A Conducted Data Tour in the Health Domain

Anjan Kr. Dasgupta adgcal@gmail.com

June 2020

The future in a post-Coronavirus world

Anjan Kr. Dasgupta
adgcal@gmail.com

- The value of open innovation to solve a crisis
- Unexpected lessons from an unprecedented crisis
- How do we keep innovation projects going over distance and through a crisis?
- The future of healthcare and medical tech

How to rearchitect?

Anjan Kr. Dasgupta
adgcal@gmail.com

- How can health rearchitect themselves to be more crisis-proof?
- How will medical care approach risk as the pandemic settles down?
- Smart biomedical investment strategies during the Coronavirus pandemic
- What priorities must change, and what must stay the same.

E-Words on or before 1999

Anjan Kr. Dasgupta
adgcal@gmail.com

- e-Commerce
- e-Business
- e-Solutions

E-Words after 2000

Anjan Kr. Dasgupta
adgcal@gmail.com

- e-Health
- e-Agriculture
- e-Environment

Components of e-health

Anjan Kr. Dasgupta
adgcal@gmail.com

- Telemedicine
- Health Informatics
- MHealth
- Evidence Based Medicine

Telemedicine

Anjan Kr. Dasgupta
adgcal@gmail.com

- Capture readings from medical devices at a faraway location.
- Using telemedicine software, patients can see a doctor for diagnosis and treatment without having to wait for an appointment.
- Patients can consult a physician at the comfort of their home
- https://www.healthline.com/health/best-telemedicine-companies

Telemedicine Companies

Anjan Kr. Dasgupta
adgcal@gmail.com

- US based
  - Teladoc, MeMD, iCliniq Amwell,PlushCare,HealthTap, Takeaway
- India based
  - https://vhealth.io

Health Informatics

Anjan Kr. Dasgupta
adgcal@gmail.com

Health informatics uses ICT to organize and analyze health records to improve healthcare outcomes.
- https://www.iihmrdelhi.edu.in/health-information-technology?gclid=CjwKCAjw8pH3BRAXEiwA1pvMsT10OtbAhawHxXu9ZLMGp2QFddVlIP9Bmm8xl4Ienzt0kBfRhAGpUxoCN1QQAvD_BwE

MHealth (mobile health)

Anjan Kr. Dasgupta
adgcal@gmail.com

- Refers to use of mobile communication devices, such as mobile phones,personal digital assistants (PDAs), and wearable devices such as smart watches, for health services.
- https://equityhealthj.biomedcentral.com/articles/10.1186/s12939-020-01175-7

Evidence Based Medicine

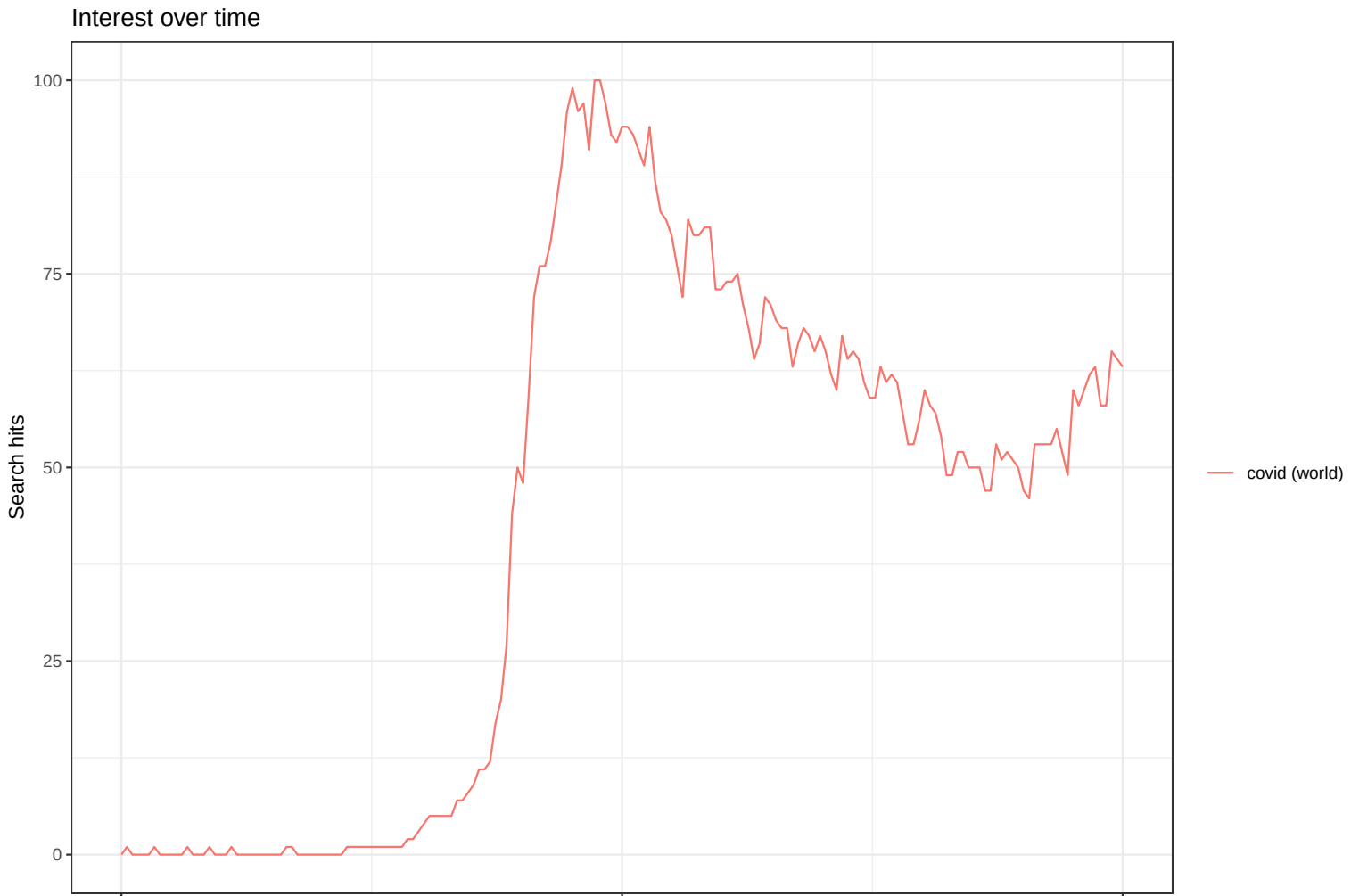
Anjan Kr. Dasgupta
adgcal@gmail.com

- Evidence based medicine (EBM) uses current DATA SCIENCE based approaches in decisions making particularly in the care of individual patients.
- EBM integrates clinical experience and patient values with the best available research information.
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3789163/#:~:text=Evidence%20based%20medicine%20(EBM)%20is,the%20best%20available%20research%20information.

Trends Inspection

Anjan Kr. Dasgupta
adgcal@gmail.com

```
library(gtrendsR)
tr <- gtrends("covid",geo="",time="2020-01-01 2020-07-01")
plot(tr)
```



Competeing synonmns

Anjan Kr. Dasgupta
adgcal@gmail.com

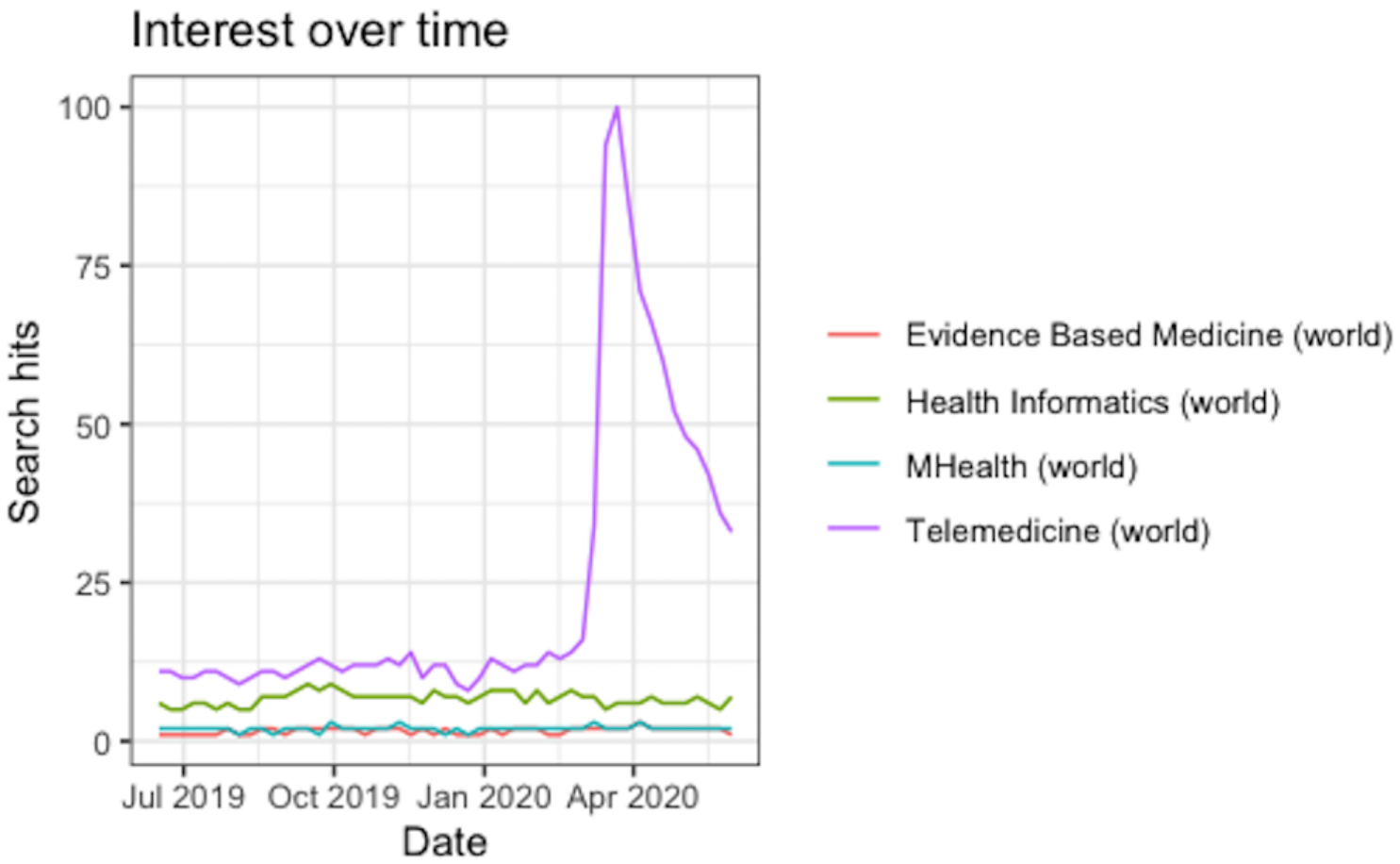


Figure 1: Comparative Synonymns- A sudden peak appears during the COVID 19 outbreak

