Synthetic Health Data Challenge Winning Solutions Webinar

Stephanie Garcia, MPH | ONC PCOR Portfolio Manager

October 19, 2021

The Office of the National Coordinator for Health Information Technology

The Office of the National Coordinator for Health Information Technology

Office of the National Coordinator for Health Information Technology (ONC)

- **Mission**
  - Improve the health and well-being of individuals and communities through the use of technology and health information that is accessible when and where it matters most
- **Strategic Goals**
  - Advance Person-Centered and Self-Managed Health
  - Transform Health Care Delivery and Community Health
  - **Foster Research, Scientific Knowledge, and Innovation**
  - Enhance Nation's Health IT Infrastructure

https://healthit.gov/

Patient-Centered Outcomes Research (PCOR)

- Produce evidence to inform health care decisions made by patients, families, and their health care providers
- Patient-Centered Outcomes Research Trust Fund (PCORTF) managed by the Assistant Secretary for Planning and Evaluation (ASPE)

https://www.healthit.gov/topic/scientific-initiatives/building-data-infrastructure-support-patient-centered-outcomes-research

ONC Synthetic Health Data Project

Accelerate ability to conduct PCOR by:

- Enhancing an open-source synthetic data generator Synthea ™, developed by The MITRE Corporation, to increase the number and variety of synthetic data
  - Opioid use
  - Pediatric populations
  - Patients with complex care needs
- Engaging broader community to validate the realism and demonstrate the potential uses of newly available synthetic data

https://www.healthit.gov/topic/research-evaluation/synthetic-health-data-generation-to accelerate-patient-centered-outcomes

Synthetic Health Data Challenge

Prize competition invited a wide array of innovators, researchers, and technology developers to **create and test innovative solutions** that enhance Synthea and the synthetic data it generates

- Advance novel uses of synthetic data for patient-centered outcomes research
- Validate the realism of Synthea-generated synthetic data

Challenge Structure

Phase I: Proposals for Innovative Models

- Written proposal describing proposed solution
- Proposals were invited from teams or individuals
- Proposals had to include methodology and intended outcomes

Phase II: Prototype/ Solution Development

- Approved proposals moved on to Phase II
- Solutions designed and tested
- Final paper describing the solution
- Video demonstration
- Evidence of validation
- Non-proprietary source code

Winning Solutions

- Total cash prize pool: $100,000
- Solutions were judged by a panel based on criteria

Winning Solutions

FIRST PLACE

$40,000

CodeRx

Medication Diversification Tool

SECOND PLACE

$15,000

The Generalistas

Virtual Generalist: Modeling Co-morbidities in Synthea

Team LMI

On Improving Realism of Disease Modules in Synthea: Social Determinant-Based Enhancements to Conditional Transition Logic

THIRD PLACE

$10,000

Particle Health	The Necessity of Realistic Synthetic Health Data Development Environments
Team TeMa	Empirical Inference of Underlying Condition Probabilities Using Synthea-Generated Synthetic Health Data
UI Health	Spatiotemporal Big Data Analysis of Opioid Epidemic in Illinois

Medication Diversification Tool (MDT)

Team CodeRx

Challenge Category and Use Case

Challenge Category 1	Enhancement to Synthea
Use Case	Pediatrics (pediatric asthma)

Source: https://www.istockphoto.com/vector/asthmatic-girl-gm185960170-27680137

Pharmacists who code and developers who (health)care

Who we are

CodeRx is a collective of pharmacists and other healthcare professionals who have a skill set in tech and apply it towards building useful things

Website: coderx.io

Founded: early 2020

Membership: 150+ (30-40 active weekly), mostly pharmacists and pharmacy students

What we do

- **Slack channel** - engage in discussions about coding / data / tech as it relates to pharmacy and healthcare
- **GitHub organization** - collaborate on open source pharm tech / health tech projects
- **Newsletter and website** - share guides, resources, and pertinent topics

Team CodeRx

For this challenge, Team CodeRx consists of six PharmDs from across the U.S.

Joseph LeGrand, **PharmD**, MS *(team lead)*	Lead Application Analyst Vanderbilt University Medical Center
Kent Bridgeman, **PharmD**, MHI	Informatics Pharmacist Allina Health
Kristen Tokunaga, **PharmD**, BCGP	Analytics Consulting Manager Komodo Health
Yevgeny Bulochnik, **PharmD**, ACE, CACP	Formulary Administration Consultant HealthPartners
Robert Hodges, **PharmD**, MSDS, MBA	Sr. Data Scientist, McKesson RelayHealth (CoverMyMeds)
Dalton Fabian, **PharmD**	Data Scientist UnityPoint Health

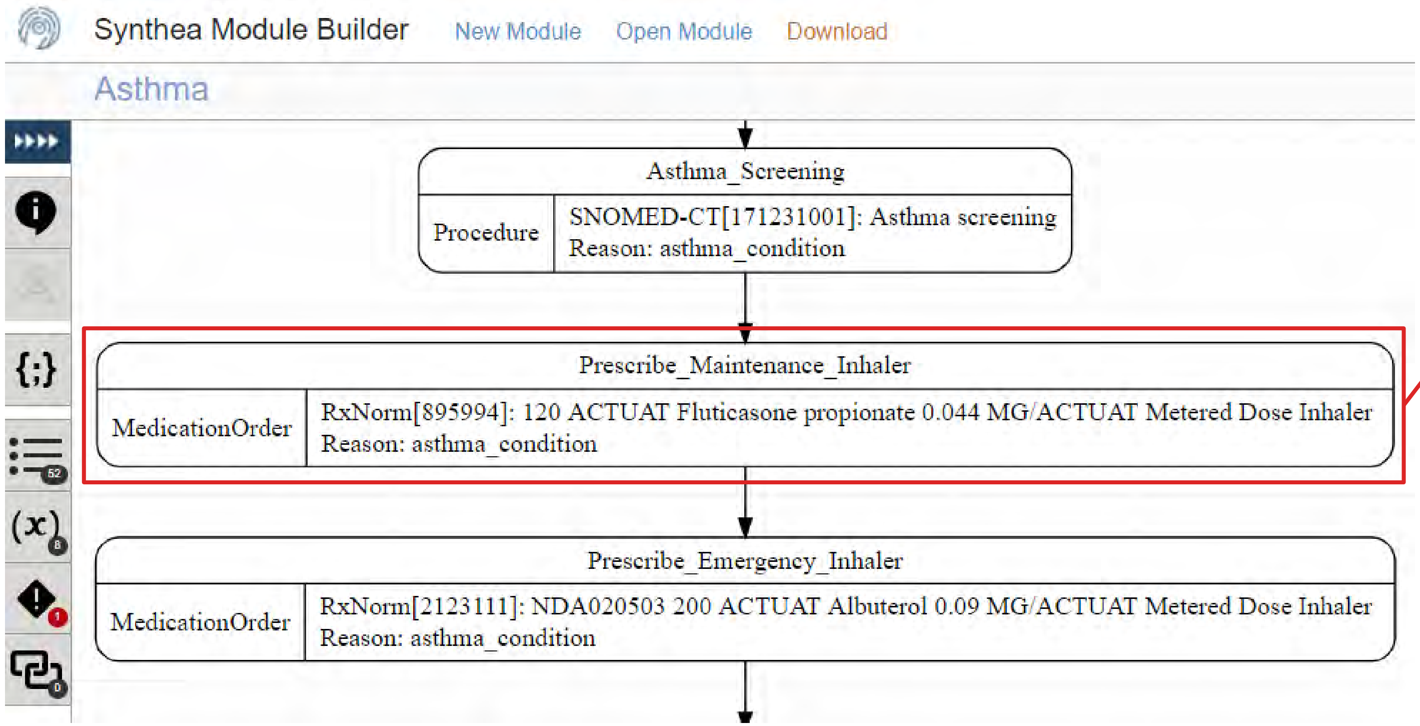
Objective

To programmatically generate Synthea medication orders through the use of RxClass and Medical Expenditure Panel Survey (MEPS) data sources

Problem Statement

In the current Synthea asthma module, 100% of Asthma patients are prescribed the same asthma maintenance inhaler (Flovent HFA 44 mcg) regardless of age.

This is not consistent with real-world clinical practice.



Source:
https://synthetichealth.github.io/module-builder/#asthma

Asthma Maintenance Inhalers on the Market

Pediatric patients are prescribed one of **many** inhaled corticosteroids as a first line therapy.

Metered Dose Inhalers
+
Dry Powder Inhalers

Single ingredient inhaled corticosteroids (ICS)

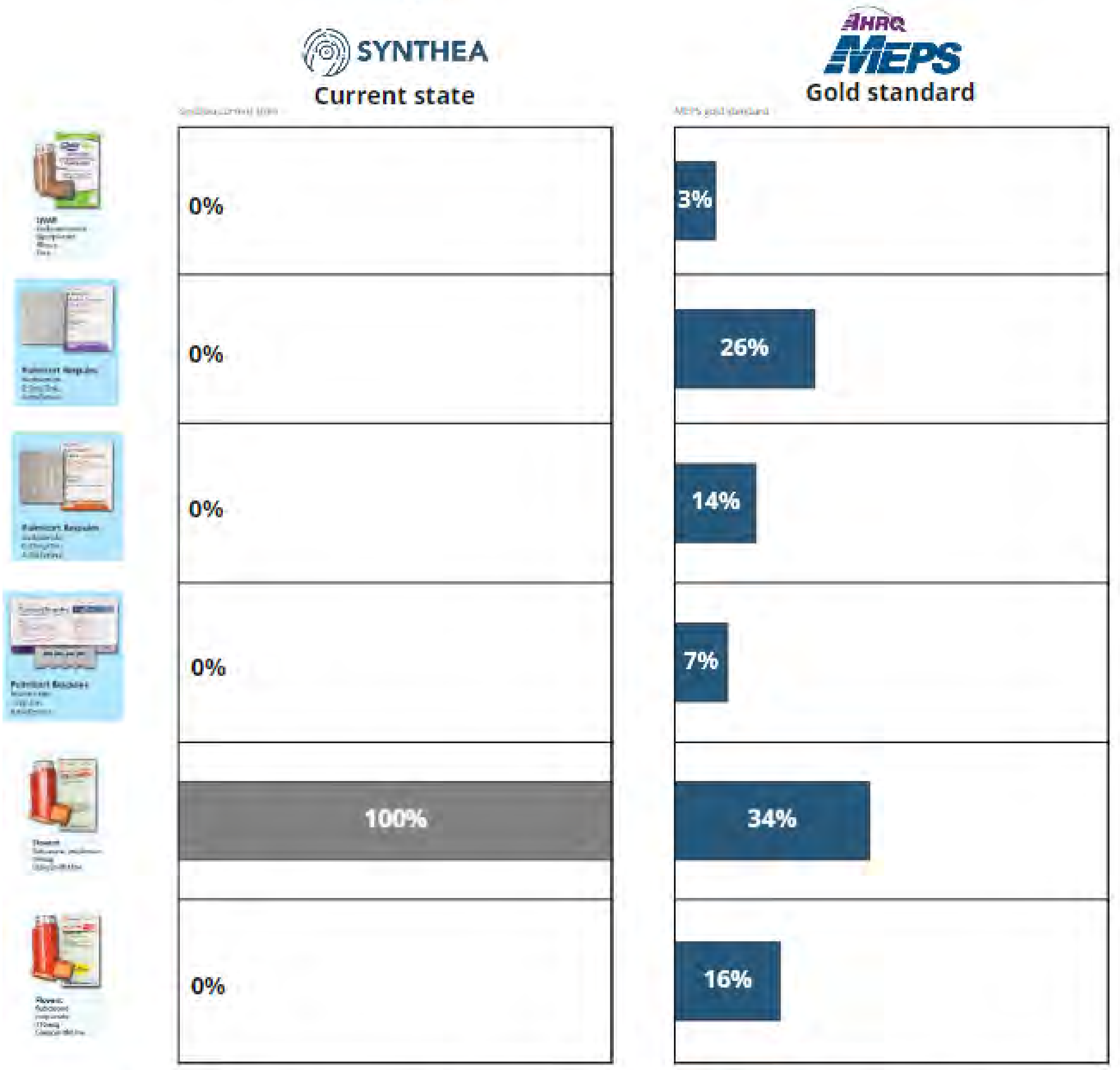
Inhalation Suspensions (Nebulized)

Source: http://www.med.umich.edu/pdf/clinical-guidelines/Asthma-medications-poster.pdf

Synthea vs MEPS

Medical Expenditure Panel Survey (MEPS) is a nationwide set of surveys of households and their medical providers.

The AHRQ conducts this survey annually to collect information on the use and cost of health care.



