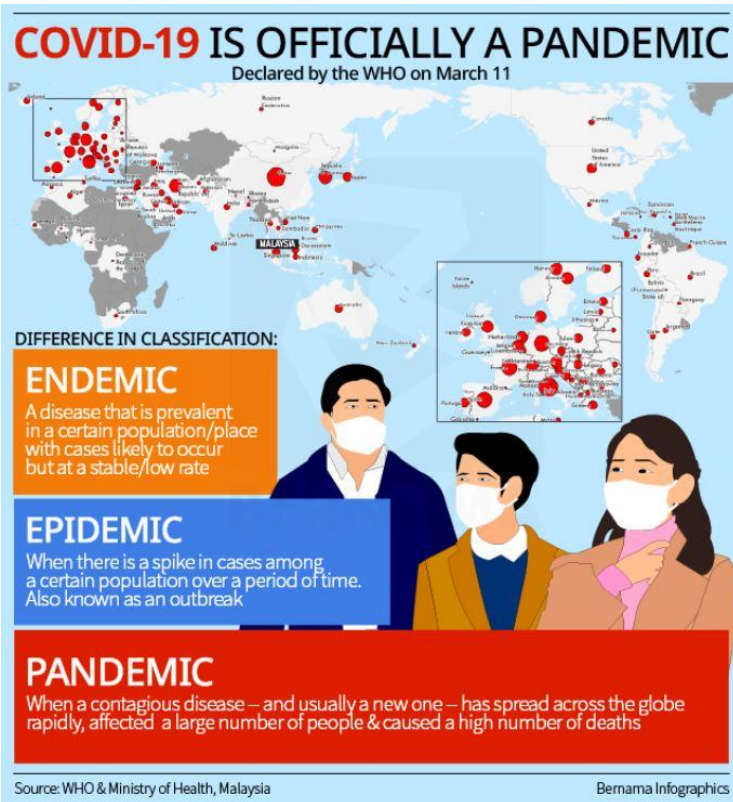

A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, whereas a correlation of 0.96 was obtained over 42 validation points. Dotted lines indicate 95% prediction intervals. The region comprises New York, New Jersey and Pennsylvania.

Case study

Figure | ILI percentages estimated by our model (black) and provided by the CDC (red) in the mid-Atlantic region, showing data available at four points in the 2007–2008 influenza season

Case studies related to SARS-CoV-2

COVID-19 IS OFFICIALLY A PANDEMIC
Declared by the WHO on March 11

DIFFERENCE IN CLASSIFICATION:

ENDEMIC
A disease that is prevalent in a certain population/place with cases likely to occur but at a stable/low rate

EPIDEMIC
When there is a spike in cases among a certain population over a period of time. Also known as an outbreak

PANDEMIC
When a contagious disease — and usually a new one — has spread across the globe rapidly, affected a large number of people & caused a high number of deaths

Source: WHO & Ministry of Health, Malaysia

Bernama Infographics

Report 2 - Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China

WHO Collaborating Centre for Infectious Disease Modelling; MRC Centre for Global Infectious Disease Analysis; Abdul Latif Jameel Institute for Disease and Emergency Analytics; Imperial College London, UK

Summary

On January 16th we released estimates of the scale of the COVID-19 (previously termed 2019-nCoV) outbreak in China based on an analysis of the number of cases detected outside mainland China. Since then, cumulative confirmed cases reported by the Chinese authorities have increased 10-fold, to 440 by January 22nd. The number of detected outside China with symptom onset by 18th January had increased to 7 in the same time. Here we report updated estimates of the scale of the epidemic in Wuhan, based on an analysis of flight and population data from that city. Our estimate of the number of cases in Wuhan with symptoms onset by January 18th is now 4,000. The uncertainty range is 1,000-9,700, reflecting

Key info

Date:
22 January 2020

Authors:
Natsuko Imai, Ilaria Dorigatti, Anne Cori, Christl Donnelly, Steven Riley, Neil M. Ferguson

Correspondence:
Professor Neil Ferguson
neil.ferguson@imperial.ac.uk

Case studies related to SARS-CoV-2

Tracing and analysis of 288 early SARS-CoV-2 infections outside China: A modeling study

Francesco Pinotti [1], Laura Di Giulia Pullano [1] [5], Eugenio

Affiliations + expand

Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts

Joel Hellewell [1],
Timothy W Russe
Centre for the M
Sebastian Funk [1]

Collaborators, A

Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak

Chad R Wells [1], Pratha Sah [1], Seyed M Moghadas [2], Abhishek Pandey [1], Affan Shoukat [1], Yaning Wang [3], Zheng Wang [4], Lauren A Meyers [5] [6], Burton H Singer [7], Alison P Galvani [1]

Case studies related to SARS-CoV-2

VacSIM: Learning Effective Strategies for COVID-19 Vaccine Distribution using Reinforcement Learning

Raghav Awasthi, Keerat Kaur Guliani, Saif Ahmad Khan, Aniket Vashishtha, Mehrab Singh Gill, Arshita Bhatt, Aditya Nagori, Aniket Gupta, Ponnurangam Kumaraguru, Tavpritesh Sethi