A complimentary trend

Anjan Kr. Dasgupta
adgcal@gmail.com

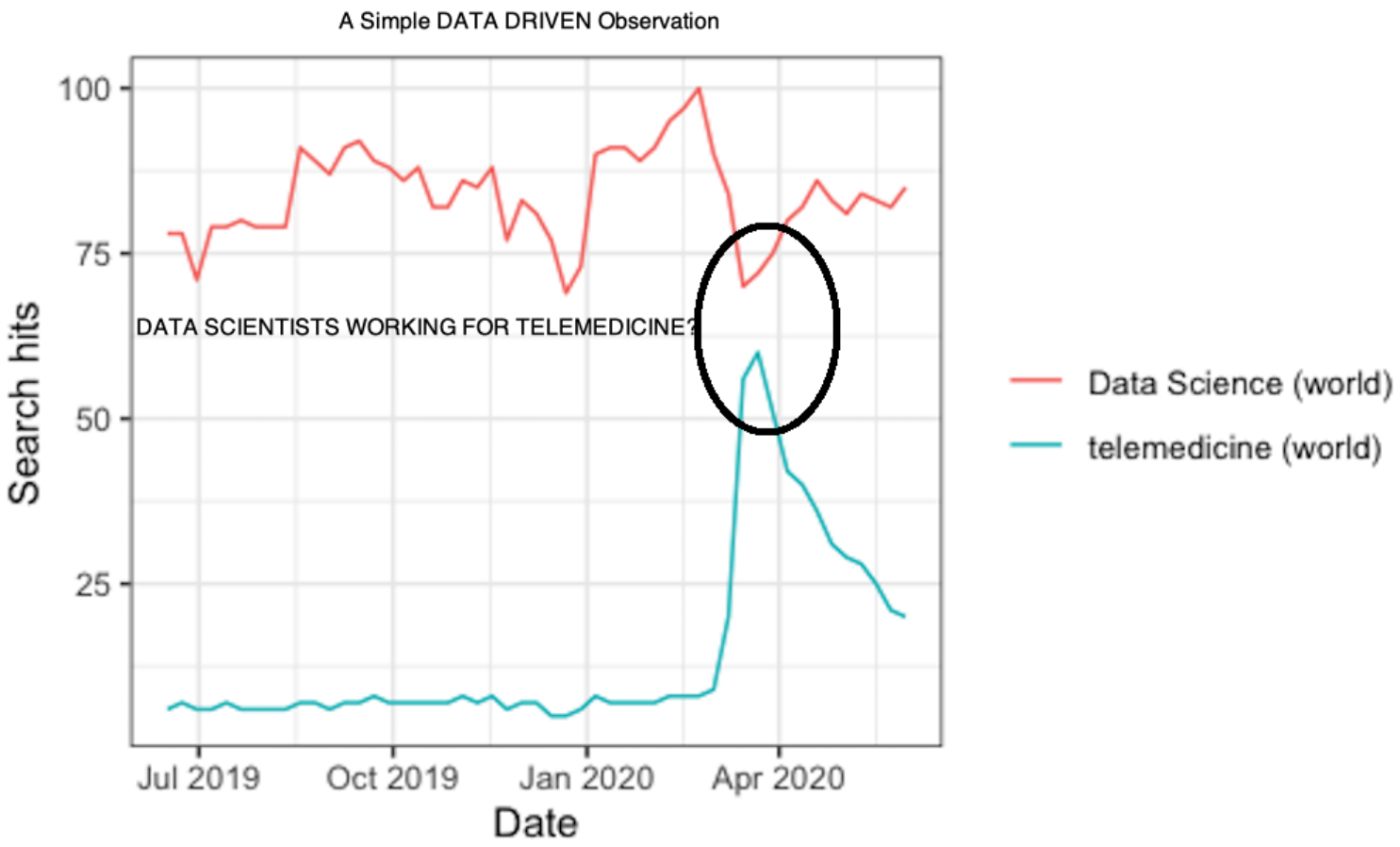


Figure 2: Data Scientsists engaged in modelling COVID? Lock Down Syndrome!

The e-health equation

Anjan Kr. Dasgupta
adgcal@gmail.com

- EH=e-health
- ICT= Information & Communication Technology(ICT)
- DS = Data Science

$$EH = ICT \bigcup DS \tag{1}$$

Moore's law in health care

Anjan Kr. Dasgupta
adgcal@gmail.com

https://thehealthcareblog.com/blog/2014/04/23/moores-law-in-healthcare-three-predictio/

- Moore's Law states that the number of transistors on a microchip doubles about every two years, though the cost of computers is halved.
- Kurzweil said, "The year 2029 is the consistent date I've predicted, when an artificial intelligence will pass a valid Turing test — achieving human levels of intelligence."

Newton's First Law applied to manufacturing and society

Anjan Kr. Dasgupta
adgcal@gmail.com

- Some things continue to advance, and that's the fact until (unless it is hit by something like CVD)

Systems Dynamics Modelling (SDM) and Agent based Models (ABM)

Anjan Kr. Dasgupta
adgcal@gmail.com

- https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-019-4627-7

SDM ADM details

Anjan Kr. Dasgupta
adgcal@gmail.com

- SDM simulates the movement of entities in a system (aggregate behaviour)
  - can be interpreted as mimicking the flow of water in and out of a bath tub.
  - how much 'water' (some resource) can leave or enter a 'bath tub' (a stock) as a function of environmental or operational) constrains
- ABM is a ground-up representation of a system, simulating the changing states of individual 'agents' in a system.

Why use SDM and ABM to model health systems?

Anjan Kr. Dasgupta
adgcal@gmail.com

- ABM is well-suited to explore systems with dynamic patient or health worker activity, a limitation of other differential equation or event-based simulation tools
- SDM is best implemented where the aim of the simulation is to examine aggregate flows, trends and sub-system behaviour as opposed to intricate individual flows of activity

Can AI ever beat the doctors ?

Anjan Kr. Dasgupta
adgcal@gmail.com

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6691444/
- Must refer to Penrose Hawkins debate ## Penrose, R. (1989). The emperor's new mind: Concerning computers, minds, and the laws of physics. Oxford University Press.

Roger Penrose

- Eminent physicist and winner, with Stephen Hawking, of the prestigious Wolf Prize—puts forward his view that **there are some facets of human thinking that can never be emulated by a machine**.
- Late Hawkins had the opposite view.