Percent of patient population age 0-5 years old with prescribed medication product

Tools

Open-source tools & code

Packages

- Python
- Pandas

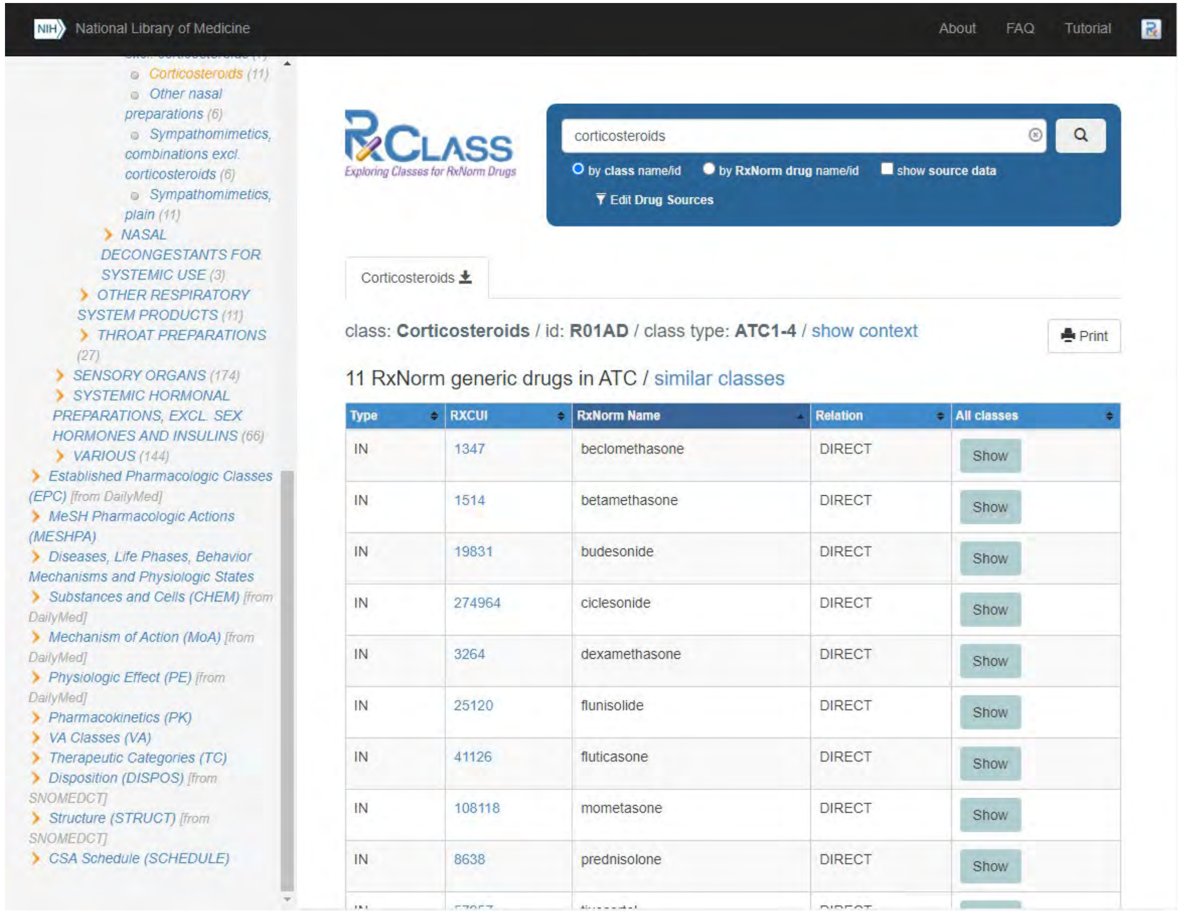
Publicly-available, government-maintained data sources

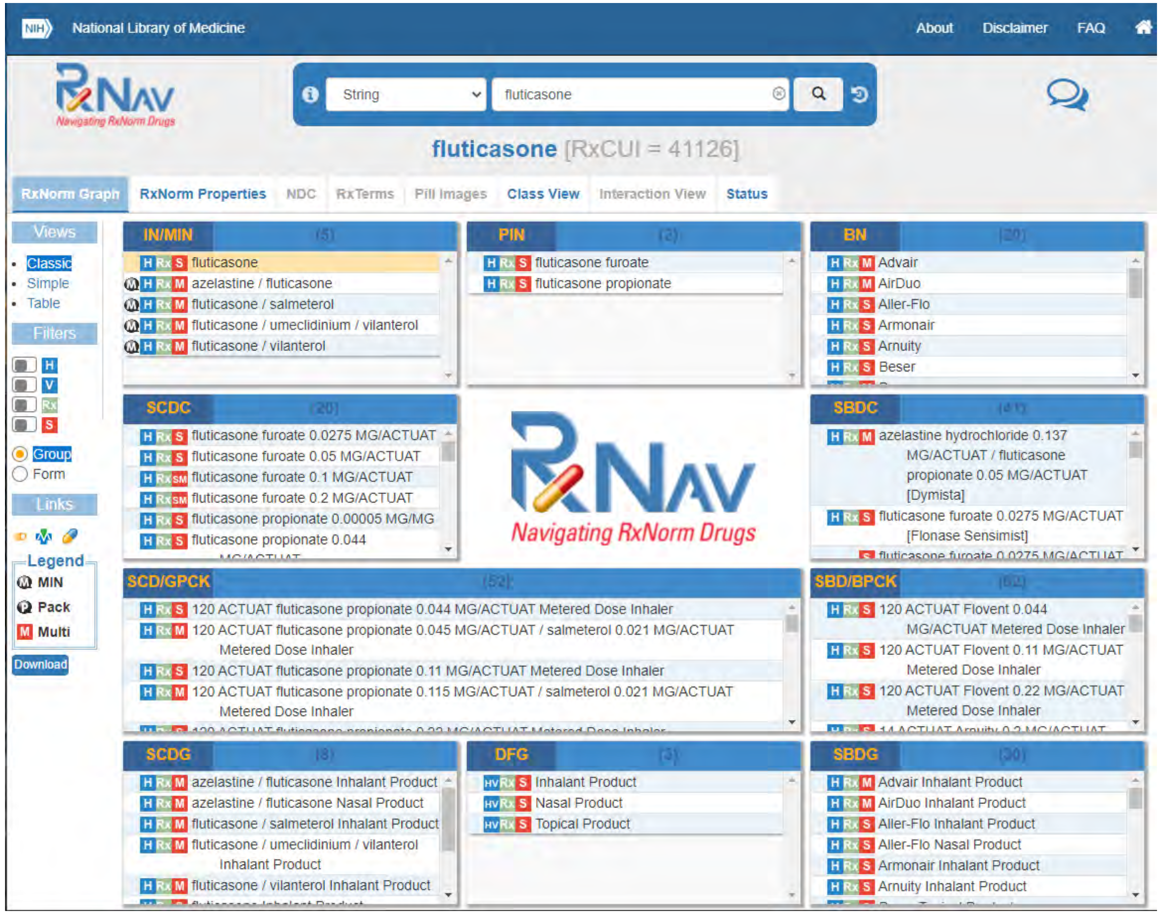
Sources

- RxNav API
- RxClass API
- MEPS

Available Online GUIs

Already exist - maintained and hosted by National Library of Medicine





Developer Inputs

Required

1. RxClass class ID(s) OR RxNorm ingredient ID(s)

Optional

1. Dose form filters
2. Patient demographic info breakpoints
   a. Age range(s)
   b. Gender M/F
   c. State of residence
3. Single vs multi ingredient drugs
4. Other Synthea settings

Methods

RxClass → (RxNorm) RxNav → MEPS

Medication class -> ingredient(s)

Medication ingredient -> product(s) -> NDC(s)

+ dose form filters

+ single vs multi ingredient filters

NDCs and demographic info and counts of patients who report taking them

Use RxClass API to return list of medication ingredients

- "Corticosteroids" = ATC Class **R01AD**

Medication ingredients

beclomethasone
betamethasone
budesonide
ciclesonide
flunisolide
fluticasone
mometasone
prednisolone
tixocortol
triamcinolone

NIH NLM

RxNav

Navigating RxNorm Drugs

Filter on RxNorm dose form in:

- Metered Dose Inhaler
- Dry Powder Inhaler
- Inhalation Suspension

Single ingredient products only

- RxNorm term type (TTY) = IN

Medication ingredients

beclomethasone	YES
~~betamethasone~~	~~NO~~
budesonide	YES
ciclesonide	YES
flunisolide	YES
fluticasone	YES
mometasone	YES
~~prednisolone~~	~~NO~~
~~tixocortol~~	~~NO~~
~~triamcinolone~~	~~NO~~

AHRQ

MEPS

Medical Expenditure Panel Survey

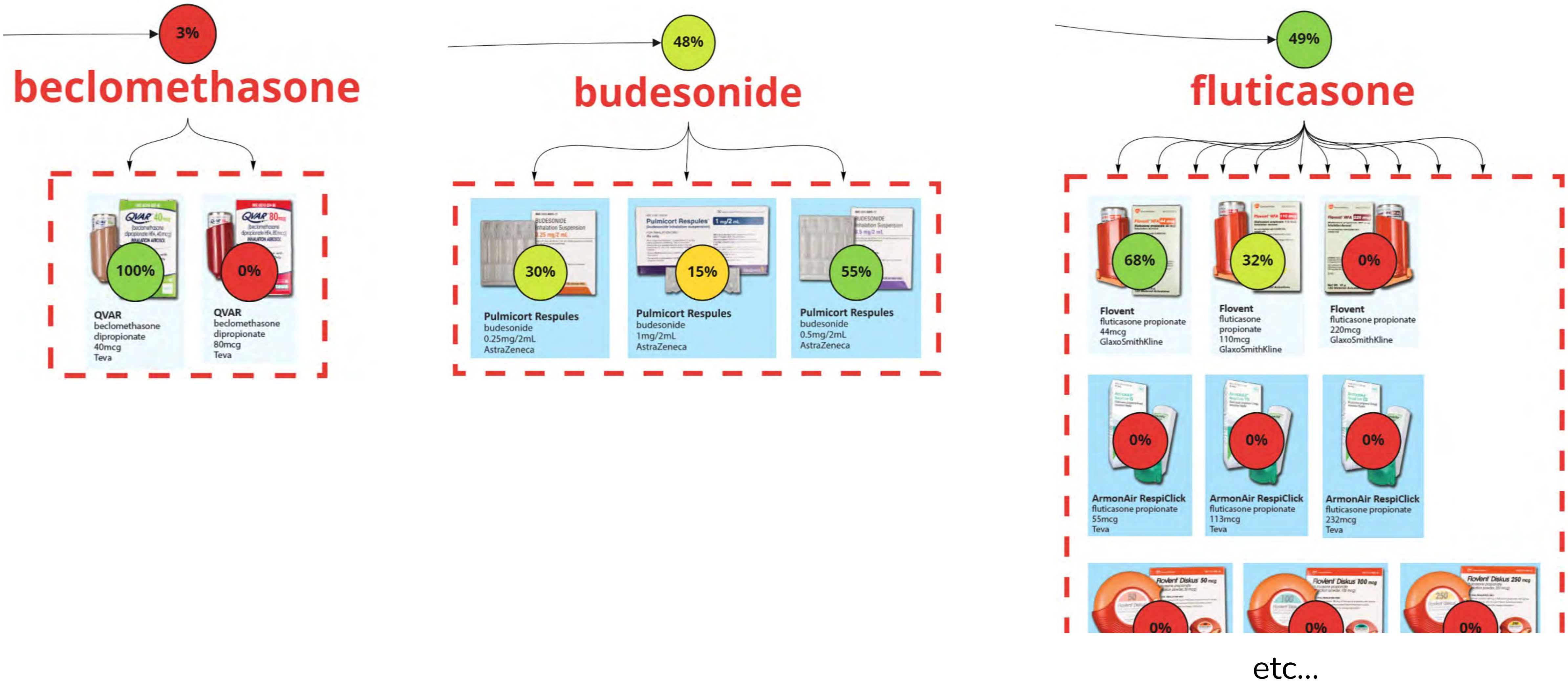
Cross-reference population prescription utilization data

- 0-5 year old age range

NOTE: for the actual submodule, we would also include the 6-103 year old age range

Medication ingredients

beclomethasone	3%
betamethasone	N/A
budesonide	48%
~~ciclesonide~~	~~0%~~
~~flunisolide~~	~~0%~~
fluticasone	49%
~~mometasone~~	~~0%~~
prednisolone	N/A
tixocortol	N/A
triamcinolone	N/A



3%
beclomethasone
100%
QVAR
beclomethasone
dipropionate
40mcg
Teva
0%
QVAR
beclomethasone
dipropionate
80mcg
Teva
48%
budesonide
30%
Pulmicort Respules
budesonide
0.25mg/2mL
AstraZeneca
15%
Pulmicort Respules
budesonide
1mg/2mL
AstraZeneca
55%
Pulmicort Respules
budesonide
0.5mg/2mL
AstraZeneca
49%
fluticasone
68%
Flovent
fluticasone propionate
44mcg
GlaxoSmithKline
32%
Flovent
fluticasone
propionate
110mcg
GlaxoSmithKline
0%
Flovent
fluticasone propionate
220mcg
GlaxoSmithKline
0%
ArmonAir RespiClick
fluticasone propionate
55mcg
Teva
0%
ArmonAir RespiClick
fluticasone propionate
113mcg
Teva
0%
ArmonAir RespiClick
fluticasone propionate
232mcg
Teva
0%
0%
0%
etc...

```
synthea/
├ src/
│  ├ main/
│  │  ├ resources/
│  │  │  ├ modules/
│  │  │  │  ├ medication/
│  │  │  │  │  ├ maintenance_inhaler.json
│  │  │  │  │  ├ ...
│  │  │  │  ├ lookup_tables/
│  │  │  │  │  ├ maintenance_inhaler_ingredient_distribution.csv
│  │  │  │  │  ├ maintenance_inhaler_fluticasone_product_distribution.csv
│  │  │  │  │  ├ maintenance_inhaler_budesonide_product_distribution.csv
│  │  │  │  │  ├ maintenance_inhaler_beclomethasone_product_distribution.csv
│  │  │  │  │  ├ maintenance_inhaler_mometasone_product_distribution.csv
│  │  │  │  │  ├ ...
│  │  │  │  ├ asthma.json
│  │  │  │  ├ ...
```

Asthma module JSON:

```
...
    "Maintenance_Medication_End": {
      "type": "MedicationEnd",
      "referenced_by_attribute": "maintenance_inhaler",
      "direct_transition": "Emergency_Medication_End"
    },
...
```

SYNTHEA

1. Place MDT module JSON file in the medications folder
2. Place MDT lookup table CSV files in the lookup_tables folder
3. Replace existing **MedicationOrder** state in asthma module JSON file with a **CallSubmodule** state referencing the MDT module
4. Ensure asthma module **MedicationEnd** states end medications by <u>attribute</u>, not by <u>name</u>