A COVID-19 vaccine is our best bet for mitigating the ongoing onslaught of the pandemic. However, vaccine is also expected to be a limited resource. An optimal allocation strategy, especially in countries with access inequities and temporal separation of hot-spots, might be an effective way of halting the disease spread. We approach this problem by proposing a novel pipeline VacSIM that dovetails Sequential Decision based RL models into a Contextual Bandits approach for optimizing the distribution of COVID-19 vaccine. Whereas the Reinforcement Learning models suggest better actions and rewards, Contextual Bandits allow online modifications that may need to be implemented on a day-to-day basis in the real world scenario. We evaluate this framework against a naive allocation approach of distributing vaccine proportional to the incidence of COVID-19 cases in five different States across India and demonstrate up to 9039 additional lives ... We also propose nov... Finally, we contribute... across the globe(this

Predicting Emerging Themes in Rapidly Expanding COVID-19 Literature with Dynamic Word Embedding Networks and Machine Learning
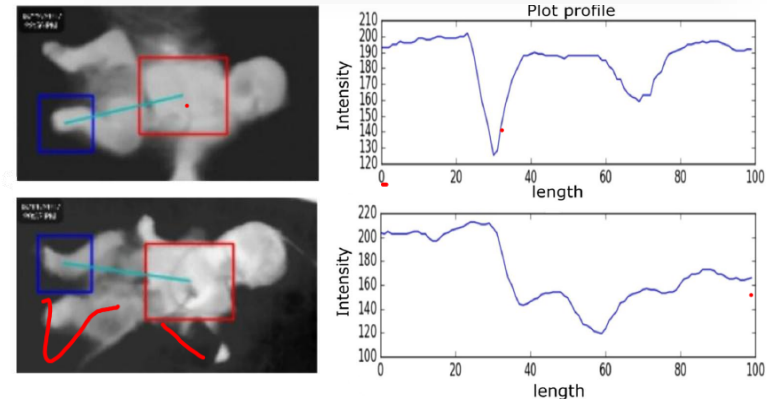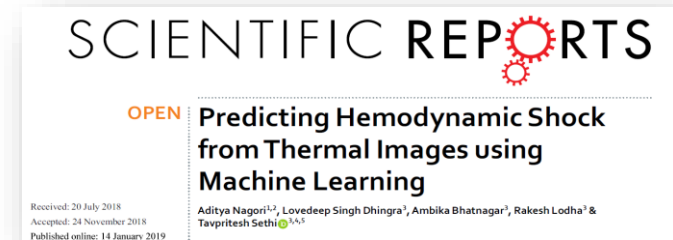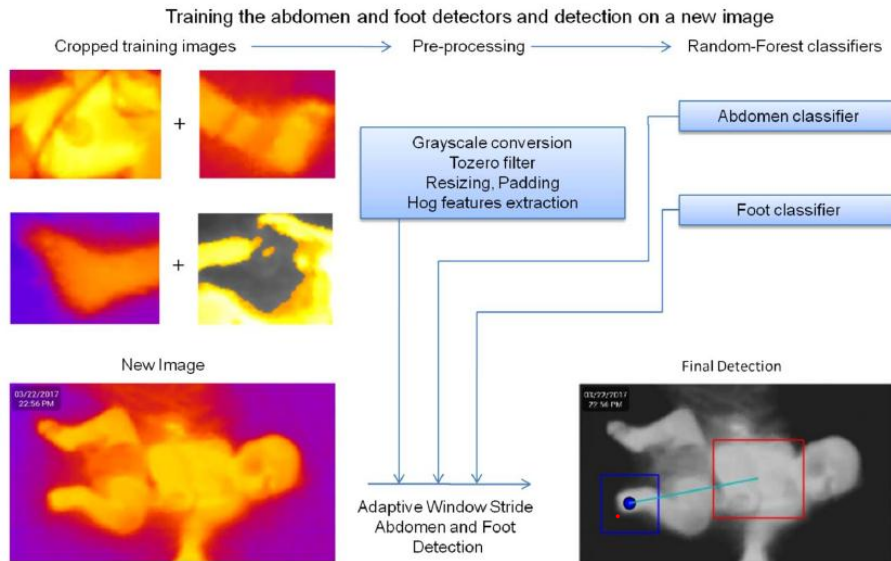
Ridam Pal, Harshita Chopra, Raghav Awasthi, Harsh Bandhey, Aditya Nagori, Amogh Gulati, Ponnurangam Kumaraguru, Tavpritesh Sethi

Case studie related to hemodynamic shock

Proactive detection of hemodynamic shock can prevent organ failure and save lives.

Thermal imaging is a non-invasive, non-contact modality to capture body surface temperature with the potential to reveal underlying perfusion disturbance in shock.

Case study related to Plant Biology

- Rice blast is caused by *Magnaporthe oryzae* or rice blast fungus
- 15-20% crop gets damaged by this disease.