Data Data Everywhere not a drop to analyze

Anjan Kr. Dasgupta
adgcal@gmail.com

ICMR DataBase

National Centre for Disease Informatics and Research

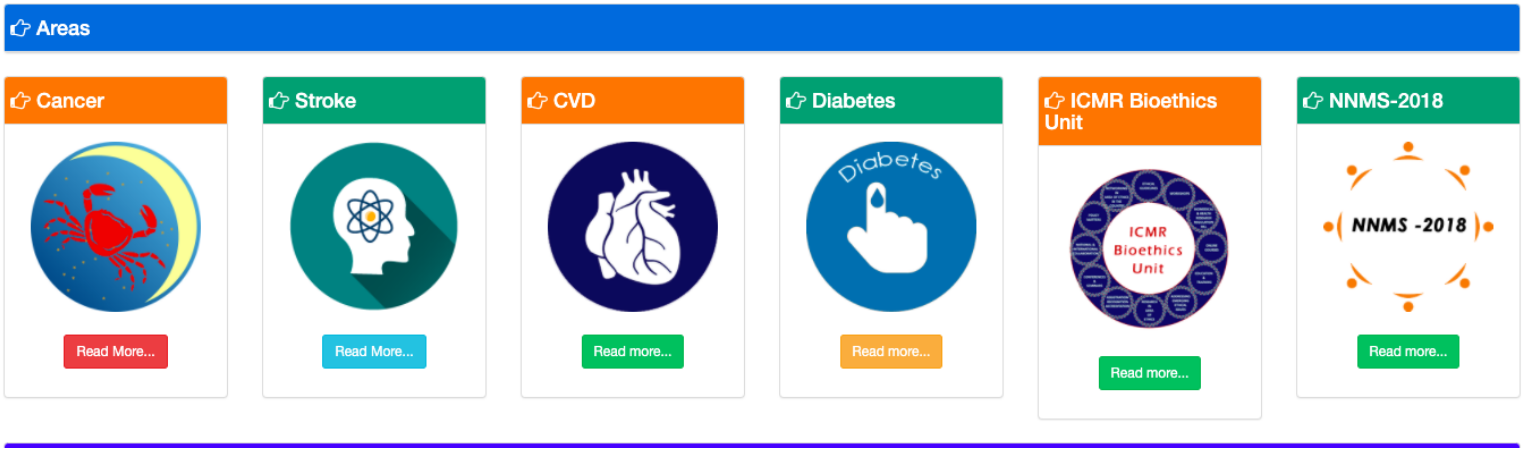


Figure 3: Intended data bases

Low volume of available data

Anjan Kr. Dasgupta
adgcal@gmail.com

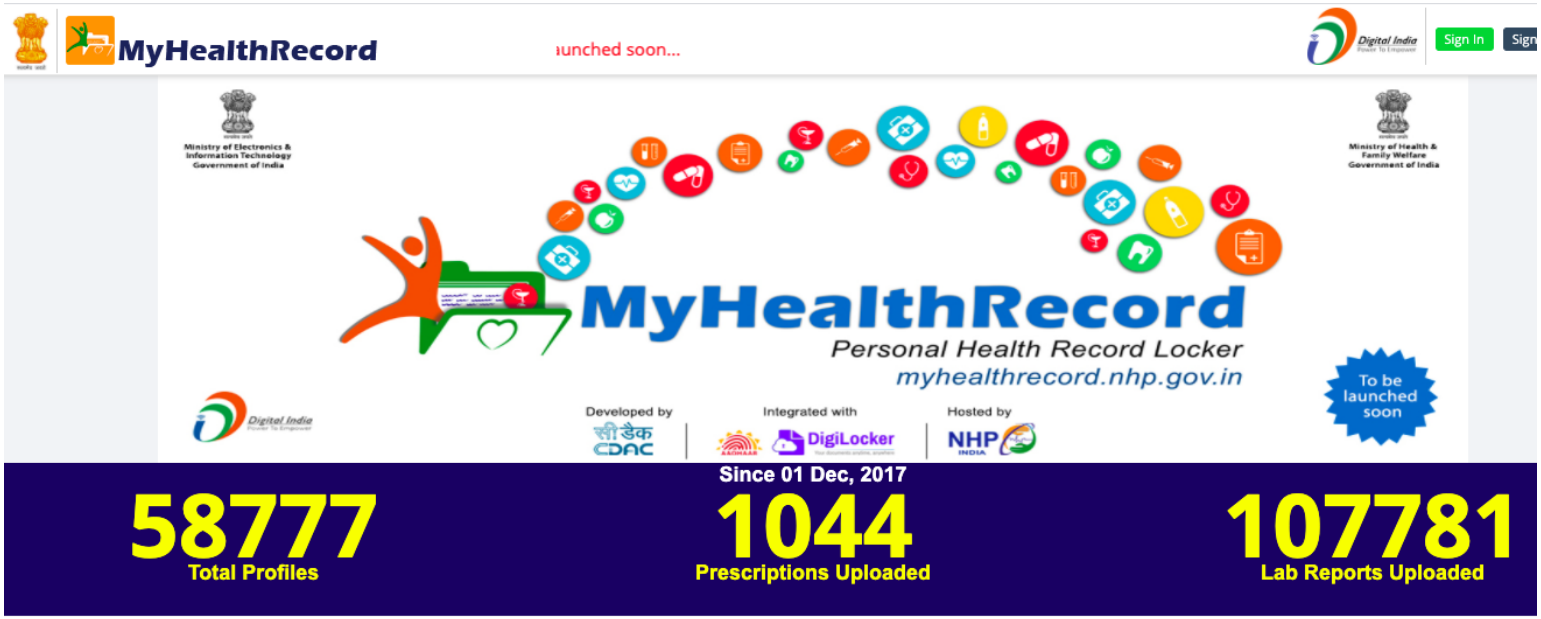
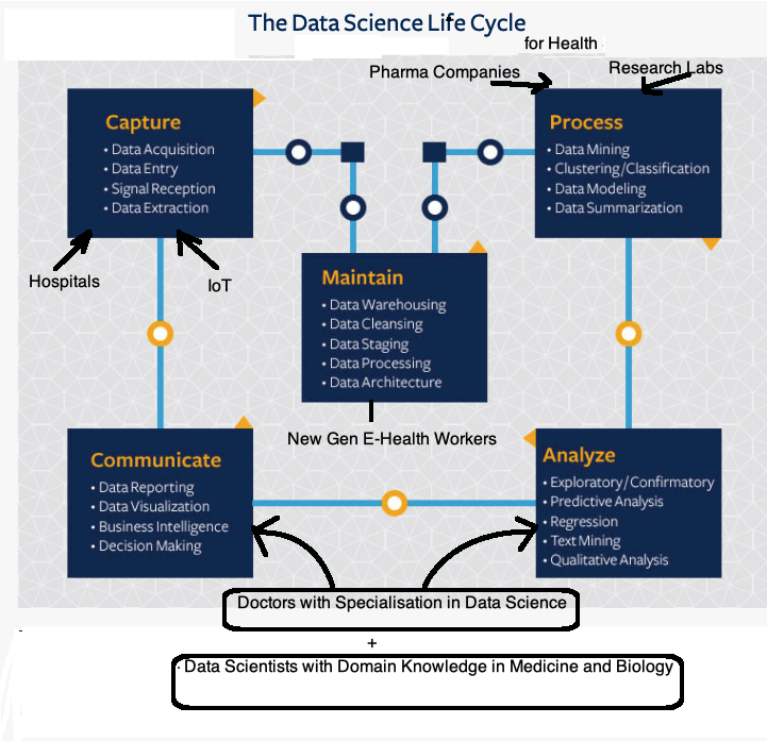


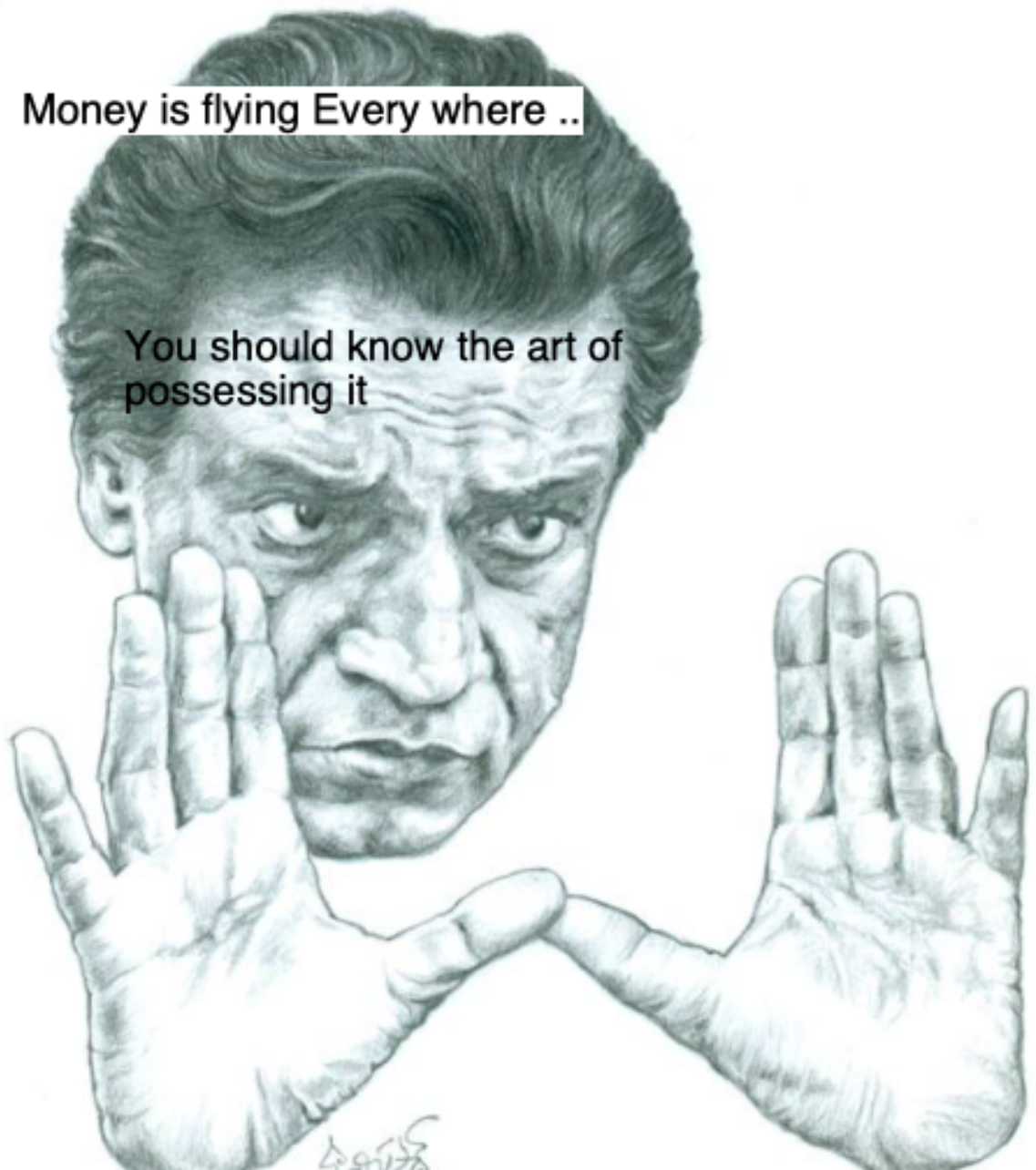
Figure 4: Low Volume records+AcessibilityIssues

Life Cycle of Health Data

Anjan Kr. Dasgupta
adgcal@gmail.com



Parash Pathar (Touch Stone)



Anjan Kr. Dasgupta
adgcal@gmail.com

The least e-Health can do - A quote from Nick Dawson

Anjan Kr. Dasgupta
adgcal@gmail.com

"Aravind Eye Hospital in India does more eye surgeries than any other place in the world. It treats nearly 2 million patients a year, for remarkably less than most hospitals in the United States, and it treats nearly two-thirds of those patients for free. If my Yelp app on my iPhone tells me FedEx can get raw fish to Aspen, we can surely get the blind to India."-ND

Pulling the COVID data

Anjan Kr. Dasgupta
adgcal@gmail.com

A queer looking output

Anjan Kr. Dasgupta
adgcal@gmail.com

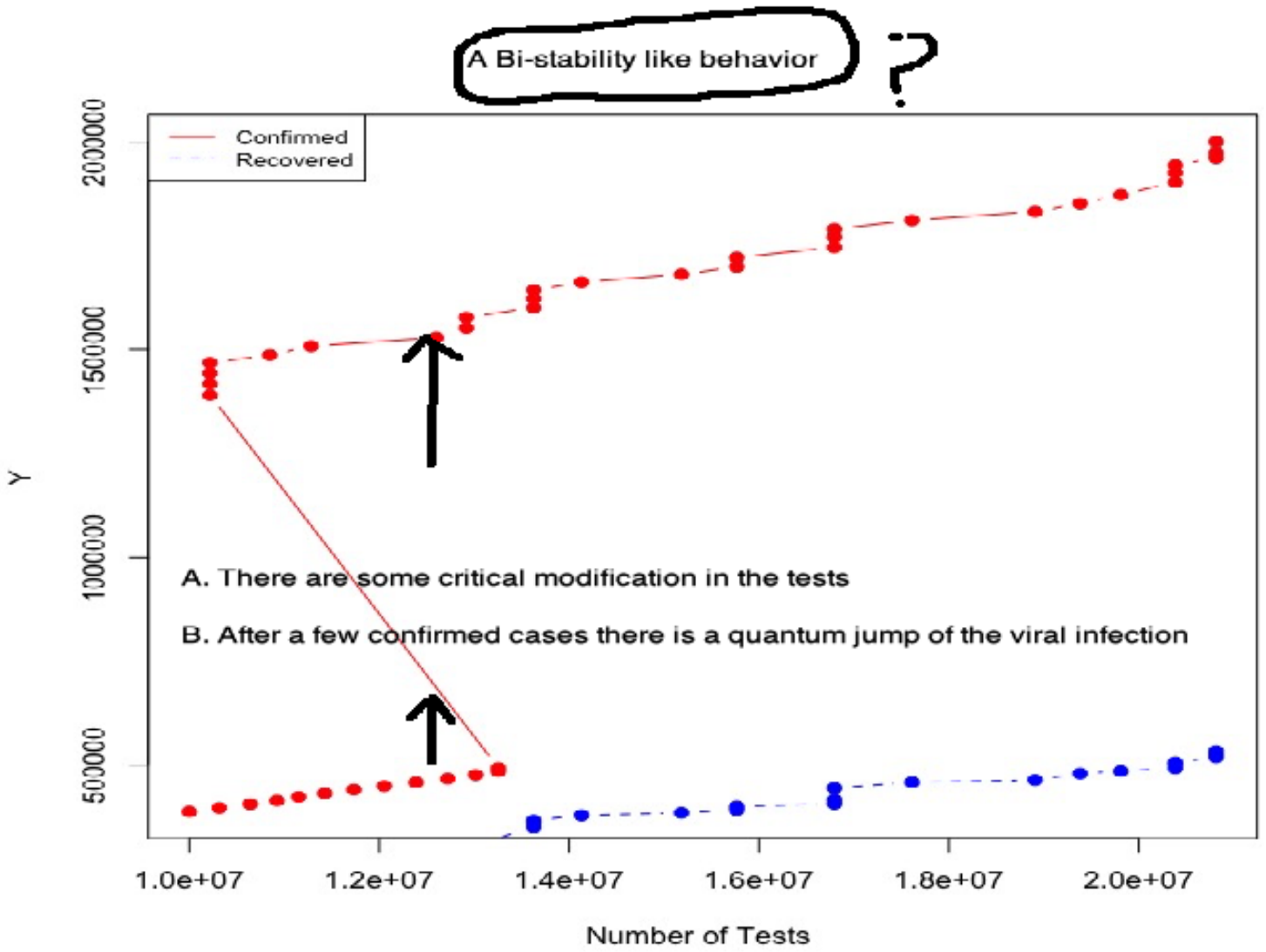


Figure 6: Ouput of the commented R code

Irreversible Bistability

Anjan Kr. Dasgupta
adgcal@gmail.com

Irreversible Bistability Posted on March 17, 2013 by Hsauro in **Analogue Machines - Cell Networks and Computation**