MDT Demo

Results

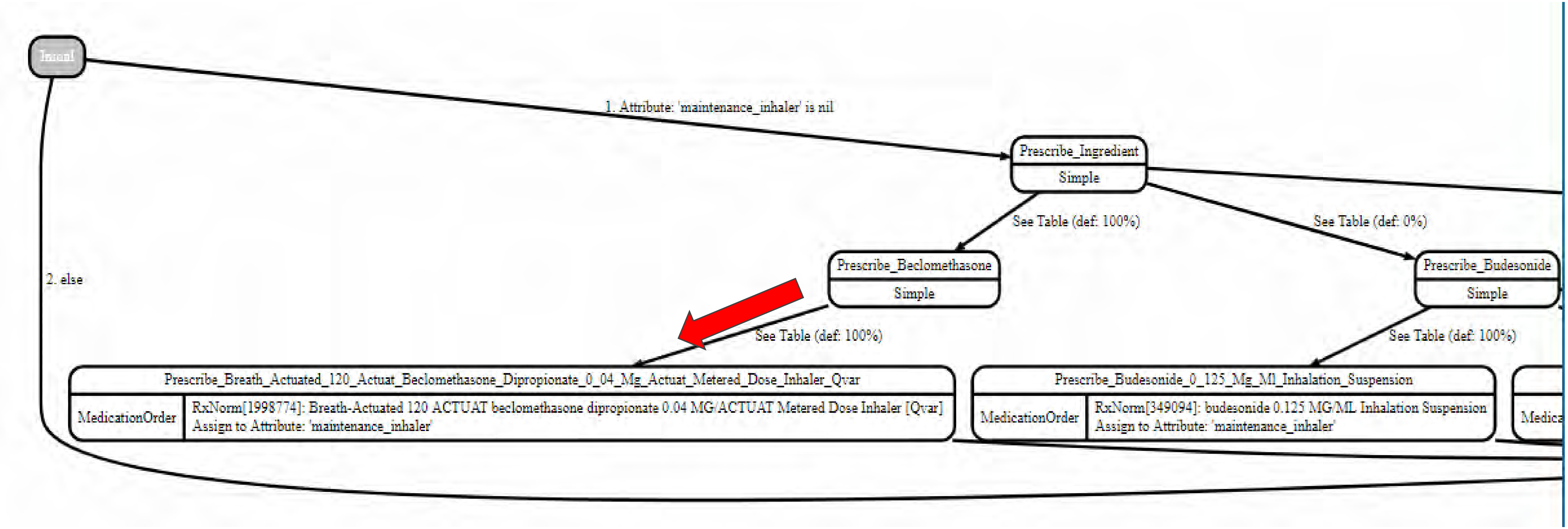
Asthma Medications using the MDT - Ingredient

```
    "==========================================================================",
    " MEDICATION INGREDIENT TABLE TRANSITION                                   ",
    "==========================================================================",
    "Ingredients in lookup table:",
    "#  [ % pop ] Name",
    "-- --------- ----",
    "1. [ 3.1% ] Beclomethasone",
    "2. [ 47.7% ] Budesonide",
    "3. [ 49.2% ] Fluticasone"
],
"type": "Simple",
"lookup_table_transition": [
    {
        "transition": "Prescribe_Beclomethasone",
        "default_probability": "1",
        "lookup_table_name": "maintenance_inhaler_ingredient_distribution.csv"
    },
    {
        "transition": "Prescribe_Budesonide",
```

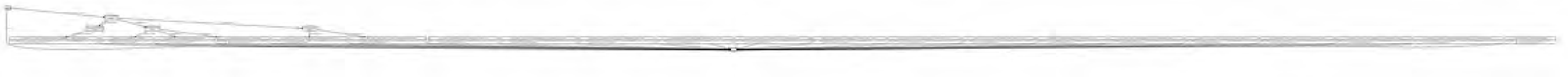
Asthma Medications using the MDT - Product

```
    "==================================================================",
    " BUDESONIDE MEDICATION PRODUCT TABLE TRANSITION  ",
    "==================================================================",
    "Products in lookup table:",
    "#  [ % pop ] Name",
    "-- --------- ----",
    "1. [ 29.5% ] Budesonide_0_125_Mg_Ml_Inhalation_Suspension",
    "2. [ 15.1% ] Budesonide_0_125_Mg_Ml_Inhalation_Suspension_Pulmicort",
    "3. [ 55.3% ] Budesonide_0_25_Mg_Ml_Inhalation_Suspension"
],
"type": "Simple",
"lookup_table_transition": [
    {
        "transition": "Prescribe_Budesonide_0_125_Mg_Ml_Inhalation_Suspension",
        "default_probability": "1",
        "lookup_table_name": "maintenance_inhaler_Budesonide_product_distribution.csv"
    },
    {
        "transition": "Prescribe_Budesonide_0_25_Mg_Ml_Inhalation_Suspension",
        "default_probability": "0",
        "lookup_table_name": "maintenance_inhaler_Budesonide_product_distribution.csv"
    },
    {
        "transition": "Prescribe_Budesonide_0_125_Mg_Ml_Inhalation_Suspension_Pulmicort",
        "default_probability": "0",
        "lookup_table_name": "maintenance_inhaler_Budesonide_product_distribution.csv"
    }
]
```

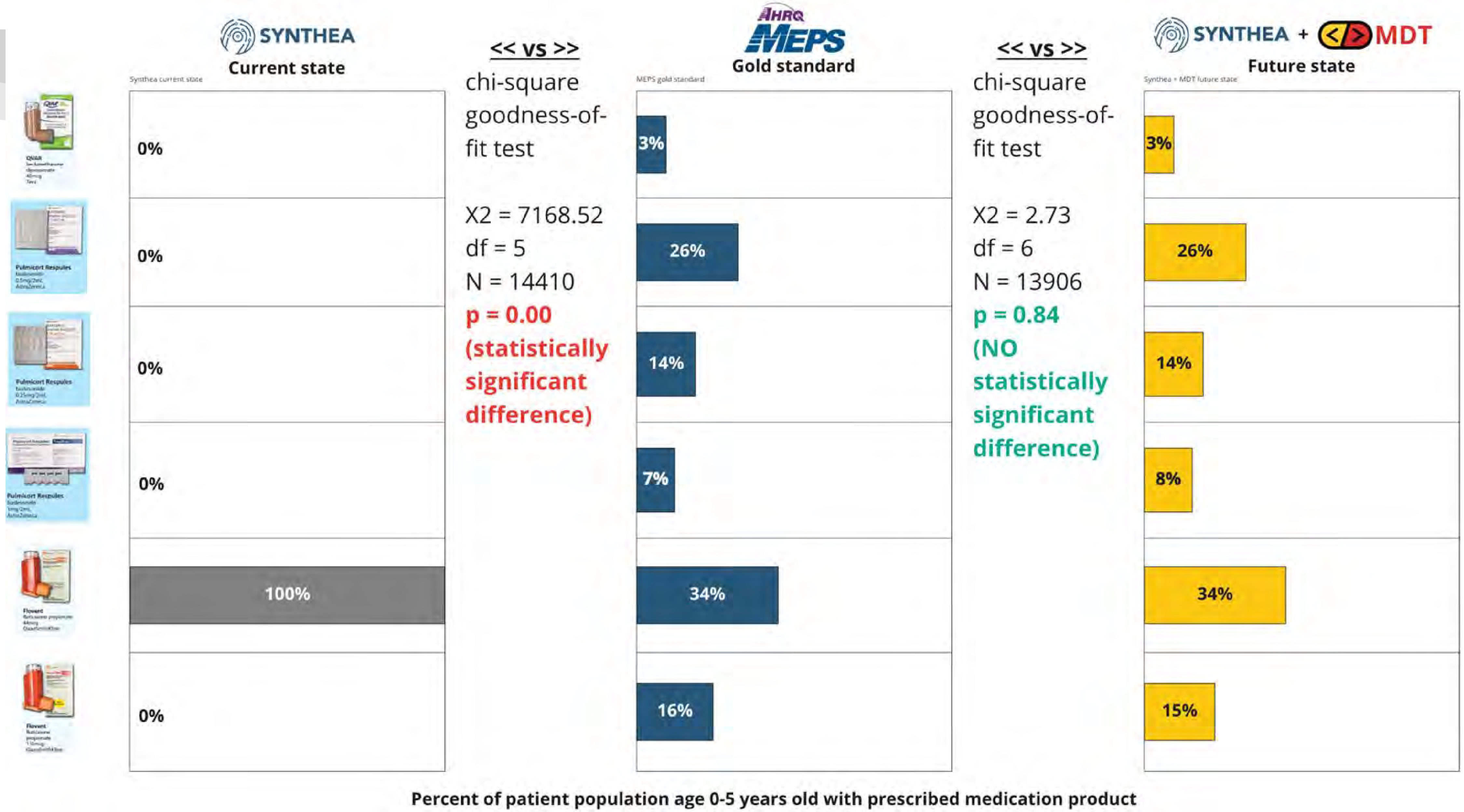
Asthma Medications In Synthea Module Builder



Asthma Medications In Synthea Module Builder



Validation



SYNTHEA
Current state
Synthea current state
<< vs >>
chi-square goodness-of-fit test
X2 = 7168.52
df = 5
N = 14410
p = 0.00
(statistically significant difference)
AHRQ MEPS
Gold standard
MEPS gold standard
<< vs >>
chi-square goodness-of-fit test
X2 = 2.73
df = 6
N = 13906
p = 0.84
(NO statistically significant difference)
SYNTHEA + MDT
Future state
Synthea + MDT future state
QVAR
Pulmicort Respules
Pulmicort Respules
Pulmicort Respules
Flovent
Flovent
0%
0%
0%
0%
100%
0%
3%
26%
14%
7%
34%
16%
3%
26%
14%
8%
34%
15%
Percent of patient population age 0-5 years old with prescribed medication product

Benefits for Researchers

- Creates micro-validated medication distributions in Synthea modules
- Loosely test hypothesis prior to obtaining access to PHI
- Identify drug trends to validate on real data

https://www.colourbox.com/vector/word-cloud-software-engineering-vector-28500240

Benefits for developers

- Encourages incorporation of complex drug treatment options into Synthea modules
- Offers complexity that developers will need to account for in software
- Allows drug related development before having PHI

Challenges & Successes

Challenges

- Complexity of data
- Difficulty in defining disease-specific drug lists

Successes

- Can be used with other Synthea modules & disease states
- Can generate medication distributions by patient age, gender, geographic location, and more
- Can be re-run with updated data
- Has flexibility for user to select drugs
- Designed by 6 Pharmacists who are experts on medication use

Future for MDT

Future Enhancements

The groundbreaking methods used in MDT to integrate public data sources allows for future enhancements:

1. Enhance Synthea with other MEPS data elements.
    a. insurance type, social determinants of health, medical conditions.
2. Model prescription fill and medication adherence through meps MEPS data.
3. Add allergy or drug-drug interactions logics through NLM[1] data.
4. Capture dose ranges (low/medium/high) and progression.
5. Improve medication costs modeling through NADAC[2] dataset.

[1]NLM [National Library of Medicine]
[2]NADAC [National Average Drug Acquisition Cost]

Thank You!!!!

Links

GitHub Repo: github.com/coderxio/medication-diversification

Joseph LeGrand, PharmD, MS linkedin.com/in/jrlegrand/

Kent Bridgeman, PharmD, MHI linkedin.com/in/kentvbridgeman/

Kristen Tokunaga, PharmD, BCGP linkedin.com/in/kristen-tokunaga-pharmd/

Robert Hodges, PharmD, MSDS, MBA linkedin.com/in/robhodgespharmd/

Dalton Fabian, PharmD linkedin.com/in/daltonfabian/

Yevgeny (Eugene) Bulochnik, PharmD ACE CACP linkedin.com/in/yevgeny-eugene-bulochnik-b429a6155/

CodeRx Website: coderx.io/

Virtual Generalist

MODELING CO-MORBIDITIES IN SYNTHEA™

Robert Horton, PhD
John-Mark Agosta, PhD

Jason Dausman, MD
Brandon DeShon
Benjamin Dummitt, PhD

Katherine Gundling, MD

SUSTAINABLE HARVEST INTERNATIONAL

github.com/rmhorton/virtual-generalist

www.sustainableharvest.org

Outline

Bayes nets & CPTs

- comorbidities matter!
- use in Synthea's new lookup_table_transition

Mapping concepts

- ICD10
- SNOMED
- Synthea attributes

Feature engineering

- SQL/Pyspark/R on Databricks
- Now demonstrated with Synthea data!

Validation

- co-occurrence matrix
- COPM

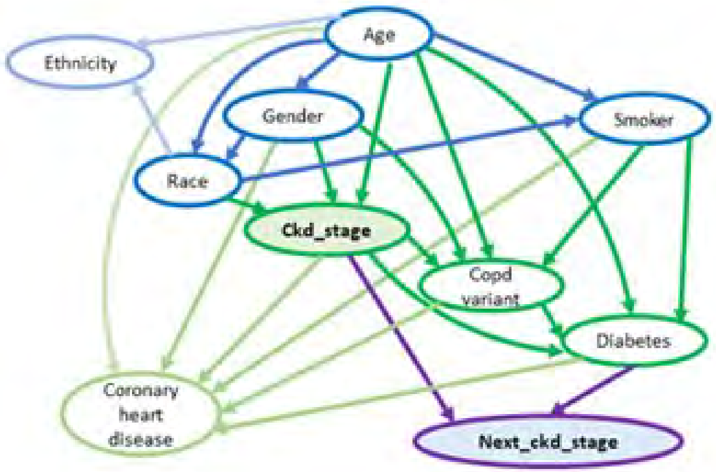
Suggested design patterns

- ***No delay:*** disease is NOT your destiny
- ***Mechanistic progression:*** allow interventions
- ***More modular modules***

  Separate treatment from disease incidence and progression

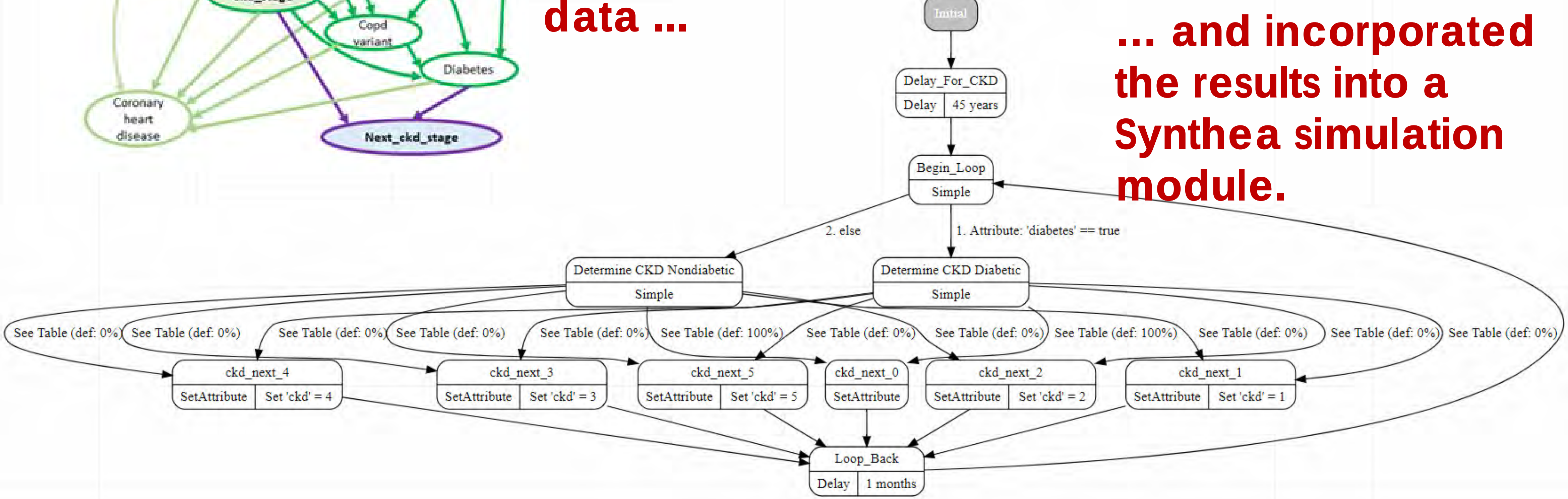
  Use attributes to communicate between modules.

