

COVID-19 RL Project Proposal

Team: Nicky Zhang, Devon Williams, Shirley (Siyue)
Huang, Matthew Taj, Andrew Massoud, Dr. Tom Woolf



Introduction: COVID-19

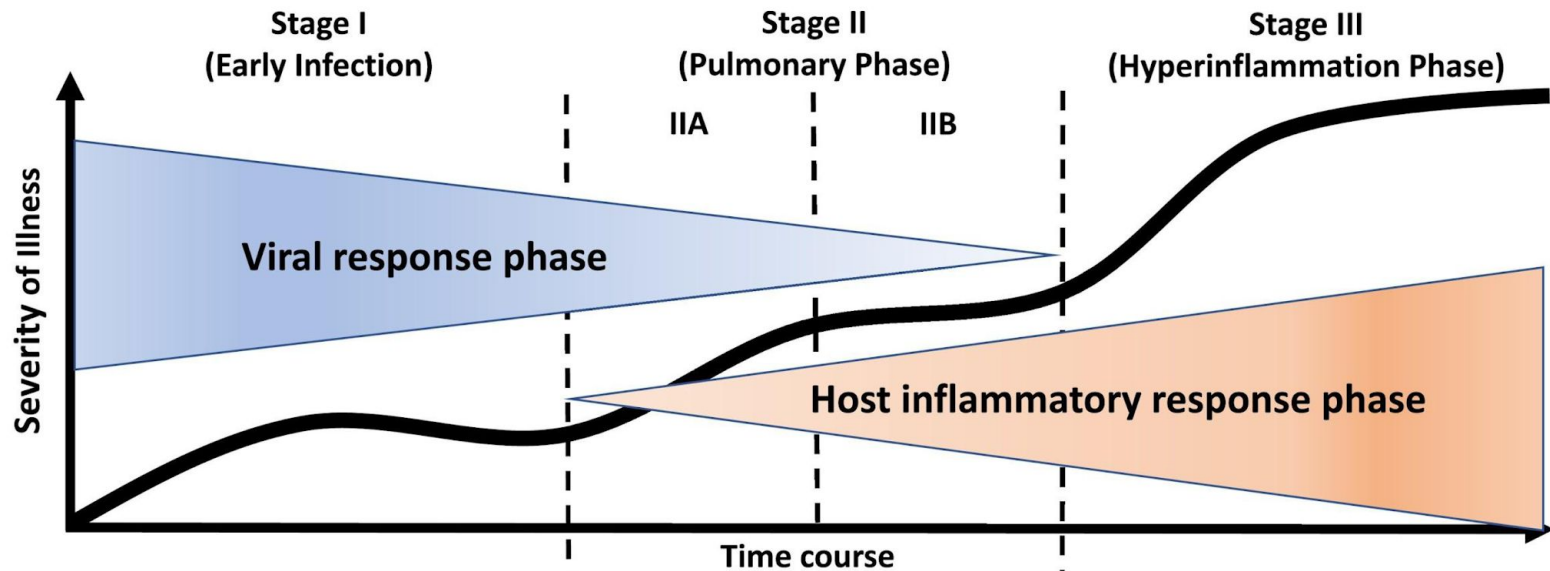
Disease Trajectory

- N3C = national registry of data
- Understand trajectory of disease
- What decisions impact the states of COVID patients the most?



National
COVID
Cohort
Collaborative





Clinical Symptoms	Mild constitutional symptoms Fever >99.6°F Dry Cough	Shortness of Breath without (IIA) and with Hypoxia (IIB) (PaO ₂ /FiO ₂ ≤ 300 mmHg)	ARDS SIRS/Shock Cardiac Failure
Clinical Signs	Lymphopenia	Abnormal chest imaging Transaminitis Low-normal procalcitonin	Elevated inflammatory markers (CRP, LDH, IL-6, D-dimer, ferritin) Troponin, NT-proBNP elevation
Potential Therapies	Remdesivir, chloroquine, hydroxychloroquine, convalescent plasma transfusions		
	Reduce immunosuppression (avoid excess steroids)	Careful use of Corticosteroids; statins; human immunoglobulin, IL-1/IL-2/IL-6/JAK inhibitors/GM-CSF Inhibitors	