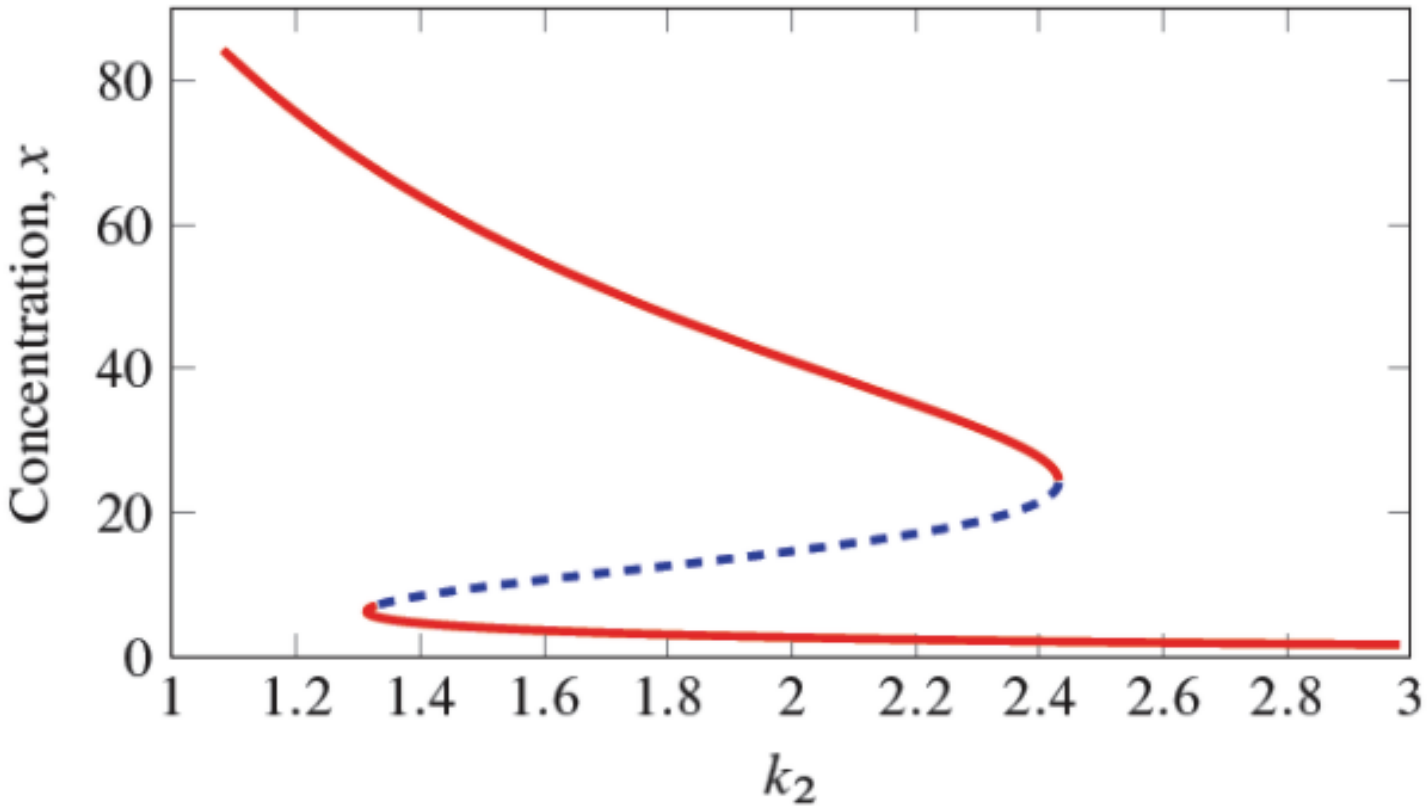


Figure 7: Here we see the state of the system (vertical axis) plotted as a function of a parameter in the system

A delay

Anjan Kr. Dasgupta
adgcal@gmail.com

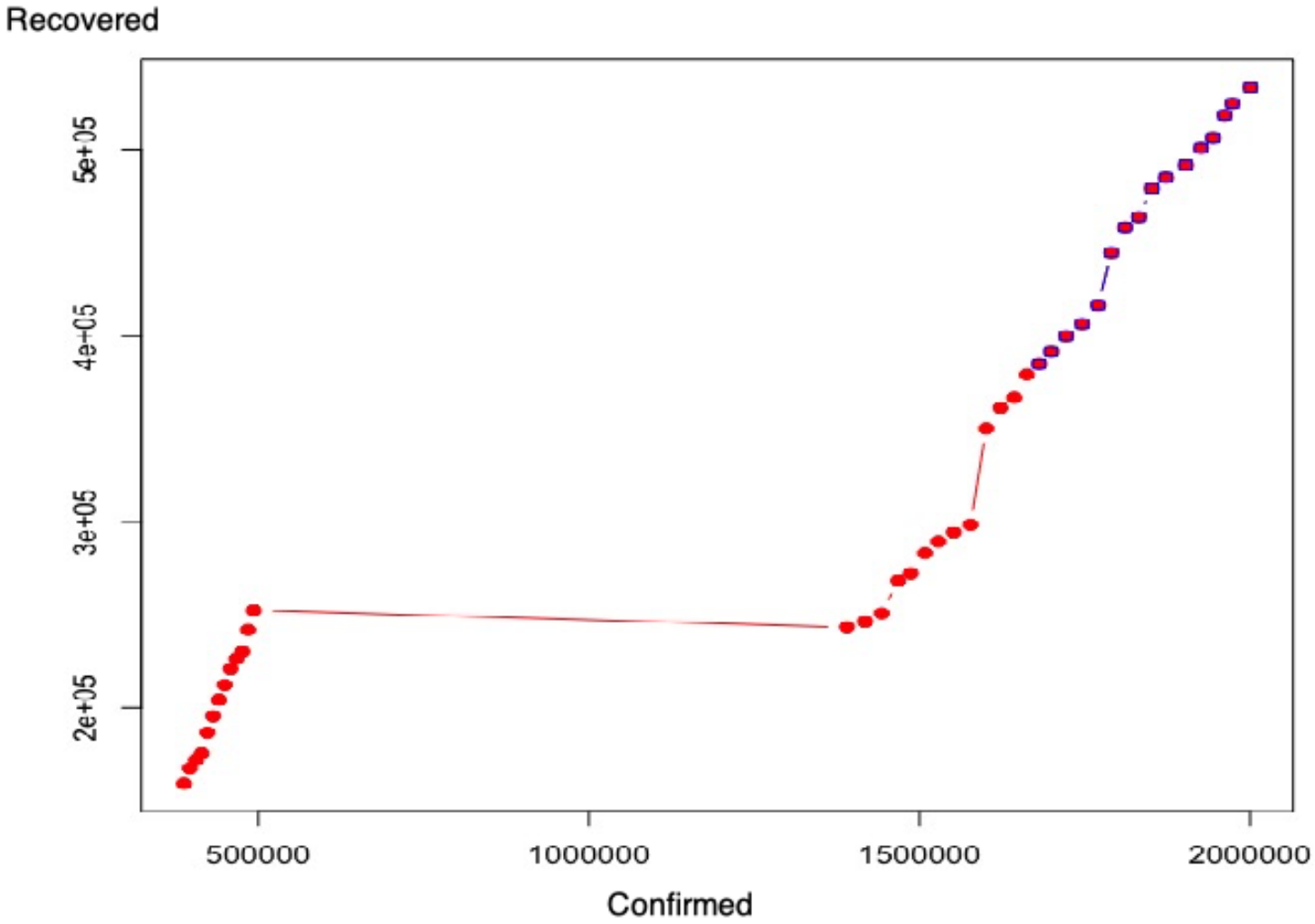
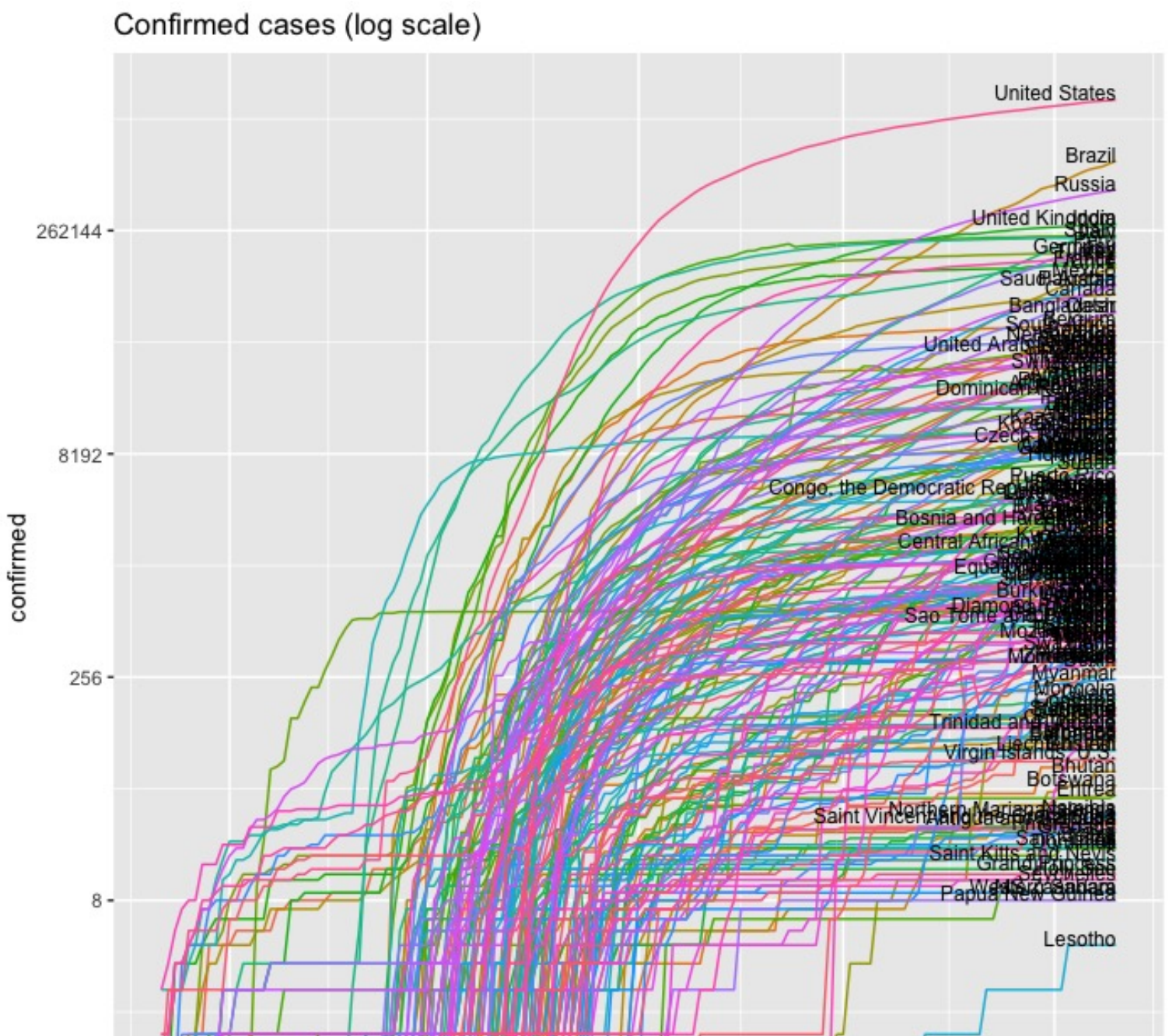