Bayes Nets decompose joint probabilities into a network of conditional probabilities



We trained a Bayes Net model on real hospital data ...

... and incorporated the results into a Synthea simulation module.



Each node in the Bayes Net contains a Conditional Probability Table

```
%r
# patients with diabetes
fit$next_ckd_stage$prob[,'T',] %>% as.matrix %>% t %>% format(digits=3, scientific=FALSE)
```

```
         next_ckd_stage
ckd_stage ckd_0      ckd_1      ckd_2      ckd_3      ckd_4      ckd_5
    ckd_0 "0.986301" "0.000363" "0.001291" "0.010941" "0.000720" "0.000384"
    ckd_1 "0.000000" "0.956890" "0.008477" "0.031727" "0.002180" "0.000727"
    ckd_2 "0.000000" "0.001905" "0.938423" "0.055385" "0.002926" "0.001361"
    ckd_3 "0.000000" "0.000577" "0.004083" "0.977222" "0.014840" "0.003279"
    ckd_4 "0.000000" "0.000432" "0.001105" "0.075977" "0.878830" "0.043656"
    ckd_5 "0.000000" "0.000115" "0.000804" "0.011950" "0.019496" "0.967634"
```

```
%r
# patients without diabetes
fit$next_ckd_stage$prob[,'F',] %>% as.matrix %>% t %>% format(digits=3, scientific=FALSE)
```

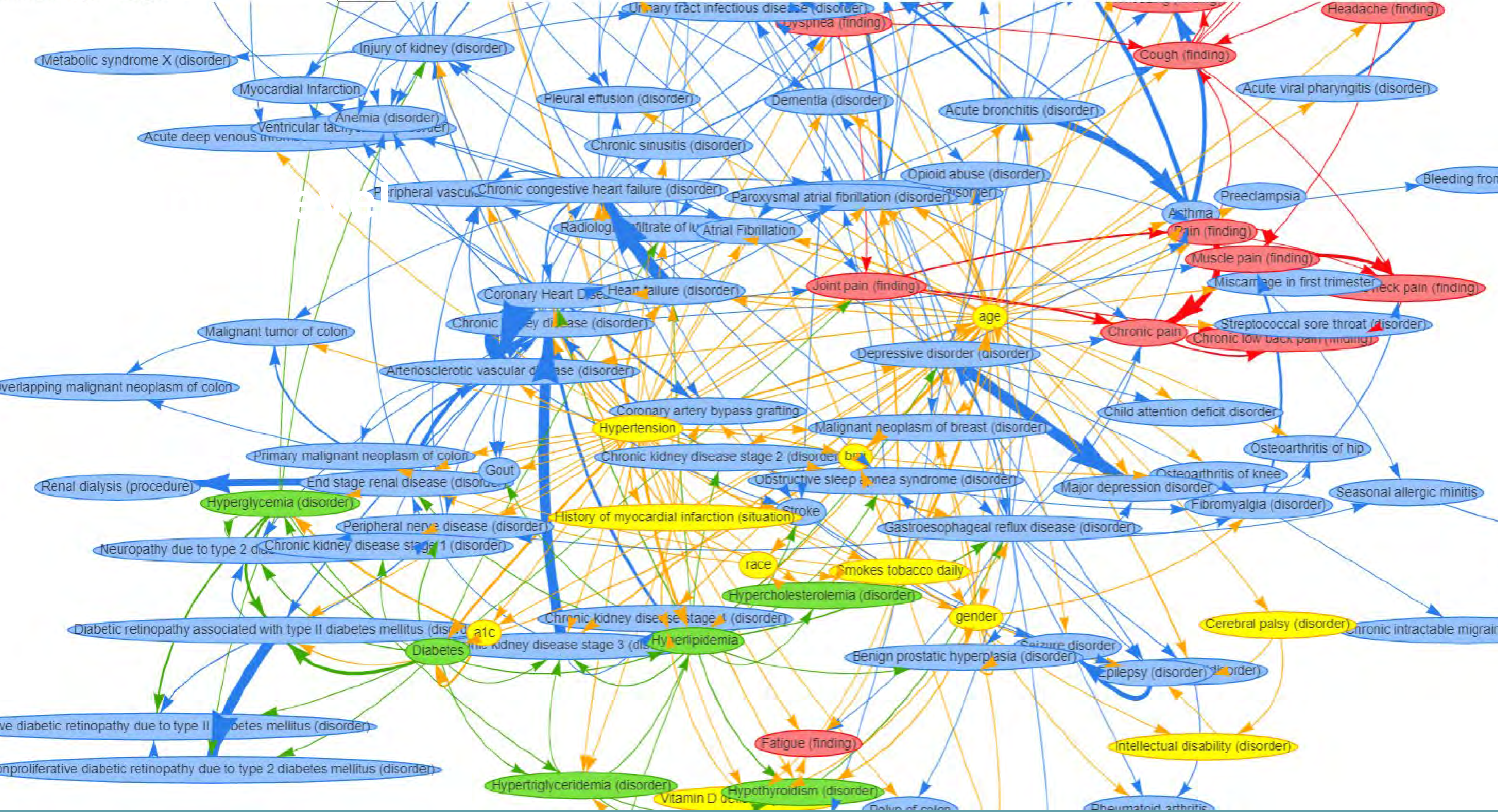
```
         next_ckd_stage
ckd_stage ckd_0       ckd_1       ckd_2       ckd_3       ckd_4       ckd_5
    ckd_0 "0.9922398" "0.0001239" "0.0007213" "0.0063517" "0.0003788" "0.0001846"
    ckd_1 "0.0000000" "0.9690204" "0.0053652" "0.0228453" "0.0020768" "0.0006923"
    ckd_2 "0.0000000" "0.0007380" "0.9623602" "0.0341439" "0.0018645" "0.0008934"
    ckd_3 "0.0000000" "0.0002917" "0.0025664" "0.9877009" "0.0078259" "0.0016151"
    ckd_4 "0.0000000" "0.0001632" "0.0011015" "0.0518951" "0.9211374" "0.0257027"
    ckd_5 "0.0000000" "0.0000676" "0.0005068" "0.0089204" "0.0116236" "0.9788816"
```

Mapping concepts

ICD10 → SNOMED → attributes

	attribute	test	attribute_type	icd_pattern	icd_name	snomed_concept_name
1	ckd_1	==1	integer	N18.1	Chronic_kidney_disease_stage_1	Chronic kidney disease stage 1 (disorder)
2	ckd_2	==2	integer	N18.2	Chronic_kidney_disease_stage_2	Chronic kidney disease stage 2 (disorder)
3	ckd_3	==3	integer	N18.3	Chronic_kidney_disease_stage_3	Chronic kidney disease stage 3 (disorder)
4	ckd_4	==4	integer	N18.4	Chronic_kidney_disease_stage_4	Chronic kidney disease stage 4 (disorder)
5	ckd_5	==5	integer	N18.[56]	Chronic_kidney_disease_stage_5	End stage renal disease (disorder)
6	smoker	is true	boolean	F17	Nicotine_dependence	Smokes tobacco daily
7	diabetes	is true	boolean	E11	Type_2_diabetes_mellitus	Diabetes
8	coronary_heart_disease	is true	boolean	I25	Chronic_ischemic_heart_disease	Coronary Heart Disease
9	copd_variant	is not nil	ConditionOnset	J44	Chronic_obstructive_pulmonary_disease	Chronic obstructive bronchitis (disorder)

inimum edge strength: 0



Synthea-Specific ICD10 to SNOMED Mapping

icd_set_id	snomed_concept_id	snomed_concept_name	icd_code	icd_description
43	44054006	Diabetes	E11	Type 2 diabetes mellitus
446	46177005	End stage renal disease (disorder)	N18.5	Chronic kidney disease, stage 5
447	46177005	End stage renal disease (disorder)	N18.6	End stage renal disease
899	75498004	Acute bacterial sinusitis (disorder)	B96.89	Other specified bacterial agents as the cause of diseases classified elsewhere
899	75498004	Acute bacterial sinusitis (disorder)	J01.90	Acute sinusitis, unspecified
1323	301011002	Escherichia coli urinary tract infection	B96.20	Unspecified Escherichia coli [E. coli] as the cause of diseases classified elsewhere
1323	301011002	Escherichia coli urinary tract infection	N39.0	Urinary tract infection, site not specified

Patterns

E11 matches E11.0, E11.1, etc

Sets

All ICD elements of the set must be present for the SNOMED concept to apply

Feature engineering

Now with Synthea data!

	patient	month_number	age_group	gender	race	ethnicity	ckd_stage	next_				smoker
726	00037be2-d64b-adb8-c7e7-90dcffa144bf	515	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	[illegible]	[illegible]	F	F
727	00037be2-d64b-adb8-c7e7-90dcffa144bf	516	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	[illegible]	[illegible]	F	F
728	00037be2-d64b-adb8-c7e7-90dcffa144bf	517	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	[illegible]	[illegible]	F	F
729	00037be2-d64b-adb8-c7e7-90dcffa144bf	518	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	[illegible]	F	F	F
730	00037be2-d64b-adb8-c7e7-90dcffa144bf	519	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	[illegible]	F	F	F
731	00037be2-d64b-adb8-c7e7-90dcffa144bf	520	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
732	00037be2-d64b-adb8-c7e7-90dcffa144bf	521	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
733	00037be2-d64b-adb8-c7e7-90dcffa144bf	522	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
734	00037be2-d64b-adb8-c7e7-90dcffa144bf	523	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
735	00037be2-d64b-adb8-c7e7-90dcffa144bf	524	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
736	00037be2-d64b-adb8-c7e7-90dcffa144bf	525	age_45_64	F	white	nonhispanic	ckd_3	ckd_3	F	F	F	F
737	00037be2-d64b-adb8-c7e7-90dcffa144bf	526	age_45_64	F	white	nonhispanic	ckd_3	ckd_4	F	F	F	F
738	00037be2-d64b-adb8-c7e7-90dcffa144bf	527	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
739	00037be2-d64b-adb8-c7e7-90dcffa144bf	528	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
740	00037be2-d64b-adb8-c7e7-90dcffa144bf	529	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
741	00037be2-d64b-adb8-c7e7-90dcffa144bf	530	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
742	00037be2-d64b-adb8-c7e7-90dcffa144bf	531	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
743	00037be2-d64b-adb8-c7e7-90dcffa144bf	532	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
744	00037be2-d64b-adb8-c7e7-90dcffa144bf	533	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
745	00037be2-d64b-adb8-c7e7-90dcffa144bf	534	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
746	00037be2-d64b-adb8-c7e7-90dcffa144bf	535	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
747	00037be2-d64b-adb8-c7e7-90dcffa144bf	536	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
748	00037be2-d64b-adb8-c7e7-90dcffa144bf	537	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
749	00037be2-d64b-adb8-c7e7-90dcffa144bf	538	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
750	00037be2-d64b-adb8-c7e7-90dcffa144bf	539	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F
751	00037be2-d64b-adb8-c7e7-90dcffa144bf	540	age_45_64	F	white	nonhispanic	ckd_4	ckd_4	F	F	F	F

Validation

```sql
select description, count(*) tally from conditions
  where description rlike('(Chronic kidney disease|End stage renal)')
  group by description order by description
```

missouri_pre

description	tally
Chronic kidney disease stage 1 (disorder)	3024
Chronic kidney disease stage 2 (disorder)	436
Chronic kidney disease stage 3 (disorder)	25

missouri_test

description	tally
Chronic kidney disease stage 1 (disorder)	3837
Chronic kidney disease stage 2 (disorder)	6557
Chronic kidney disease stage 3 (disorder)	18607
Chronic kidney disease stage 4 (disorder)	10585
End stage renal disease (disorder)	17272