Role of ML in COVID-19

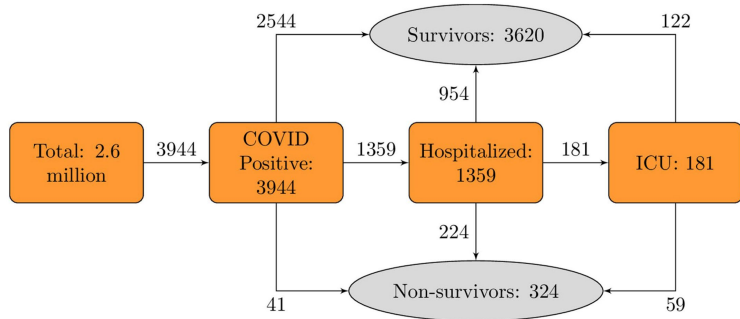
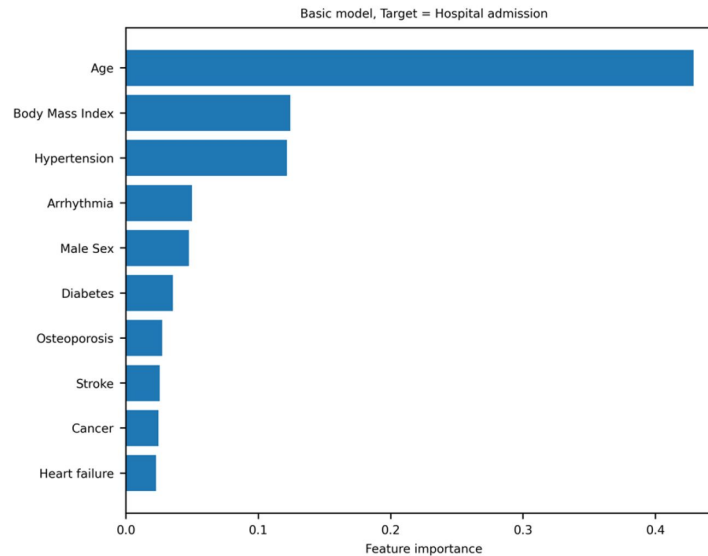


Figure 1. Flowchart of patient selection and identification for the Danish cohort. ICU intensive care unit.





Data source: Electronic Health Records (EHRs)