Figure 8: Rcovery process

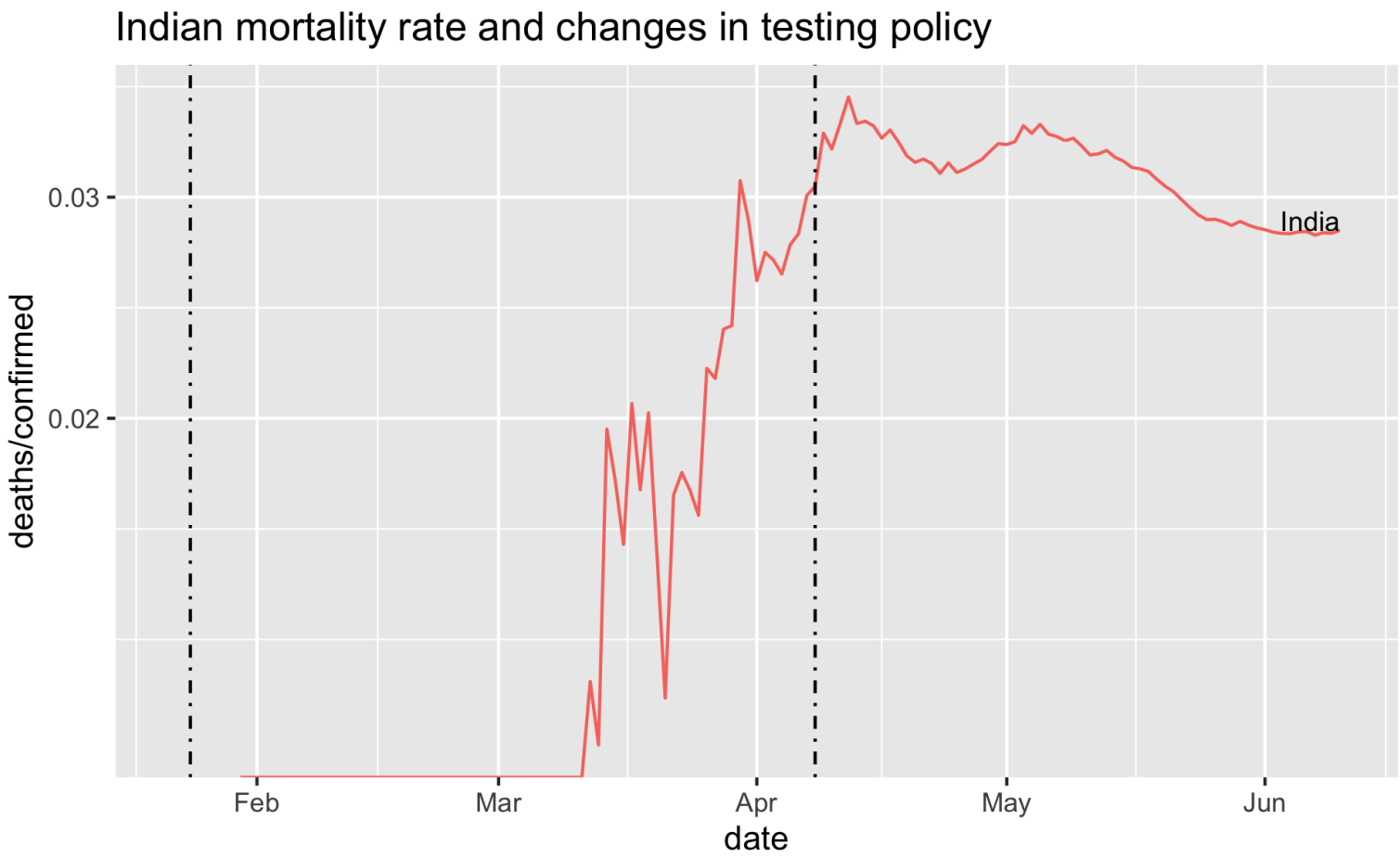
A Cumulative picture



Anjan Kr. Dasgupta
adgcal@gmail.com

The policy change effect

Anjan Kr. Dasgupta
adgcal@gmail.com



SIR Model

Anjan Kr. Dasgupta
adgcal@gmail.com

A simple mathematical description of the spread of a disease in a population is the so-called *SIR model*. (Kermack and McKendrick ,**Proc. R. Soc. A, 115, 772 1927**)

The story behind the SIR model

Anjan Kr. Dasgupta
adgcal@gmail.com

A simple mathematical description of the spread of a disease in a population is the so-called *SIR model*, ime t. - S(t) are those susceptible but not yet infected with the disease; - I(t) is the number of infectious individuals; - R(t) are those individuals who have recovered from the disease and now have immunity to it.

Equations

Anjan Kr. Dasgupta
adgcal@gmail.com

$$\frac{dS}{dt} = -\beta \cdot S \cdot I/N \tag{2}$$

$$\frac{dI}{dt} = \beta \cdot S \cdot I/N - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

Running Python as Markdown chunk

Anjan Kr. Dasgupta
adgcal@gmail.com

```
library(reticulate)
```

Importing python libraries

Anjan Kr. Dasgupta
adgcal@gmail.com

```
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt
```

Initialization

Anjan Kr. Dasgupta
adgcal@gmail.com

```
# Total population, N.
N = 1000
# Initial number of infected and recovered individuals, I0 an
I0, R0 = 1, 0
# Everyone else, S0, is susceptible to infection initially.
S0 = N - I0 - R0
# Contact rate, beta, and mean recovery rate, gamma, (in 1/da
beta, gamma = 0.8, 3./10
t = np.linspace(0, 40, 160)
```

The model in python

```
def deriv(y, t, N, beta, gamma):
    S, I, R = y
    dSdt = -beta * S * I / N
    dIdt = beta * S * I / N - gamma * I
    dRdt = gamma * I
    return dSdt, dIdt, dRdt
```


Anjan Kr. Dasgupta
adgcal@gmail.com

Simulating the SIR model

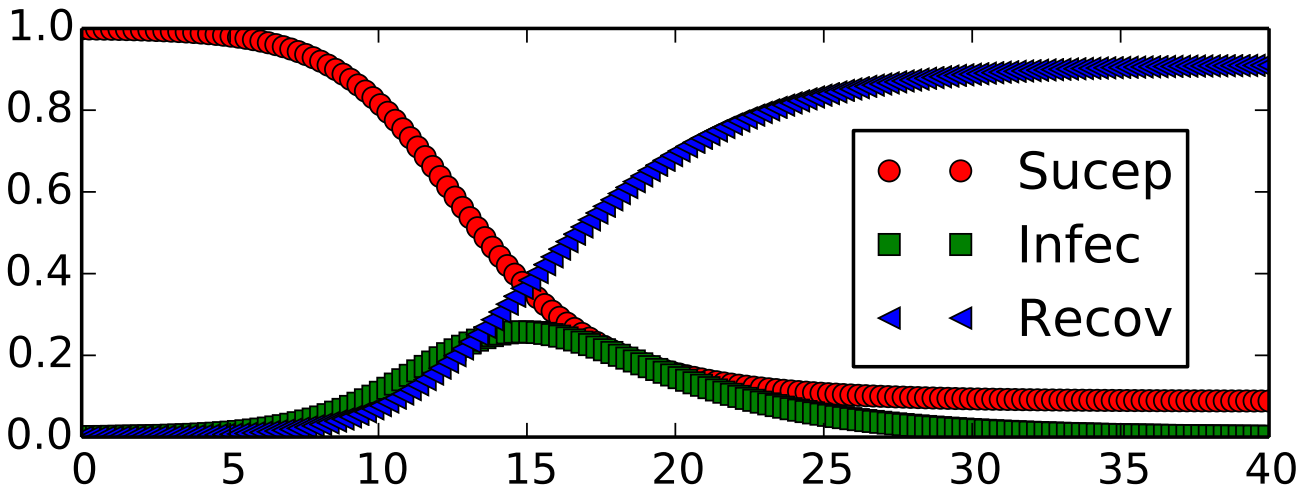
Anjan Kr. Dasgupta
adgcal@gmail.com

```
y0 = S0, I0, R0
# Integrate the SIR equations over the time grid, t.
ret = odeint(deriv, y0, t, args=(N, beta, gamma))
S, I, R = ret.T
```

Ploting the model variables against TIME (AU)

Anjan Kr. Dasgupta
adgcal@gmail.com

```
plt.figure(figsize=(6,2))
plt.plot(t,S/1000,'or',label='Sucep')
plt.plot(t,I/1000,'sg',label='Infec')
plt.plot(t,(1000-S-I)/1000,'<b',label='Recov')
plt.legend(bbox_to_anchor=(0.65, 0.75),\
  loc='upper left', borderaxespad=0.)
```



Sympy

Anjan Kr. Dasgupta
adgcal@gmail.com

The analytical expression for Jacobian of the SIR model assumes

$$\begin{bmatrix} -\frac{I\beta}{N} & -\frac{S\beta}{N} \\ \frac{I\beta}{N} & \gamma + \frac{S\beta}{N} \end{bmatrix}$$

The steady stae of the SIR

Anjan Kr. Dasgupta
adgcal@gmail.com

s_S = sympy.Eq(dSdt, 0) s_I = sympy.Eq(dIdt, 0) s_R = sympy.Eq(dRdt, 0)

s_S_sol = sympy.solve(s_S, S) s_I_sol = sympy.solve(s_I,I) s_R_sol = sympy.solve(s_R,R)

print('The steady state solution for the different populations are:') s_S_sol,s_I_sol,s_R_sol

The steady state solution for the different populations are:0,0. This actually implies S=S(ss),I=I(ss)