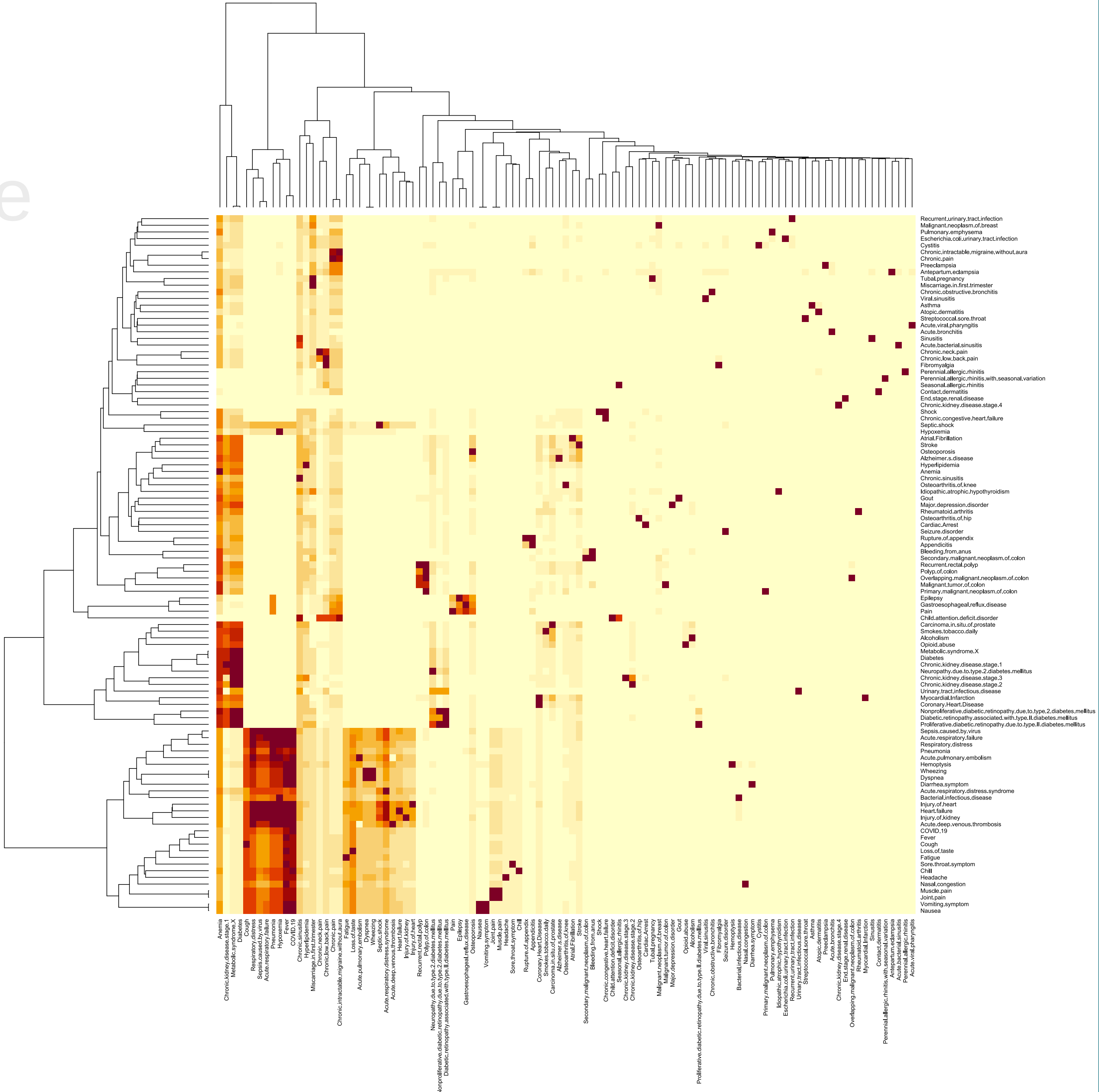
Individual concepts

Including Covid | Mercy EHR | Pre-Covid

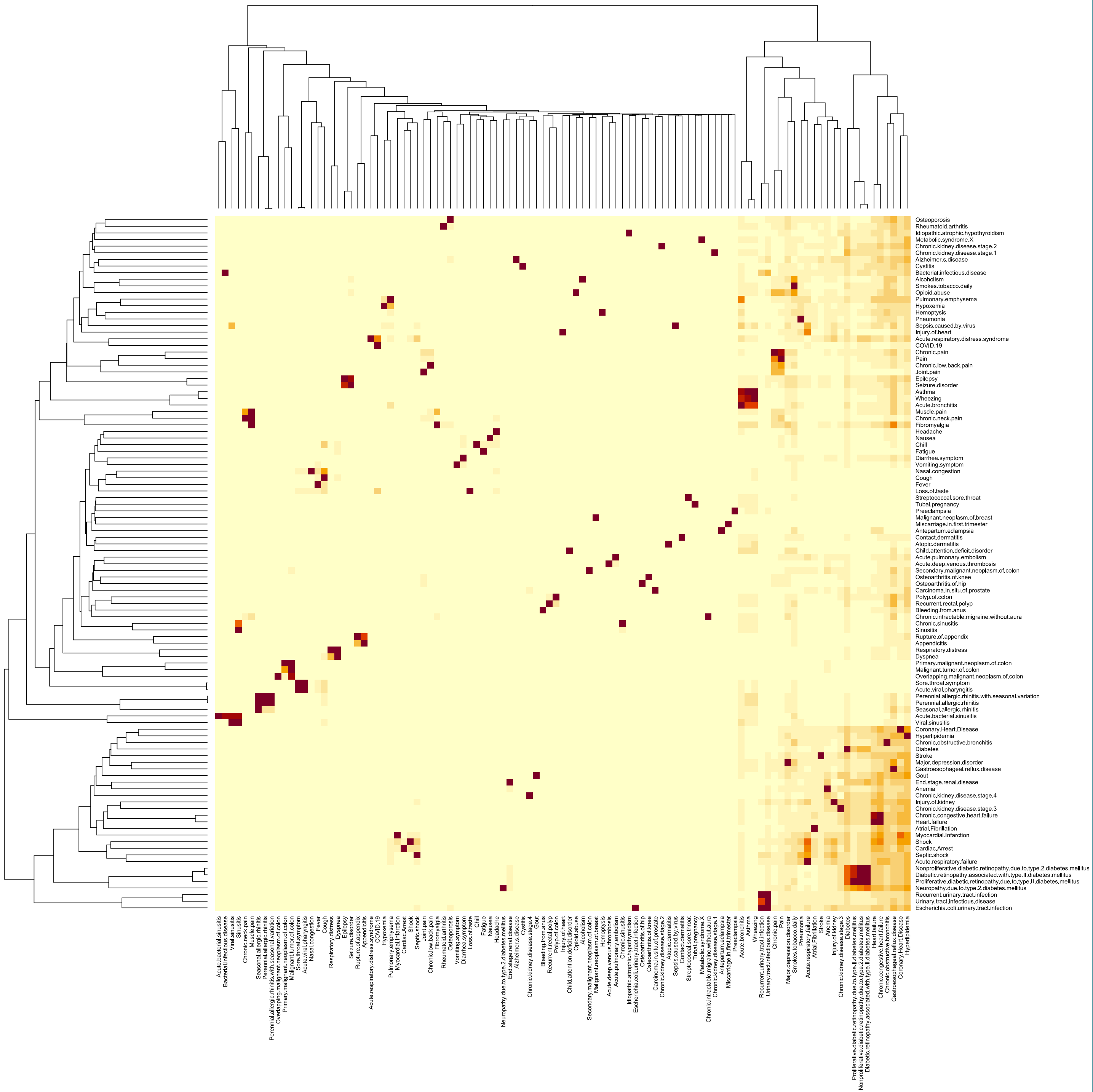
Rank	Synthea Concept	Mercy EHR Concept	Pre-Covid Synthea Concept
1	Suspected COVID-19	Essential hypertension (disorder)	Viral sinusitis (disorder)
2	COVID-19	Diabetes mellitus type 2 (disorder)	Acute viral pharyngitis (disorder)
3	Fever (finding)	Hyperlipidemia (disorder)	Acute bronchitis (disorder)
4	Cough (finding)	Cough (finding)	Normal pregnancy
5	Loss of taste (finding)	Asthma (disorder)	Streptococcal sore throat (disorder)
6	Viral sinusitis (disorder)	Gastroesophageal reflux disease (disorder)	Otitis media
7	Fatigue (finding)	Coronary arteriosclerosis (disorder)	Unhealthy alcohol drinking behavior (finding)
8	Sputum finding (finding)	Hypertriglyceridemia (disorder)	Severe anxiety (panic) (finding
9	Hypoxemia (disorder)	Joint pain (finding)	Sprain of ankle
10	Respiratory distress (finding)	Body mass index 30+ - obesity (finding)	Prediabetes

dominated by Covid-19 | common chronic conditions | largely acute conditions

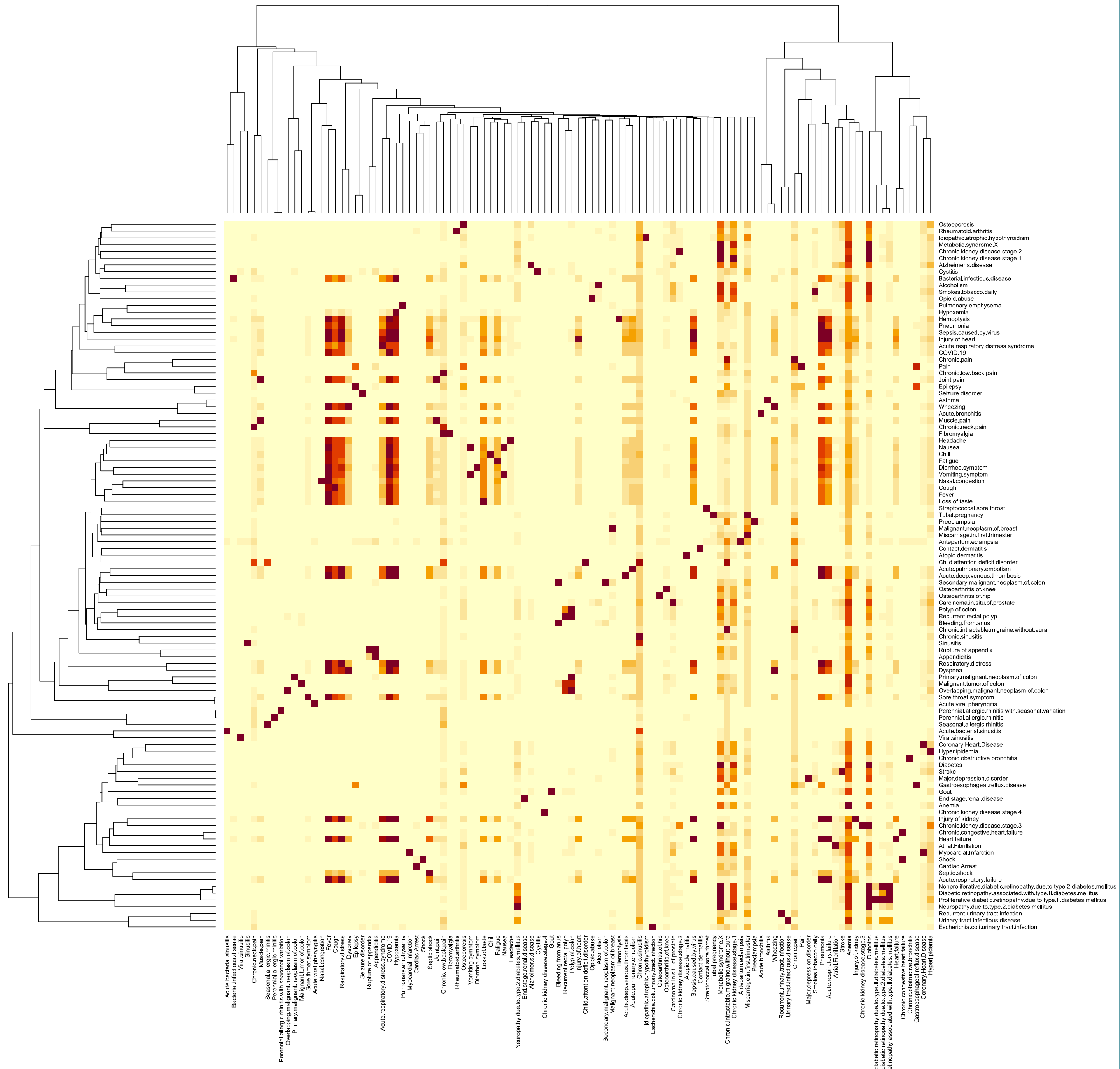
Co-occurrence heatmap for Synthea data



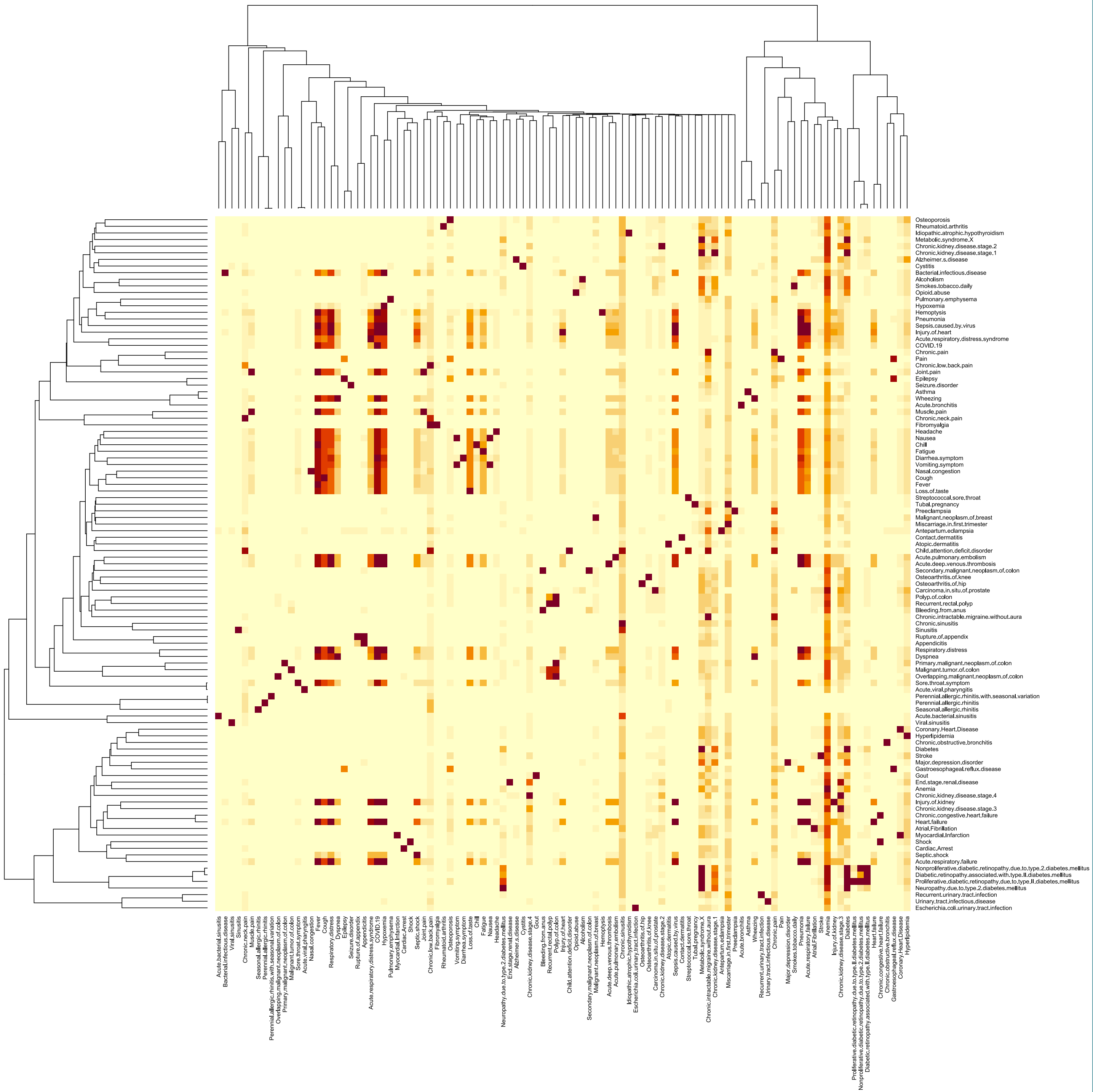
EMR data



Synthea data



Synthea data after *Virtual Generalist* CKD



Co-Occurrence Probability Measure (COPM)