- Interconnected aggregate of all the patients' health records
- Can be used to track disease progression -> a possibility to be utilized in RL to help clinicians design better policies (i.e., treatment plans).

```
GpIda945 O'Haral6
-----
Race: White
Ethnicity: Non-Hispanic
Gender: F
Age: 45
Birth Date: 1971-10-04
Marital Status: M
-----
ALLERGIES: N/A
-----
MEDICATIONS:
2015-09-14 [CURRENT] : 3 ML liraglutide 6 MG/ML Pen Injector
2014-11-23 [STOPPED] : canagliflozin 100 MG Oral Tablet
2014-11-23 [STOPPED] : 3 ML liraglutide 6 MG/ML Pen Injector
2014-11-23 [CURRENT] : 24 HR Metformin hydrochloride 500 MG Extended Release Oral Tablet
2010-11-30 [STOPPED] : Amoxicillin 250 MG / Clavulanate 125 MG [Augmentin] for Viral sinusitis (disorder)
2007-07-05 [STOPPED] : Amoxicillin 250 MG / Clavulanate 125 MG [Augmentin] for Sinusitis (disorder)
-----
CONDITIONS:
2014-11-23 - : Diabetes
2014-01-10 - 2014-02-05 : Viral sinusitis (disorder)
2010-11-22 - 2010-12-10 : Viral sinusitis (disorder)
2007-06-28 - 2007-07-22 : Sinusitis (disorder)
1998-04-22 : Prediabetes
1990-08-29 - : Hypertension
-----
CARE PLANS:
1998-04-22 [CURRENT] : Diabetes self management plan
Reason: Diabetes
Activity: Diabetic diet
Activity: Exercise therapy
-----
OBSERVATIONS:
2016-11-14 : Body Height 157.5 cm
2016-11-14 : Body Weight 104.3 kg
2016-11-14 : Body Mass Index 42.0 kg/m2
2016-11-14 : Systolic Blood Pressure 198.0 mmHg
2016-11-14 : Diastolic Blood Pressure 107.0 mmHg
2016-11-14 : Hemoglobin A1c/Hemoglobin.total in Blood 8.3 %
2016-11-14 : Glucose 133.0 mg/dL
2016-11-14 : Urea Nitrogen 13.0 mg/dL
2016-11-14 : Creatinine 1.0 mg/dL
2016-11-14 : Calcium 9.4 mg/dL
2016-11-14 : Sodium 136.0 mmol/L
2016-11-14 : Potassium 4.5 mmol/L
2016-11-14 : Chloride 102.0 mmol/L
2016-11-14 : Carbon Dioxide 27.0 mmol/L
2016-11-14 : Basic Metabolic Panel
2016-11-14 : Total Cholesterol 243.0 mg/dL
2016-11-14 : Triglycerides 340.0 mg/dL
2016-11-14 : Low Density Lipoprotein Cholesterol 145.0 mg/dL
2016-11-14 : High Density Lipoprotein Cholesterol 30.0 mg/dL
2016-11-14 : Lipid Panel
2016-11-14 : Microalbumin Creatine Ratio 2.0 mg/g
2016-11-14 : Estimated Glomerular Filtration Rate >60 mL/min/(1.73 m2)
-----
PROCEDURES:
2014-11-23 : Documentation of current medications
2011-01-02 : Documentation of current medications
2007-11-19 : Documentation of current medications
-----
ENCOUNTERS:
2016-11-14 : Outpatient Encounter
2015-09-14 : Outpatient Encounter
2015-03-23 : Outpatient Encounter
2014-11-23 : Outpatient Encounter
2014-01-15 : Encounter for Viral sinusitis (disorder)
2011-01-02 : Outpatient Encounter
2010-11-30 : Encounter for Viral sinusitis (disorder)
2007-11-19 : Outpatient Encounter
2007-07-05 : Encounter for Sinusitis (disorder)
```



Synthea: Synthetic Patient Data Generator

- Synthetic Patient Simulation
- High Quality Health Records
- Freely available

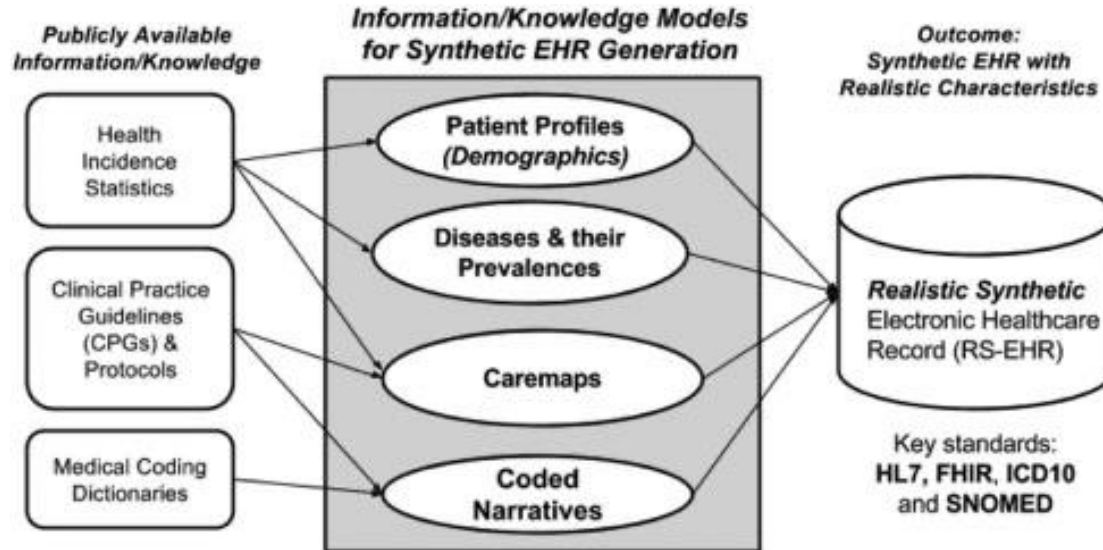
Synthea™ Novel coronavirus (COVID-19) model and synthetic data set