Extended SIR Model

Anjan Kr. Dasgupta
adgcal@gmail.com

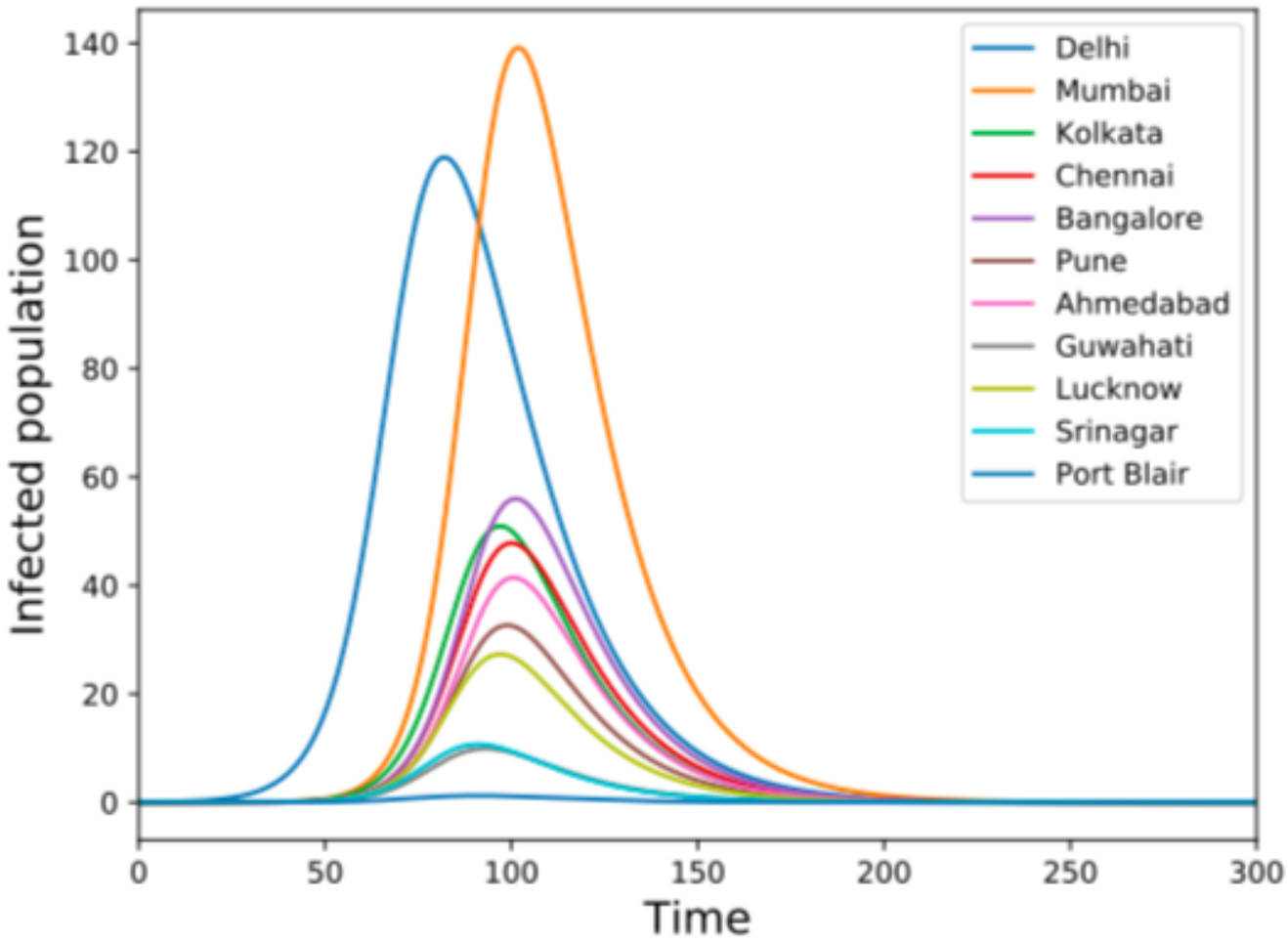
It may be noted that though SIR model model is simplistic the pattern remains similar for complicated cases (Multi-city modeling of epidemics using spatial networks: Application to 2019-nCov (COVID-19) coronavirus in India< https://doi.org/10.1101/2020.03.13.20035386>)

$$\frac{ds_i}{dt} = -\beta \frac{s_i(t)x_i(t)}{\phi_i} - \eta_i s_i(t) + \sum_j \frac{A_{ij}\eta_j}{k_j} s_j(t-\delta_{ij}),$$

$$\frac{dx_i}{dt} = \beta \frac{s_i(t)x_i(t)}{\phi_i} - \gamma x_i(t) - \eta_i x_i(t) + \sum_j \frac{A_{ij}\eta_j}{k_j} x_j(t-\delta_{ij}),$$

$$\frac{dr_i}{dt} = \gamma x_i(t) - \eta_i r_i(t) + \sum_j \frac{A_{ij}\eta_j}{k_j} r_j(t-\delta_{ij}),$$

Output of the connected model



Anjan Kr. Dasgupta
adgcal@gmail.com

The SEIR Story

Anjan Kr. Dasgupta
adgcal@gmail.com

Many diseases have a latent phase during which the individual is infected but not yet infectious. This delay between the acquisition of infection and the infectious state can be incorporated within the SIR model by adding a latent/exposed population, E, and letting infected (but not yet infectious) individuals move from S to E and from E to I.

SEIR without vital dynamics

Anjan Kr. Dasgupta
adgcal@gmail.com

In a closed population with no births or deaths, the SEIR model becomes:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{5}$$

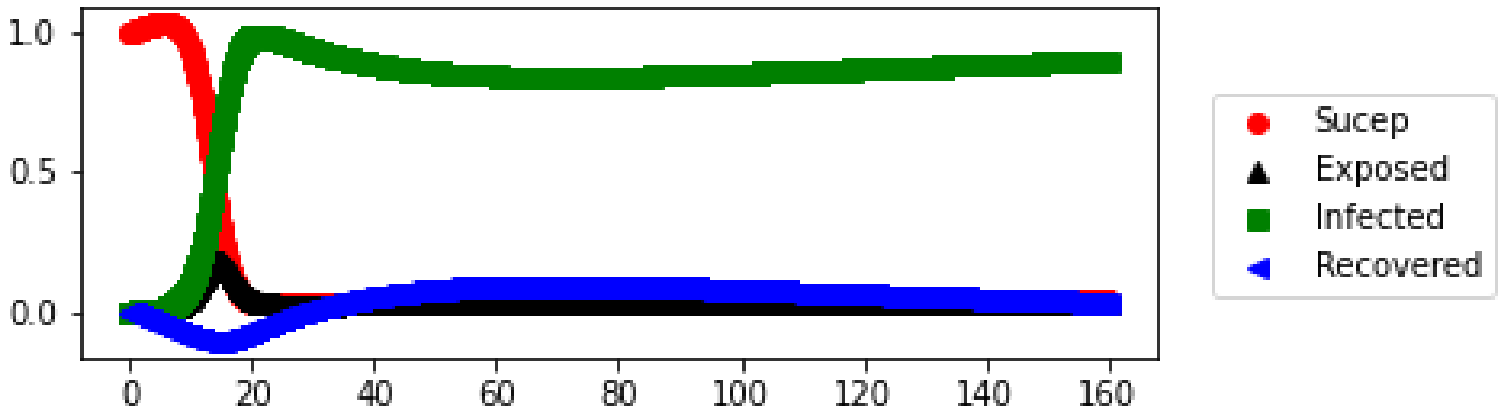
$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{6}$$

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{7}$$

$$\frac{dR}{dt} = \gamma I \tag{8}$$

where $N = S + E + I + R$ is the total population.

Numerical Simulation to the SEIR model



Anjan Kr. Dasgupta
adgcal@gmail.com

Analytical Solution to the SEIR model

Anjan Kr. Dasgupta
adgcal@gmail.com

The steady state solution for the different populations are:

$$\left(\left[\frac{N\left(N\mu + R\zeta\right)}{I\beta + N\nu}\right], \quad \left[\frac{IS\beta}{N\left(\nu + \sigma\right)}\right], \quad \left[\frac{E\sigma}{\gamma + \nu}\right], \quad \left[\frac{I\gamma}{\nu + \zeta}\right]\right)$$

- Point to note : excepting for S , no dependence on N.

Implications of the SEIR

Anjan Kr. Dasgupta
adgcal@gmail.com

- Longer latency period will result in slower initial growth of the outbreak.
- However, since the model does not include mortality, the basic reproductive number, $R0 = \beta/\gamma$, does not change.
- After the initial fast growth, the epidemic depletes the susceptible population.
- Eventually the virus cannot find enough new susceptible people and dies out.
- Introducing the incubation period does not change the cumulative number of infected individuals.

Jacobian of the SEIR

$$\begin{bmatrix} -\frac{I\beta}{N} - \nu & 0 & -\frac{S\beta}{N} & \zeta \\ \frac{I\beta}{N} & -\nu - \sigma & \frac{S\beta}{N} & 0 \\ 0 & \sigma & -\gamma - \nu & 0 \\ 0 & 0 & \gamma & -\nu - \zeta \end{bmatrix}$$

Anjan Kr. Dasgupta
adgcal@gmail.com

Other Models (SIR & SEIR type - E=EXPOSED)

Anjan Kr. Dasgupta
adgcal@gmail.com

- The first is a model developed by scientists at the Indian Council of Medical Research (ICMR) and their collaborators. *ICMR study=SEIR*.

Other Models (SIR & SEIR type - E=EXPOSED)

Anjan Kr. Dasgupta
adgcal@gmail.com

- The first is a model developed by scientists at the Indian Council of Medical Research (ICMR) and their collaborators. *ICMR study=SEIR*.
- The second is a model produced by a group of epidemiologists and statisticians largely from the University of Michigan. *Michigan study=SIR*.

Other Models (SIR & SEIR type - E=EXPOSED)

Anjan Kr. Dasgupta
adgcal@gmail.com

- The first is a model developed by scientists at the Indian Council of Medical Research (ICMR) and their collaborators. *ICMR study=SEIR*.
- The second is a model produced by a group of epidemiologists and statisticians largely from the University of Michigan. *Michigan study=SIR*.
- The third is a set of reports published by the Centre for Disease Dynamics, Economics and Policy *(CDDEP) Hopkins study=ABM=Data driven*.

Other Models (SIR & SEIR type - E=EXPOSED)

Anjan Kr. Dasgupta
adgcal@gmail.com

- The first is a model developed by scientists at the Indian Council of Medical Research (ICMR) and their collaborators. *ICMR study=SEIR*.
- The second is a model produced by a group of epidemiologists and statisticians largely from the University of Michigan. *Michigan study=SIR*.
- The third is a set of reports published by the Centre for Disease Dynamics, Economics and Policy *(CDDEP) Hopkins study=ABM=Data driven*.
- Finally, there is a recent study from scientists at Cambridge University *Cambridge study=SIR*.