a) across all 107 conditions

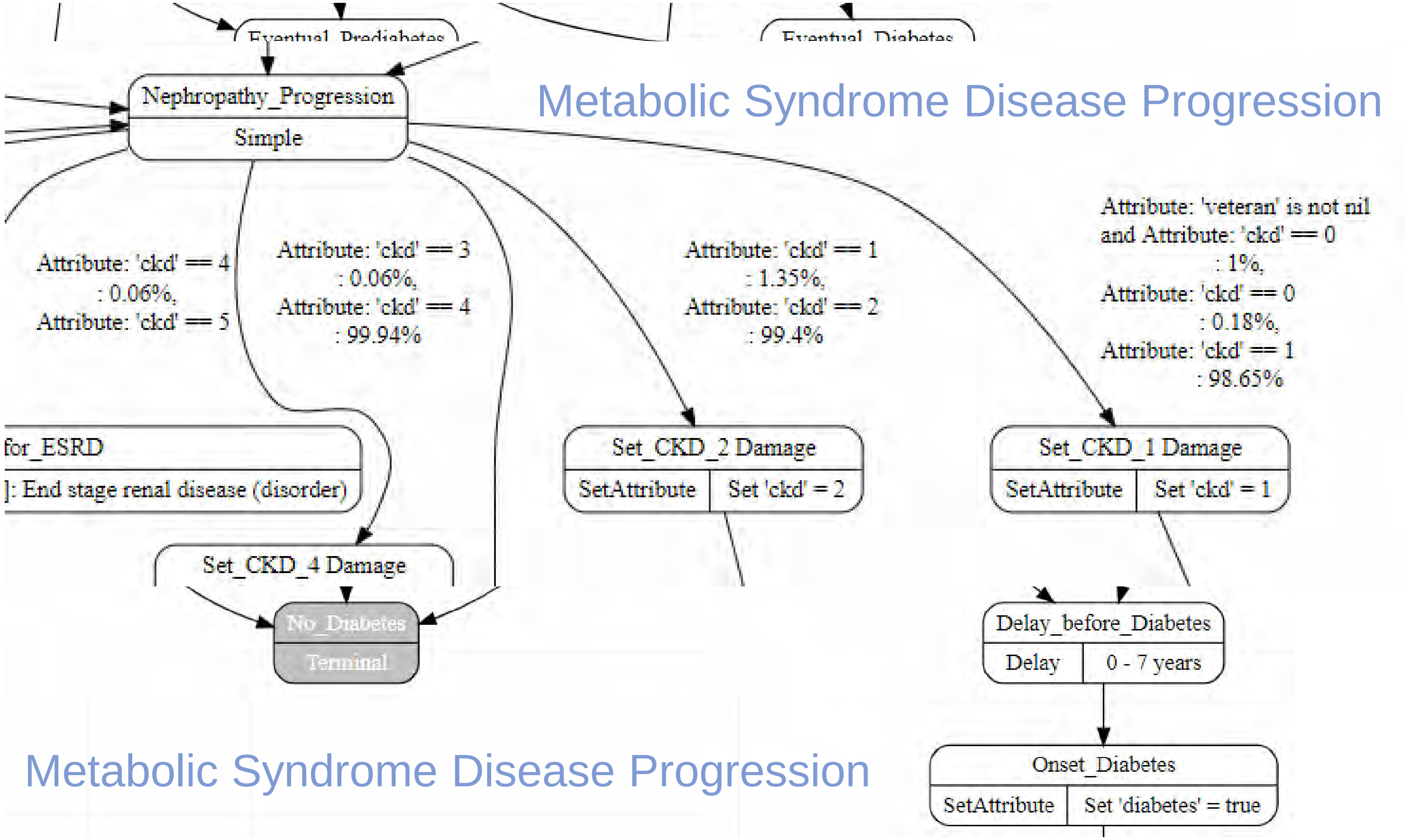
From \ To	actual	simulation_pre	simulation_post
actual	0	0.2123	0.2008
simulation_pre		0	0.0110
simulation_post			0

b) focused on just 6 CKD-related conditions

From \ To	actual	simulation_pre	simulation_post
actual	0	0.8680	0.2430
simulation_pre		0	0.1207
simulation_post			0

Design pattern recommendations

- ***No delay***
  - disease is not your destiny
- ***Mechanistic progression***
  - allow interventions
- ***More modular modules***
  - separate treatment from disease incidence and progression
  - use attributes to communicate between modules



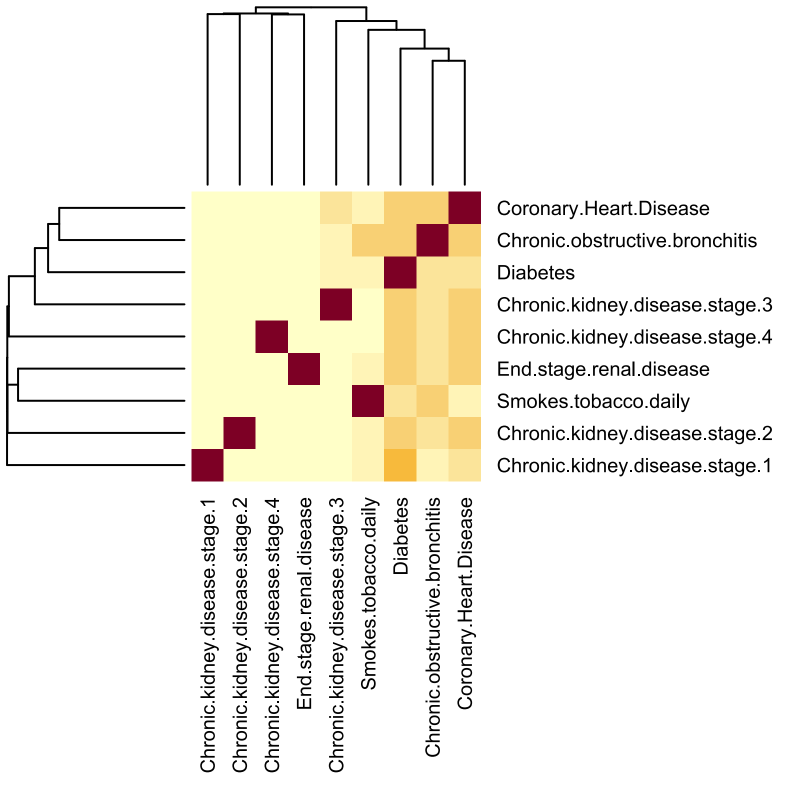
Bonus Slides!

data visualization helps find bugs

Debugging Synthea with co-occurrence

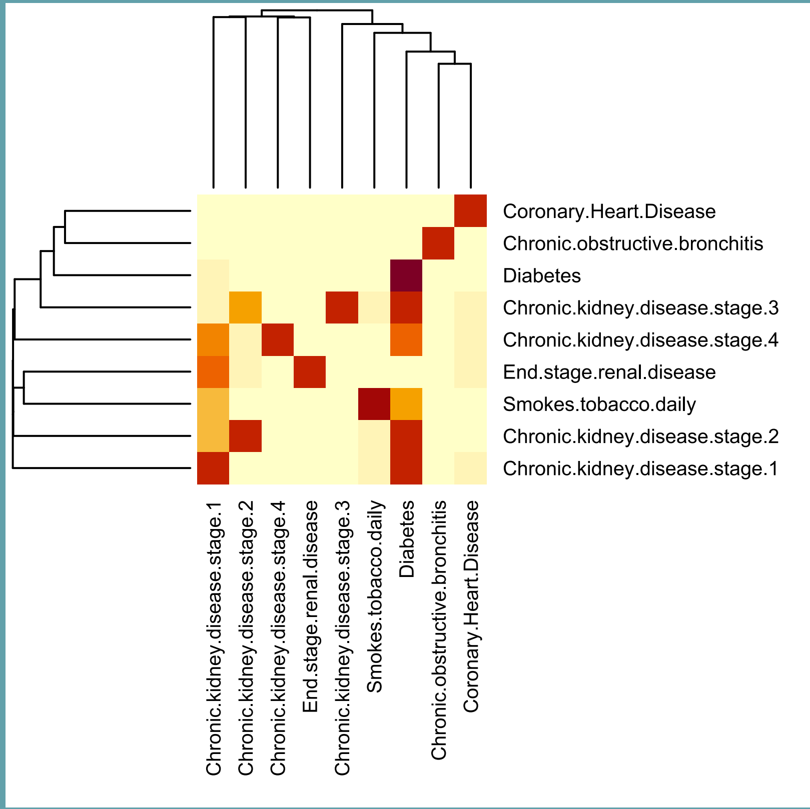
EMR data:

mutually exclusive stages



Synthea data:

stages can co-occur

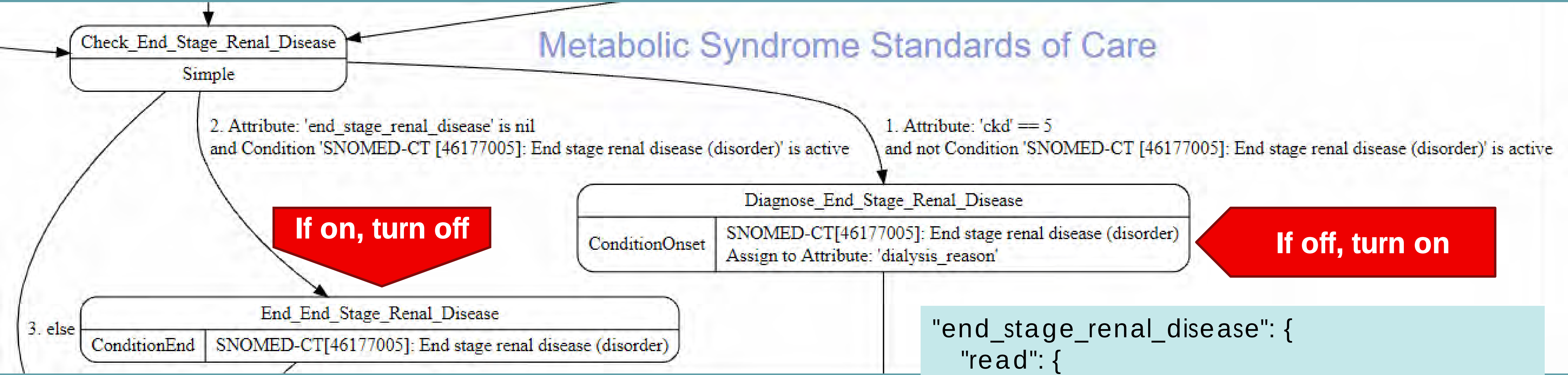


‘End stage renal disease’ condition toggles on and off

```
select patient, description, start, stop, encounter, code from conditions
  where description rlike('(Chronic kidney disease|End stage renal)')
  order by patient, start
```

patient	description	start	stop	encounter
00118915-7610-1be1-fd03-21811ac23b71	Chronic kidney disease stage 1 (disorder)	2014-09-17	null	cdfd0045-026b-a
00154e43-88d2-f074-2810-23c2ed04235f	Chronic kidney disease stage 3 (disorder)	2010-08-26	null	ac3401ab-e826-7
001792aa-daec-beab-969b-1a7a98c0dc67	Chronic kidney disease stage 3 (disorder)	2005-04-17	null	6b2c1374-3280-f
001792aa-daec-beab-969b-1a7a98c0dc67	End stage renal disease (disorder)	2008-07-13	2009-07-19	f6787445-a37a-a
001792aa-daec-beab-969b-1a7a98c0dc67	End stage renal disease (disorder)	2010-07-25	2011-07-31	4f70d0ec-3965-1
001792aa-daec-beab-969b-1a7a98c0dc67	End stage renal disease (disorder)	2012-08-05	2013-08-11	93ce8bb5-5949-1

End Stage Renal Disease




```
"end_stage_renal_disease": {
  "read": {
    "HealthInsuranceModule": [],
    "Metabolic Syndrome Standards of Care": [
      "Check_End_Stage_Renal_Disease"
    ]
  },
  "write": {},
  "example_values": []
},
```

Nothing writes this attribute; it is always nil

ON IMPROVING REALISM OF DISEASE MODULES IN SYNTHEA™

Social Determinant-Based Enhancements to Conditional Transition Logic

Team LMI

Brant Horio
Greg Pekar
Simon Whittle
Maureen Merkl
Linna Qiao

Response to the 2021 HHS Synthetic Health Data Challenge

Category I: Enhancements to Synthea™
Opioids Use Case

October 19, 2021

LMI

RESEARCH QUESTION AND KEY TECHNICAL IDEA

- Opioid Use Disorder (OUD) is a crisis
  - **Deadly outcomes:** 69,700+ Americans dying from opioid overdoses in 2020 (36% increase from prior year)
  - **Hard to fix:** OUD is a highly individual and complex care condition and highly influenced by SDOH
  - **Ongoing problem:** U.S. CBP confiscated more fentanyl in first half 2021 than last 3 years (Kaminsky 2021)
- Our research focused on one pathway to OUD
  - **Prescribed opioids may result in OUD:** often due to long term need to help with chronic pain
  - **SDOH (occupation) can drive onset of chronic pain:** particularly for heavy manual labor in rural areas
- The key technical idea was to demonstrate
  - a **software workflow** that mines open-source secondary data sources for SDOH (occupation)
  - modifications to the Synthea codebase to **operationalize the SDOH** (following published findings that could bridge occupation to relevant transition logic in Synthea's OUD-related state machines, and
  - **increased realism** in Synthea's generated populations with respect to opioid prescriptions