Jason Walonoski ^{a,*}, Sybil Klaus ^a, Eldesia Granger ^a, Dylan Hall ^a, Andrew Gregorowicz ^a,
George Neyarapally ^a, Abigail Watson ^b, Jeff Eastman ^c

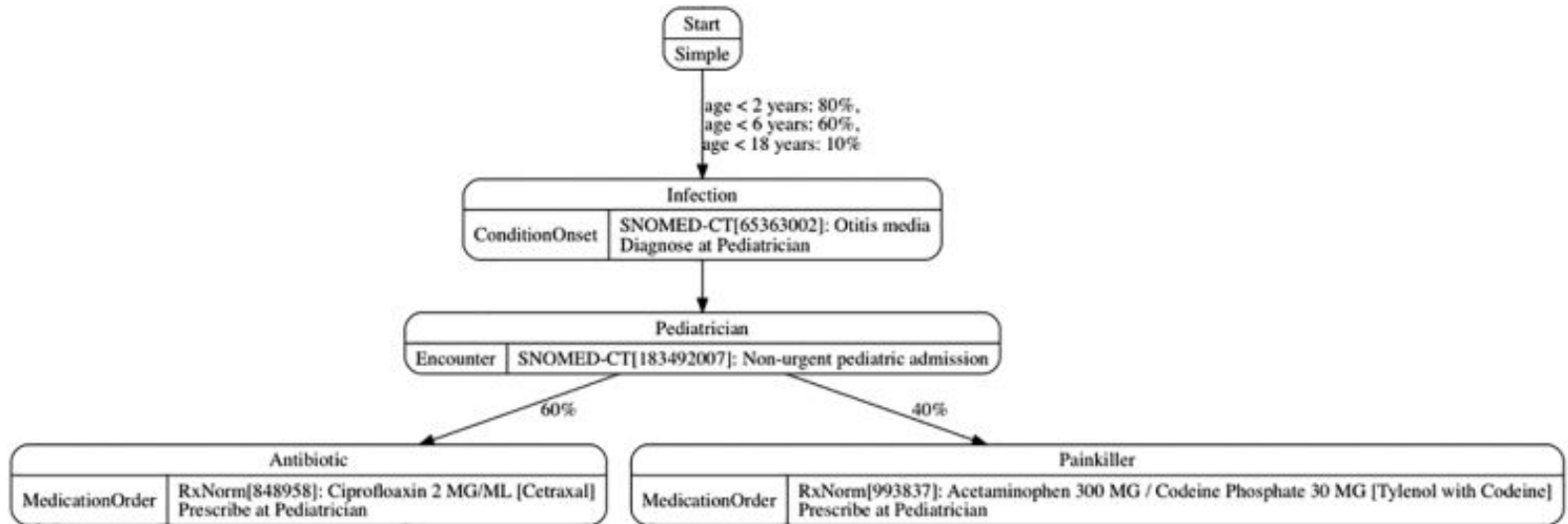
- Later transform to N3C platform (The National COVID Cohort Collaborative)

Synthea: Conceptual framework

- Bases on Publicly Available Data Approach to the Realistic Synthetic EHR, PADARSER



Synthea: A simplified example





Synthea: Partial JSON representation

```
"Infection": {
  "type": "ConditionOnset",
  "target_encounter": "Pediatrician",
  "codes": [ { "system": "SNOMED-CT", "code": "65363002", "display": "Otitis media" } ],
  "direct_transition": "Pediatrician"
},
"Pediatrician": {
  "type": "Encounter",
  "encounter_class": "ambulatory",
  "codes": [ { "system": "SNOMED-CT", "code": "183492007",
    "display": "Non-urgent pediatric admission" } ],
  "distributed_transition": [
    { "distribution": 0.6, "transition": "Antibiotic" },
    { "distribution": 0.4, "transition": "Painkiller" } ]
}
```