Fitzhugh-Nagumo model: an excitable system

Anjan Kr. Dasgupta
adgcal@gmail.com

The Fitzhugh-Nagumo model of an excitable system is a two-dimensional simplification of the Hodgkin-Huxley model of spike generation in squid giant axons.

$$\frac{dv}{dt} = v - v^3 - w + I_{ext} \tag{9}$$

$$\tau \frac{dw}{dt} = v - a - bw \tag{10}$$

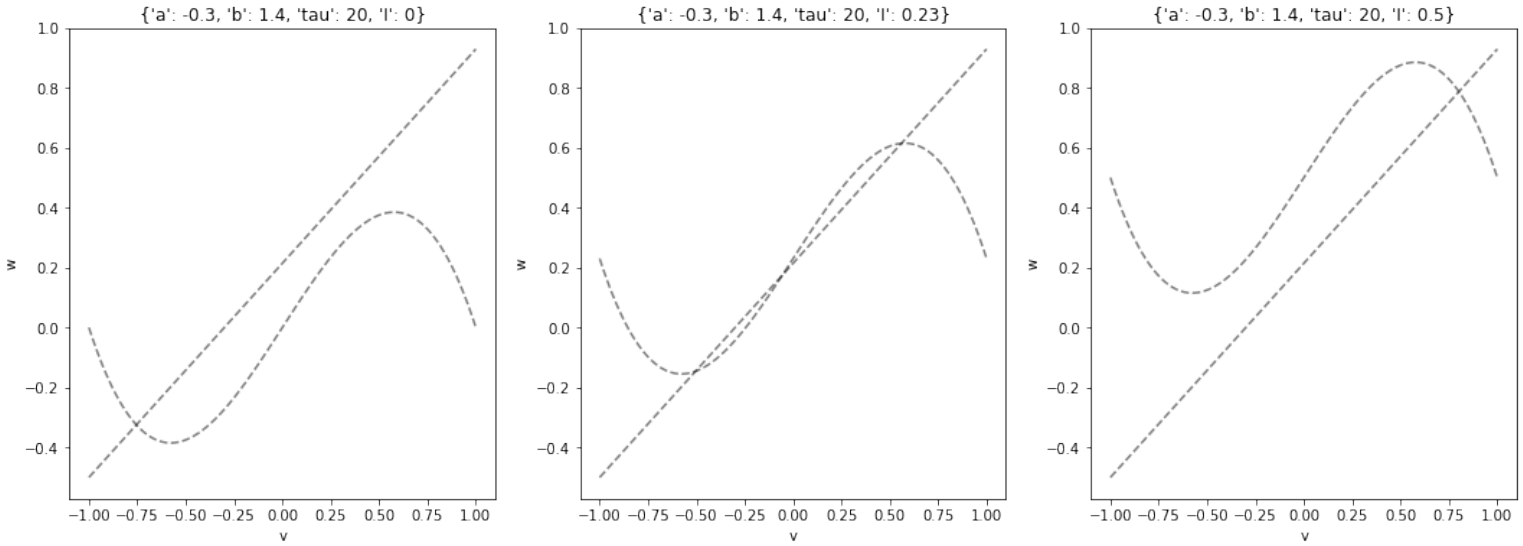
Isocline plot

Anjan Kr. Dasgupta
adgcal@gmail.com

```
def plot_isocline(ax, a, b, tau, I, color='k', style='-',
opacity=.5, vmin=-1,vmax=1): """Plot the null iscolines of
the Fitzhugh nagumo system""" v =
np.linspace(vmin,vmax,100) ax.plot(v, v - v**3 + I, style,
color=color, alpha=opacity) ax.plot(v, (v - a)/b, style,
color=color, alpha=opacity)
```

The output

Anjan Kr. Dasgupta
adgcal@gmail.com



Digitization of ECG Signal

Anjan Kr. Dasgupta
adgcal@gmail.com

- Researchers want to analyze the actual value of signals, i.e., the value of this ECG signal in milivolts.
- But to process information using computers, they must collect the signals via some capturing device which discretely samples, and digitizes the signals into 2^N levels, where N is the resolution of the device.
- Each sample captured requires N bits to store, and takes one of 2^N possible integer values. - There is also information stored which allows the user/program to map these integers back to the physical values the device managed to capture given its resolution.
- For example if they have a 12 bit oscilloscope, they have 4096 levels to capture the range and details of the signal.
- A higher N allows us to resolve finer details, but requires more storage space per sample.

PhysioNet

Anjan Kr. Dasgupta
adgcal@gmail.com

- For e-health this is what NCBI is to *Bioinformatics*.This is supported by the **National Institute of General Medical Sciences (NIGMS) and the National Institute of Biomedical Imaging and Bioengineering (NIBIB)** under NIH grant number 2R01GM104987-09.
- The following site provides deeper insights into the data format used in physionet

http://archive.physionet.org/faq.shtml#physiobank-formats

Overview of Physionet

Anjan Kr. Dasgupta
adgcal@gmail.com

- Clinical Databases: Data from critical care clinical settings
- Waveform Databases: High resolution continuous recordings of physiological signals
- Waveform databases are organized according to their signal and annotation types.
- Multi-Parameter Databases: Available signals vary, but may include ECG, continuous invasive blood pressure, respiration, oxygen saturation, and EEG, among others.
- Physionet provides the **WFDB software package** highly useful for reading, writing, and processing the above described WFDB files.
- The popular WFDB platforms are MATLAB & PYTHON

MIT Format

Anjan Kr. Dasgupta
adgcal@gmail.com

- MIT Signal files **(.dat)** are binary files containing samples of digitized signals. These store the waveforms, but they cannot be interpreted properly without their corresponding header files.
- MIT Header files **(.hea)** are short text files that describe the contents of associated signal files.
- Annotation files **(.atr)** . Thr filr record.dat may be accompanied with RECORDNAME.atr.

European Data Format (EDF)

Anjan Kr. Dasgupta
adgcal@gmail.com

- EDF files contain digital signals stored in their standard international format.
- EDF files store their header information at the beginning of the file, as opposed to MIT format which has a separate header file.
- EDF is a WFDB and PhysioBank-compatible format.
- If a directory contains RECORDNAME.edf and RECORDNAME.edf.qrs, the .qrs file is the annotation file associated with the record.

A visit

https://physionet.org/lightwave/?db=mitdb/1.0.0

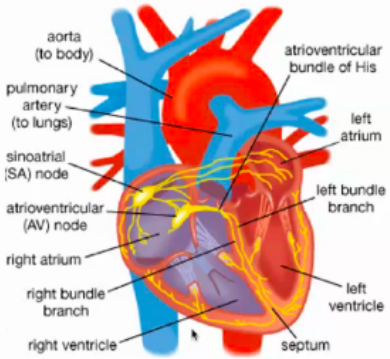
Anjan Kr. Dasgupta
adgcal@gmail.com

ECG digitization - A success story

Anjan Kr. Dasgupta
adgcal@gmail.com

Modelling framework

Model the pacemaker and the heart, compose and verify



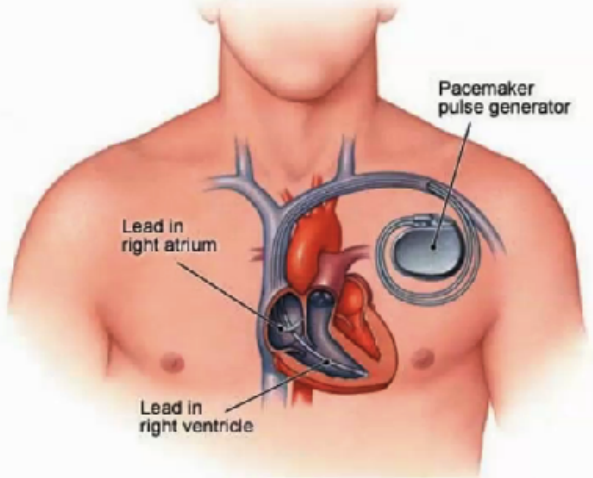
All may not be well for AI

Anjan Kr. Dasgupta
adgcal@gmail.com

Viewing Marta Kwiatkow...

Case study: Cardiac pacemaker

- How it works
  - reads electrical signals through sensors in the right atrium and right ventricle
  - monitors the timing of heart beats and local electrical activity
  - generates artificial pacing signal as necessary



- Safety-critical system!
- The guarantee
  - (basic safety) maintain 60–100 beats per minute

– Killed by code: FDA recalls 23 defective pacemaker devices because of adverse health consequences or death, six likely caused by software defects (2010)

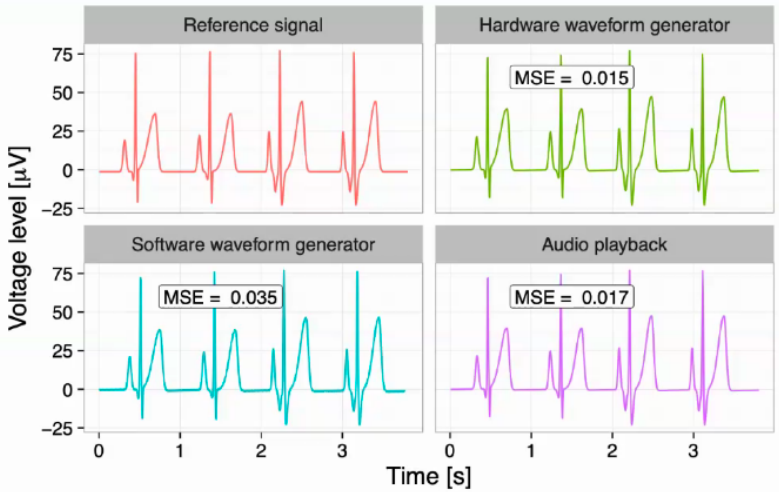
From heart to cybersecurity

Anjan Kr. Dasgupta
adgcal@gmail.com

Viewing Marta Kwiatkow...

Attack on ECG biometrics

- We use synthetic ECGs to impersonate a user
  - build model from data, 41 volunteers
  - inject synthetic signals to break authentication
  - 80% success rate
- Results
  - serious weakness
  - countermeasures needed



- Modelling essential, good for attacks...

Can AI ever beat the doctors ?

Anjan Kr. Dasgupta
adgcal@gmail.com

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6691444/
- https://interestingengineering.com/these-7-ai-powered-doctor-phone-apps-could-be-the-future-of-healthcare

Data Silos

Anjan Kr. Dasgupta
adgcal@gmail.com

- A data silo is a collection of information in an organization that is isolated from and not accessible by other parts of the organization.
- Removing data silos can help you get the right information at the right time so you can make good decisions.
- And, you can save money by reducing storage costs for duplicate information.