OUR APPROACH MINED SECONDARY DATA FOR OUD-RELEVANT SDOH TO AUGMENT SIMULATED PERSON JOURNEYS

SYNTHEA

User specifications inform Synthea runs

User-specifications
e.g., location, population, diseases

Synthea draws from publicly available information and knowledge

Publicly-available demographic data, e.g., US Census

Health incidence statistics, e.g., CDC prevalence and incidence rates, NIH reports

Clinical practice guidelines and protocols

Medical coding dictionaries

Data is transformed

PATIENT PROFILE
Demographics
Residence location

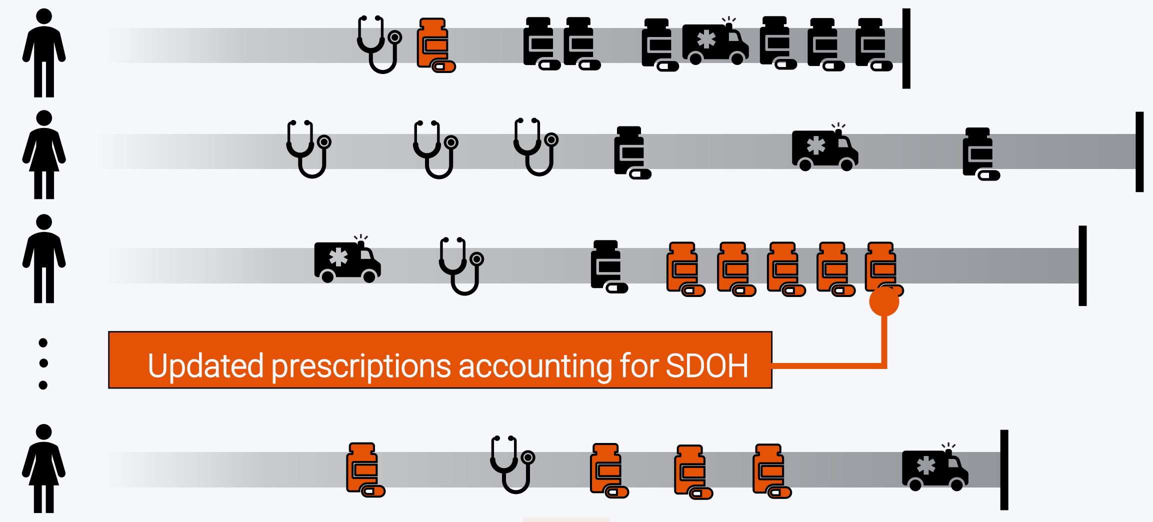
HEALTH PROFILE
Health condition and disease

CAREMAPS

CODED NARRATIVES

SOCIAL DETERMINANTS OF HEALTH
Household financial status
Insurance
Occupation (w.r.t. chronic pain)

An agent-based approach is used to generate a fully-synthetic population, simulating individuals from birth-to-death while tracking their healthcare journey



Our technical approach seeks to automate data mining of publicly available data sources for SDOH that better characterize communities at the local level

Lat/Lon → Census Tract → Localized SDOH (i.e., occupation)

Repeatable and accessible enhancement to Synthea codebase that introduces SDOH (or other Census tract-related Person attributes) for more realistic modeling

WE OPERATIONALIZE SDOH DATA AS AGENT ATTRIBUTES TO ENHANCE TRANSITION LOGIC IN SYNTHEA'S MODULE BUILDER

Prepare SDOH Data

Mine data and create census tract-indexed file to occupation information for Bangor, ME

Assign SDOH Attributes to Person Agents

Assign an occupation to each person based on Census statistics for our targeted occupation classes in their Census tract communities

Modify Relevant Disease Modules

Adjust transition probabilities for a chronic low back pain condition, based on occupation and gender

Run Synthea and Validate

Run ten trials of 32,000 patients with legacy_Synthea and LMI_Synthea

Compare simulated outcomes for number of opioid prescriptions per capita

PREPARE SDOH DATA

- We scoped research to Bangor, ME due to higher than national average prevalence of OUD and predominant SDOH that strongly influence OUD
- Based on collaboration with University of Maine OUD researchers and published literature, we focused on
  - Maine's prevalence of forestry and fishing occupations
  - Correlation of these occupations to chronic musculoskeletal pain conditions
  - Chronic pain as a pathway to prescribed opioids, potential abuse, and OUD

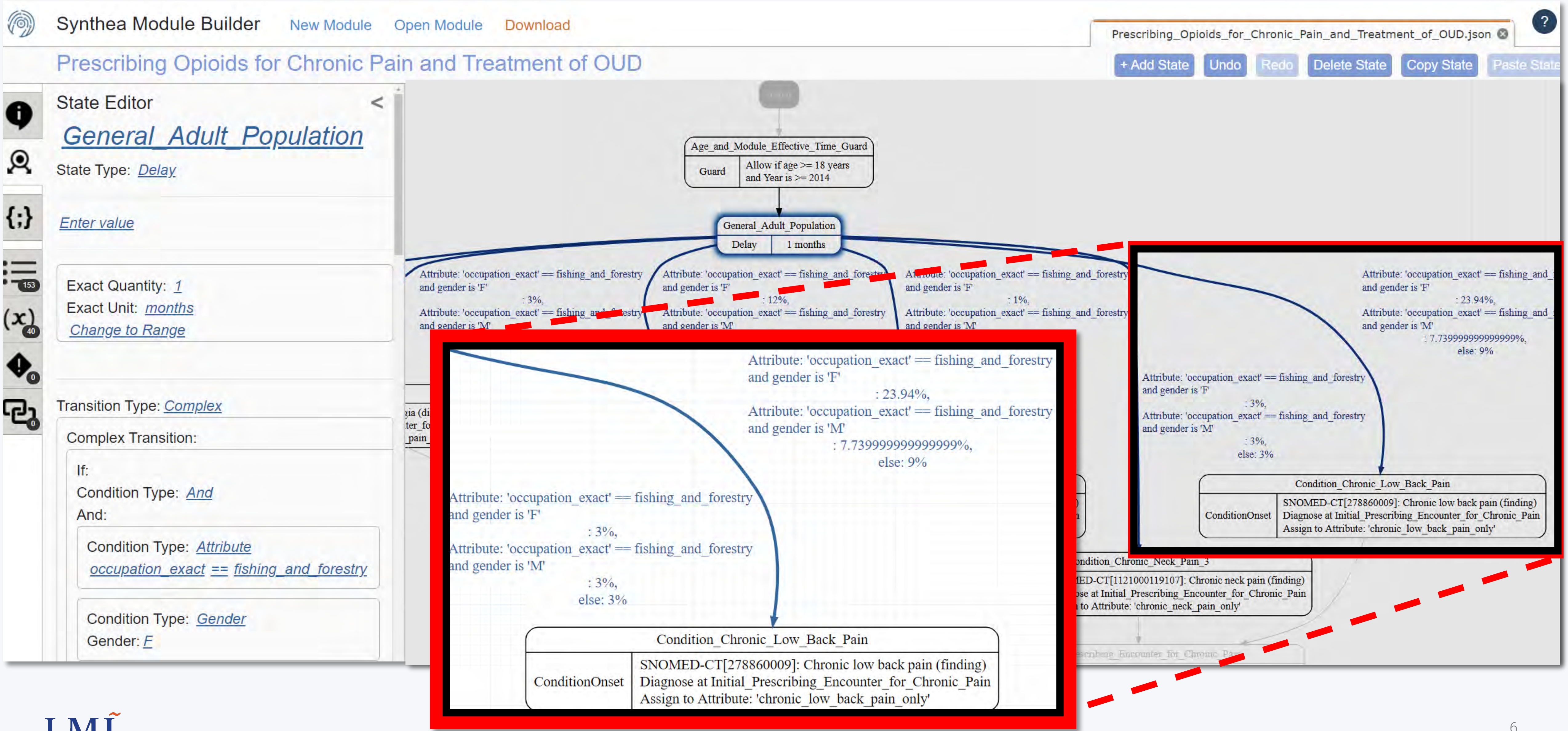
1865 THE UNIVERSITY OF MAINE

- We processed American Community Survey data to derive likelihood of individuals having forestry and fishing occupations by Census tract in Penobscot County, ME (where Bangor is located) to build a tract-indexed file to assign Person occupations
- Drawing from literature (Yang, Halderman, Lu, and Baker 2016) assessing chronic low back pain risk associated with occupations and gender, to inform state machine transition probabilities
- Validation patterns were collaborated on with University of Maine researchers and focused on opioid prescription counts for State and County

ASSIGN TO AGENTS

- Built a data ingestion method to access the tract-indexed file we created for occupation data
- Assigned Persons to a Census tract
  - For greater localized level of detail than zip code and to align with our data sources of interest
  - Based on shortest distance between Synthea's provided latitude-longitude residence coordinates to the centroid of the nearest Census block
  - Given Block assignment, then we assigned the Person to the appropriate Census tract
- Assigned occupation to Persons
  - Using assigned tract number, probabilistically assign occupation to Person by referencing occupation file
- Person owns one new attribute for occupation which can now be used for transitional conditional logic in the disease module state machines

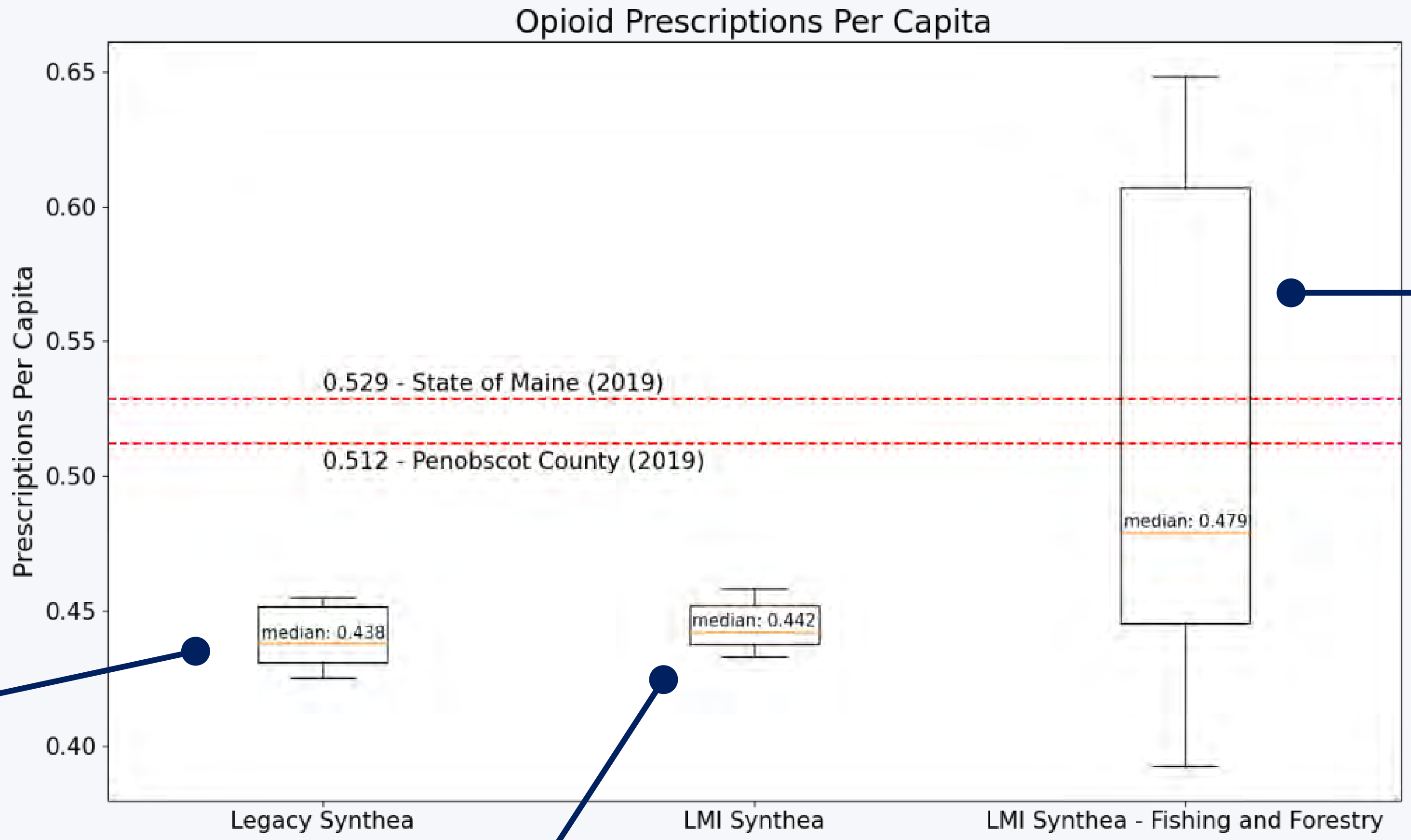
MODIFICATION OF THE DISEASE MODULES



RUN SYNTHEA AND VALIDATE

- To validate our modifications to Synthea:
  - We ran 10 trials, each with different random seeds
  - Each run generated 32,000 patients to approximately represent the entire City of Bangor, ME
  - We used the same 10 seeds for legacy_Synthea and LMI_Synthea to allow comparison between software versions
- We used the Prescription Monitoring Program Annual Report 2020 (Maine HSS 2021) and data from the CDC (CDC 2020) as our ground truth data for the number of opioid prescriptions in Maine

VALIDATION RESULTS SHOW BETTER OUTCOMES FOR OUR SUBPOPULATION OF INTEREST



Legacy Synthea results for Bangor, ME fall short of both County and State benchmarks.

LMI Synthea shows slight improvement, but as the subpopulation we augmented was very small compared to the overall population, it did little to shift metrics at the Bangor scale.

Evaluating only the subpopulation we changed in the fishing and forestry occupation, both State and County benchmarks fall within the interquartile range of our simulated results.

The results are promising in that outcomes are demonstrative of how population level representation might be improved by adjusting many relevant subpopulations.