Reinforcement Learning

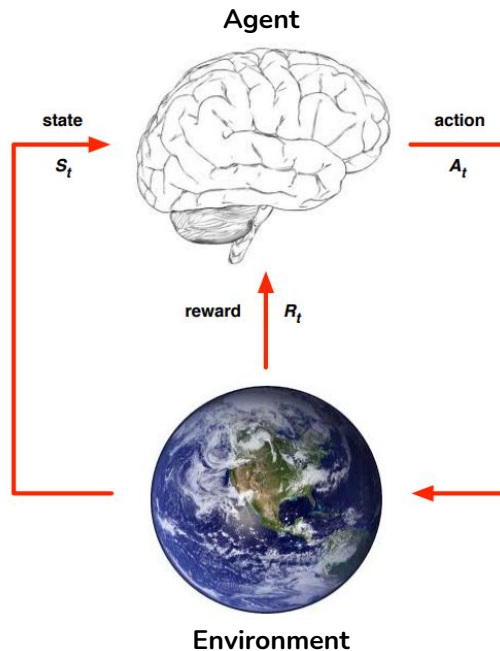
Actions: Decisions that influence the state and “give” a reward

Rewards: Scalar values that the agent maximizes

- “Reward Hypothesis”: All goals can be described by the maximization of expected cumulative reward

States: A summary of our agent at some time t

- More formally, any function of the history of our agent (actions, rewards, observations/states)

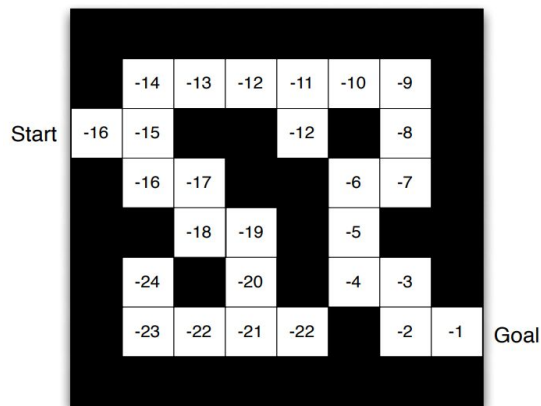




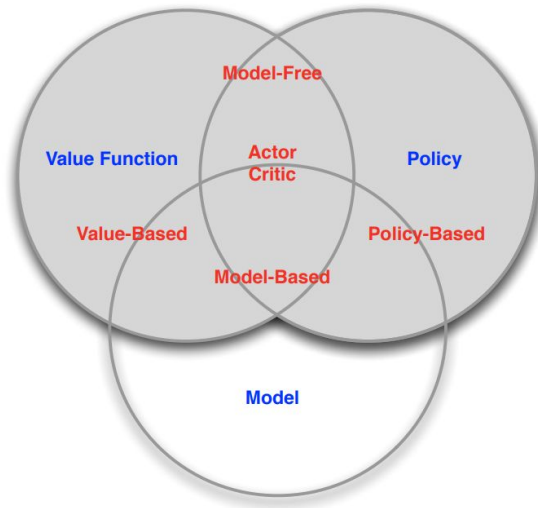
Reinforcement Learning

Types:

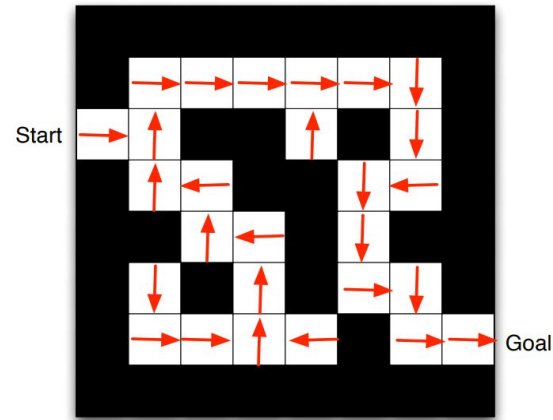
Value-Based



“What is the reward for each state?”



Policy-Based



“Given my state, what is my next action?”