About healthcare today, we talk about the silos a lot

Anjan Kr. Dasgupta
adgcal@gmail.com

- Very rarely does infrastructure technology create the hype that blockchain currently has
- The challenges that blockchain addresses in healthcare are very disruptive

Block Chain in Healthcare

Anjan Kr. Dasgupta
adgcal@gmail.com

- Synaptic Health Alliance=Optum+Humana+MultiPlan+Quest
- Aetna+ Ascension
- Indian telemedicine https://vhealth.io/ by Aetna

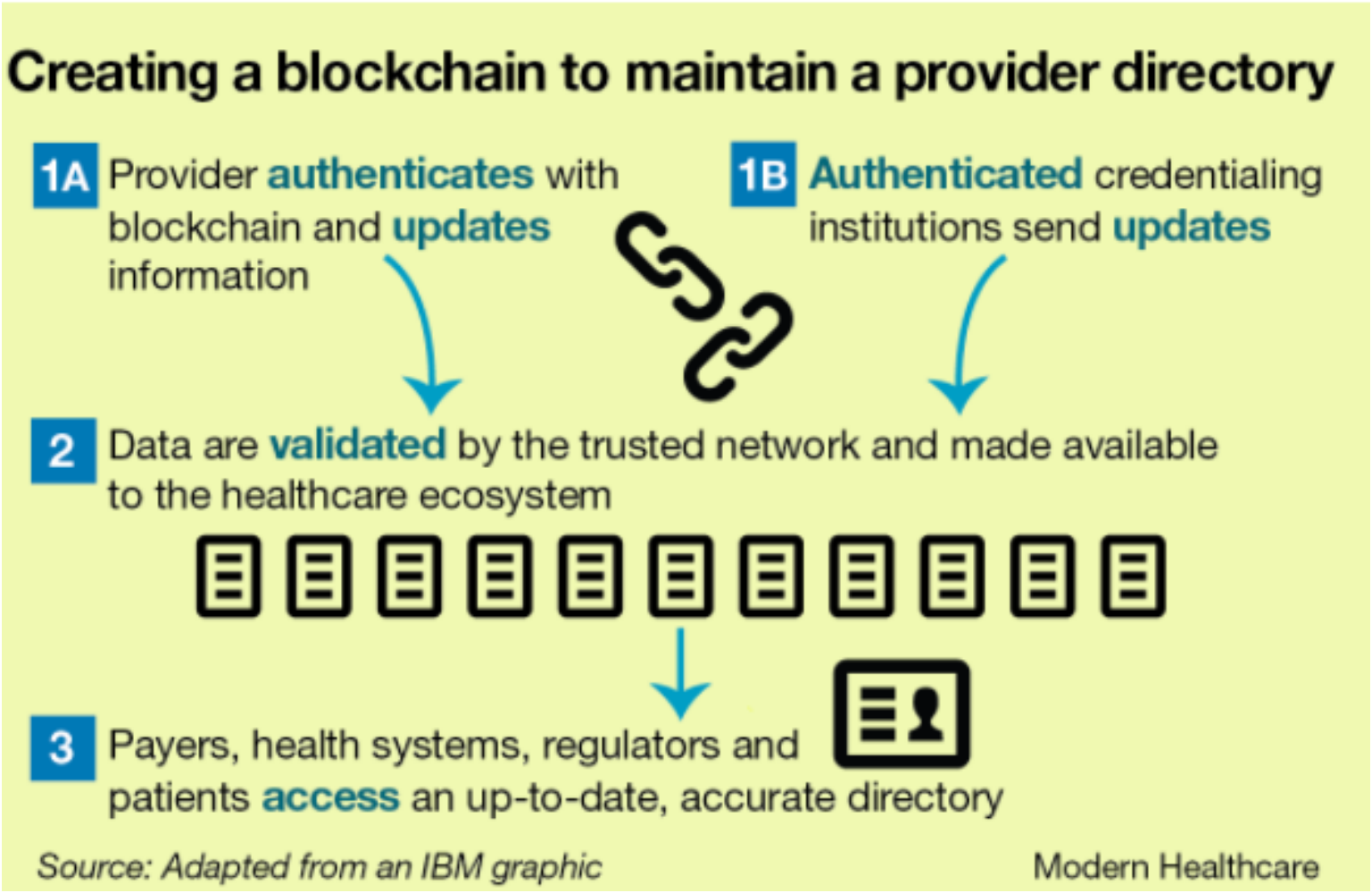
Prophecy:Blockchain next dramatic innovation in Healthcare

Anjan Kr. Dasgupta
adgcal@gmail.com

- Multiple people looking at the same information
- Quality of that information should go up
- Operational costs for the provider and payers lesser
- should go down because there's less-frequent contact being done between those two stakeholders. -blockchain allows us to connect those silos and . . . enable new capabilities (so that) access to information no longer is where we compete

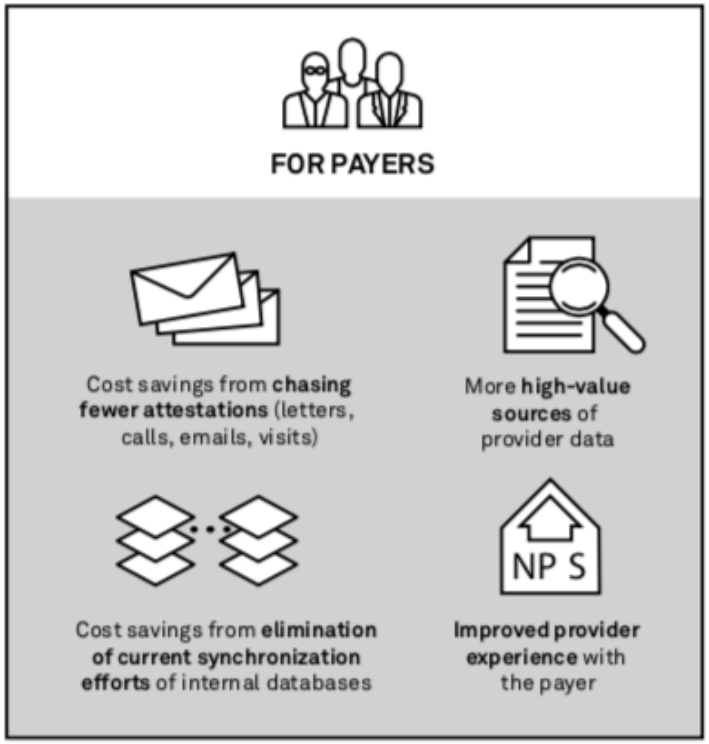
Potential benifits of Blockchain Technology in Healthcare

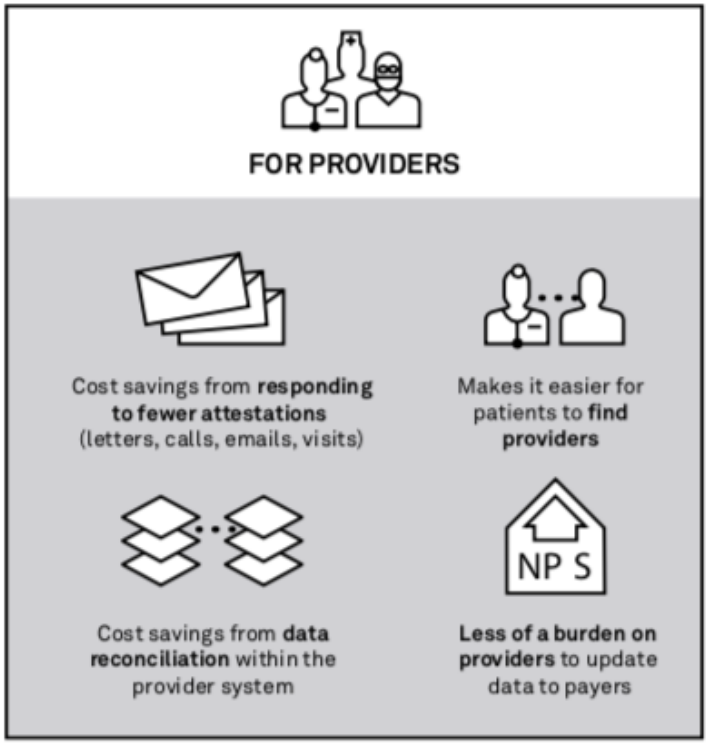
Anjan Kr. Dasgupta
adgcal@gmail.com



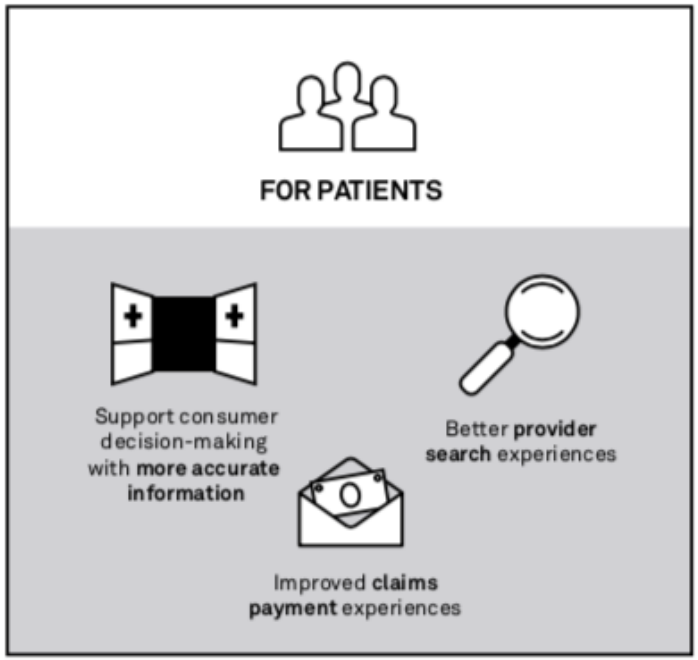
Advantages

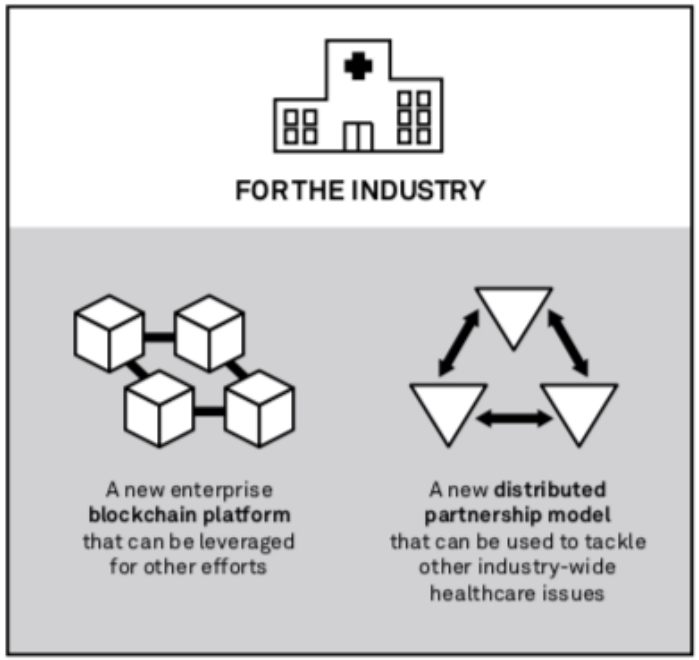
Anjan Kr. Dasgupta
adgcal@gmail.com





Advantages





Anjan Kr. Dasgupta
adgcal@gmail.com

An interactive Shiny Application

Anjan Kr. Dasgupta
adgcal@gmail.com

https://adgcal.shinyapps.io/covid/

A quiz for all

Anjan Kr. Dasgupta
adgcal@gmail.com

https://adgcal.shinyapps.io/EH_mcq/

Lastly

Anjan Kr. Dasgupta
adgcal@gmail.com