Challenge: Mapping States, Actions, and Rewards

- Actions = clinical interventions and actions intended to alter outcome
- States = measures captured by EHRs, states of patients
- Rewards = longevity, minimize death, maximize health and utility
- **Ultimate question:** Given a particular state, what's the action leading to optimal policy?
 - Given: history of states and actions, self-defined reward
 - Challenge #1: learn the policy from the data
 - Challenge #2: defining metrics of rewards

ML study example takeaways

scientific reports

OPEN Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients

[Check for updates](#)

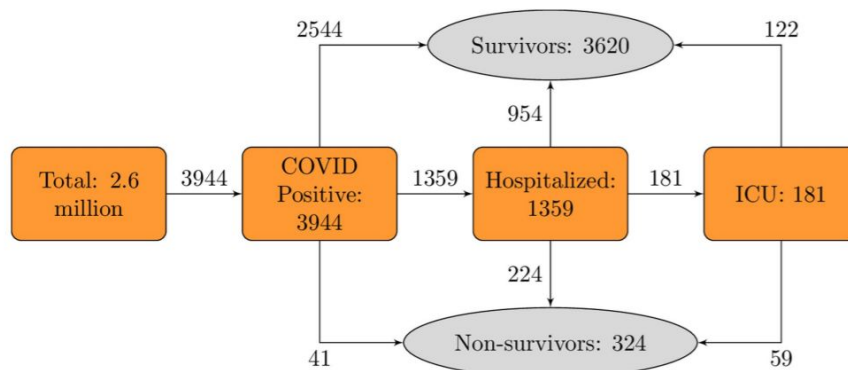


Figure 1. Flowchart of patient selection and identification for the Danish cohort. *ICU* intensive care unit.



Division of Initial Tasks and Roles

- **Andrew:** Organizing literature, entry point for N3C data querying, Synthea exploration
- **Nicky:** EHR data, entry point for N3C data querying, Synthea exploration
- **Devon & Matthew:** RL deep-dive and specific model exploration, codesets and coding leads
- **Shirley:** Synthea exploration, mathematical understanding and background for models

Roughly our next few weeks



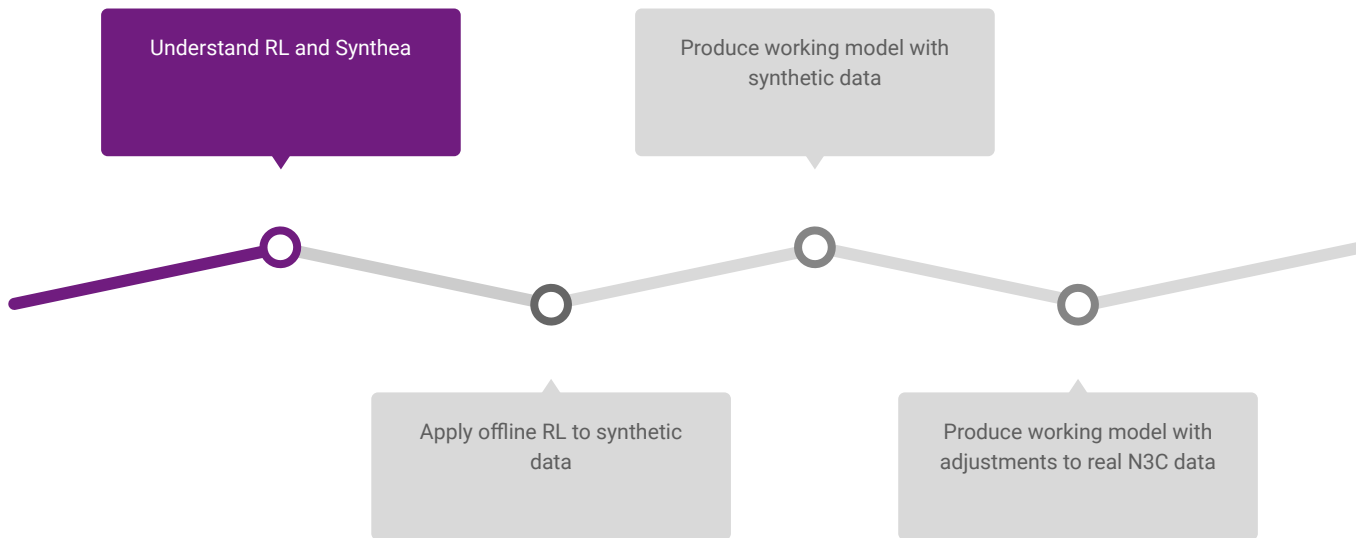
Division of Tasks and Roles - mid to final

- **Andrew:** RL code running
- **Nicky:** RL code running
- **Devon & Matthew:** RL code development
- **Shirley:** RL algorithm

Aiming for start of March -- 2nd step on next slide



Next Steps





Questions?