LMI

BENEFITS TO HEALTHCARE COMMUNITY

- Benefits to Synthea
  - Minimal codebase modifications to allow an easy pull request
  - Method of ingesting Census tract level data and assigning to patients can be generalized to other similar data sets at a local scale
  - Focus on enhancing Person attributes to integrate SDOH allows Synthea ecosystem to be fully leveraged (e.g., ModuleBuilder)
- Benefits to researchers
  - Provides additional experiment factors to incorporate into Synthea's disease modules for greater detail for the conditional logic and state transition path possibilities
  - Repeatable and accessible enhancement to Synthea codebase that introduces SDOH (or other Census tract-related Person attributes) for more realistic modeling
- Benefits to health IT developers
  - Provides a working framework to ingest secondary data and assign them as patient attributes
- Broader healthcare community
  - Individualized SDOH drives complex care and critically needs to be better understood, analyzed, and modeled to further advance patient-centered outcomes research

FUTURE WORK

- Additional SDOH factors based on Census tracts and blocks may be examined (e.g., homelessness, access to care, access to food)
- Work is needed to find better ways to account for relevant correlations between SDOH (e.g., our assignment of occupation neglected potentially relevant relationships to Synthea's assigned socioeconomic status or age)
- The team will be continuing to collaborate with University of Maine researchers with the objective of integrating Synthea for OUD research

SUCCESSES AND LESSONS LEARNED

- There is a lot of code
  - Big learning curve to understand the codebase and how the modules interact with one another (e.g., geography folder has a fully implemented quad tree we could have used for centroid distance calculations—we implemented a less efficient in a sorted map-based approach to finding the closest centroid)
- Successes in working with existing Synthea code base
  - Minimal modifications to preserve functioning software with least risk of “breaking it”
    - Integration of external analysis (e.g., tract and occupation) into the Synthea process
    - Use of input files and existing data loader methods
    - Primarily modified the pickDemographics( ) step of the Person generation logic
  - Preserved software practices and workflows to make sure our enhancements looked roughly like the rest of the project (would not be a stumbling block to existing Synthea developers and users)
  - Focused on capitalizing on the very polished Module Builder application to operationalize our enhancements in other modules

The Necessity of Realistic Synthetic Health Data Development Environments

Category II Entry: Novel Uses of Synthea Generated Synthetic Data

Team: Particle Health, **Submitter**: Parker Bannister

Background and Objectives

- APIs will be the tool driving national access to health data
    - 21st Century Cures Act and TEFCA
- Synthetic health data vs. Real health data
    - Synthea
        - Single CCDA documents, separately generated free-text notes
    - Access to Data
        - Download CCDA to Local Machine
        - APIs for FHIR, not CCDA
- Why is a realistic development environment needed?
    - Allows researchers and developers to seamlessly transition to real patient information
        - Time, Cost, and Liability
- The Particle Health Sandbox Objective
    - Point-In-Time documents
    - In-Document Synthetic provider notes
    - Focus on synthetic populations with specific conditions
    - Validation

Document Generation pt.1

- Synthea_Runner.py
  - Generates base Synthea documents
  - Finds patients with conditions specified to create population of interest
    - Regex - Synthea symptoms.csv
  - Stores files for population of interest and processes them with the point in time document generator
  - Finally Validates Results

Synthea_Runner.py

Particle-Health-Sandbox-Envi

github.com/ParticleHealth/Particle-Health-Sandbox-Environment/blob/master/synthea-master/Synthea_Runner.py

```python
list_of_conditions = ['covid19', 'diabetes', 'lung cancer', 'opioid addiction']
parser = argparse.ArgumentParser()
parser.add_argument("--num-patients", type=str, required=True)
parser.add_argument("--condition", type=str, required=True, choices=list_of_conditions)
args = parser.parse_args(sys.argv[1:])

num_patients = args.num_patients
condition = args.condition

list_of_modules = ["covid19", "metabolic*", "lung_cancer*", "opioid_addiction"]
if condition == 'covid19':
    module = list_of_modules[0]
if condition == 'diabetes':
    module = list_of_modules[1]
    if int(num_patients) < 10:
        num_patients = '25'
if condition == 'lung cancer':
    module = list_of_modules[2]
    num_patients = '2500'
if condition == 'opioid addiction':
    module = list_of_modules[3]
    if int(num_patients) < 10:
        num_patients = '25'



#run synthea with module and number of patients specified
print('\n' + 'Running Synthea to Generate Base Synthetic Patient Data' + '\n')
os.system('./run_synthea -p ' + num_patients + ' -m ' + module + ' -a 0-95')

#load csv produced from synthea with conditions of population generated
print('\n' + "Finding Patients with Conditiions of Interest"+ '\n')
symptoms_csv_load = []
with open('./output/symptoms/csv/symptoms.csv') as symptoms_csv:
    for i in symptoms_csv:
        symptoms_csv_load.append(i)

#define function to find condition of interest
def check_if_match(words):
    patterns = re.split('(\s|\,)', words)
    for i in range(len(patterns)):
        patterns[i] = patterns[i].lower()
    results = get_close_matches(condition, patterns)
    return results, patterns[0]

#change opioid addiction to terms found in symptom csv
if condition =='opioid addiction':
    condition = 'drug overdose'

#get unique patient ids for patients with condition of interest
```

particle

Synthea_Runner.py

```python
os.system('python point_in_time_document_generator.py')

#copy output files and directory to output folder:
for i in os.listdir('./pitd_gen_output'):
    shutil.copytree('./pitd_gen_output/' + i, '../' + condition + '_output_' + str(today) + '/generator_output/' + i)

copyfile('./output_directory.csv', '../' + condition + '_output_' + str(today) + '/output_directory.csv')

#clear notes, synthea ccds, pitd gen output_file, directory
for i in os.listdir("./Notes"):
    os.remove('./Notes/' + i)
for i in os.listdir("./Synthea_CCDs"):
    os.remove('./Synthea_CCDs/' + i)
for i in os.listdir("./output"):
    if i.startswith('.'):
        os.remove('./output/' + i)
    else:
        shutil.rmtree('./output/' + i, ignore_errors = True)
for i in os.listdir("./pitd_gen_output"):
    shutil.rmtree('./pitd_gen_output/' + i, ignore_errors = True)
os.remove('./output_directory.csv')

print('\n' + 'Generation Complete' + '\n')

#validate point in time ccda documents
print('\n' + 'Validating Output Point In Time CCDA Documents' + '\n')

#loop thru output files to validate
patients = []
val_output = []
urllib3.disable_warnings()
for i in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output'):
    if not i.startswith('.'):
        print(i)
        #loop thru folders for patient
        counter = 0
        for j in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output/' + i):
            if not j.startswith('.'):
                print('\tFOLDER:' + j)
                #loop thru files for patient
                for k in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output/' + i +"/" + j):
                    if not k.startswith('.'):
                        print('\t\tFile: ' + k)
                        patients.append(i)
                        folder_name = i
                        data_file = '../' + condition + '_output_' + str(today) + '/generator_output/' + i +"/" + j + '/' + k

                        url = "https://ccda.healthit.gov/scorecard/ccdascorecardservice2"
                        myfile = {"ccdaFile": (k, open(data_file, "rb"))}
                        r = requests.post(url, files = myfile, verify = False).json()
```

Document Generation pt.2

- Point_in_Time_Document_Generator.py
  - Parses base Synthea CCDA and stores sections relevant to point in time document types
    - LXML
  - Templates sections into new document
    - Jinja2
    - XML Templates for new point in time documents
    - Addition of in-document notes

Point_In_Time_Document_Generator.py

Particle-Health-Sandbox-Envi

github.com/ParticleHealth/Particle-Health-Sandbox-Environment/blob/master/synthea-master/point_in_time_document_generator.py

master · Particle-Health-Sandbox-Environment / synthea-master / point_in_time_document_generator.py / <> Jump to · Go to file

pbannister123 Update Synthea-Master · Latest commit d398e07 3 days ago · History

1 contributor

606 lines (552 sloc) 43.8 KB · Raw · Blame

```
#Particle Health CCDA Point In Time Document Generator

#import relevant packages
import jinja2
from jinja2 import Environment
from jinja2 import FileSystemLoader
from lxml import etree
from io import StringIO, BytesIO
import os
import csv
import string
import re


#lxml method
#parses Synthea CCD document to save variables we want to populated the template CCDAs with

#loop through n synthea files we want to populate CCDAs for
for filename in os.listdir('Synthea_CCDs'):
    #skip .DS_Store files
    if not filename.startswith('.'):
        #take file n
        with open("Synthea_CCDs/" + filename) as openfile:
            #reset variables to as if section in Synthea CCD didnt exist to avoid
            # populating template with previous ccd information
            encounter_tr = "<tr><td>Data Unknown</td></tr>"
            vital_signs_text = "<text>Data Unknown</text>"
            diagnostic_results_text = "<text>Data Unknown</text>"
            plan_of_care_text = "<text>Data Unknown</text>"
            social_history_text = "<text>Data Unknown</text>"
            problems_text = "<text>Data Unknown</text>"
            medications_text = "<text>Data Unknown</text>"
            allergies_text = "<text>Data Unknown</text>"
            immunizations_text = "<text>Data Unknown</text>"
            encounter_entry = ""
            vital_signs_entry = ""
            diagnostic_results_entry = ""
            plan_of_care_entry = ""
            social_history_entry = ""
```

particle

Point_In_Time_Document_Generator.py

Particle-Health-Sandbox-Env

github.com/ParticleHealth/Particle-Health-Sandbox-Environment/blob/master/synthea-master/point_in_time_document_generator.py

```python
                        immunizations_entry.append(etree.tounicode(entry, pretty_print=True))
                    else:
                        immunizations_entry = ""
                immunizations_entry = "".join(immunizations_entry)

        #fix null imaging section error:
        for child in tree.find("{urn:hl7-org:v3}component/{urn:hl7-org:v3}structuredBody"):
            for i in child.find("{urn:hl7-org:v3}section"):
                if i.tag == "{urn:hl7-org:v3}code":
                    if i.attrib['code'] == '18748-4':
                        child.getparent().remove(child)

        #jinja2 populate template:
        #fills blank templates with saved data parsed from synthea ccd above
        #load templates and set up environment for jinja2
        template_file_loader = FileSystemLoader('templates')
        env = Environment(loader=template_file_loader)
        encounter_summary_template = env.get_template('Encounter_Summary_Template.xml')
        refill_summary_template = env.get_template('Refill_Summary_Template.xml')
        lab_summary_template = env.get_template('Lab_Summary_Template.xml')
        immunizations_summary_template = env.get_template('Immunizations_Summary_Template.xml')

        #render encounter summary template with saved variables and save to output
        encounter_summary_output = encounter_summary_template.render(Effective_Time = effective_time,
                                                                Record_Target = record_target,
                                                                Author = author,
                                                                Custodian = custodian,
                                                                Documentation_Of = documentation_of,
                                                                Encounter_TR = encounter_tr,
                                                                Encounter_Entry = encounter_entry,
                                                                Vital_Signs_Text = vital_signs_text,
                                                                Vital_Signs_Entry = vital_signs_entry,
                                                                Diagnostic_Results_Text = diagnostic_results_text,
                                                                Diagnostic_Results_Entry = diagnostic_results_entry,
                                                                Plan_Of_Care_Text = plan_of_care_text,
                                                                Plan_Of_Care_Entry = plan_of_care_entry,
                                                                Social_History_Text = social_history_text,
                                                                Social_History_Entry = social_history_entry)

        #render refill summary template with saved variables and save to output
        refill_summary_output = refill_summary_template.render(Effective_Time = effective_time,
                                                            Record_Target = record_target,
                                                            Author = author,
                                                            Custodian = custodian,
                                                            Documentation_Of = documentation_of,
                                                            Encounter_TR = encounter_tr,
                                                            Encounter_Entry = encounter_entry,
                                                            Plan_Of_Care_Text = plan_of_care_text,
                                                            Plan_Of_Care_Entry = plan_of_care_entry,
                                                            Social_History_Text = social_history_text,
```

particle

Point_In_Time_Document_Generator.py

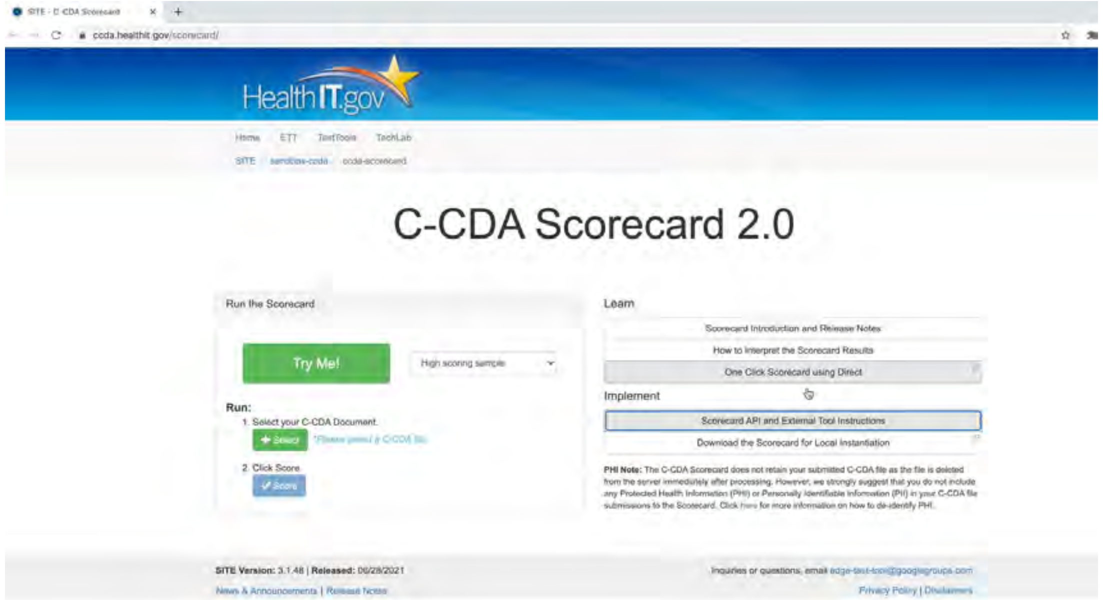
Particle-Health-Sandbox-Envi

github.com/ParticleHealth/Particle-Health-Sandbox-Environment/blob/master/synthea-master/templates/Encounter_Summary_Template.xml

```
<realmCode code="US" />
<typeId
        root="2.16.840.1.113883.1.3"
        extension="POCD_HD000040" />
<templateId root="2.16.840.1.113883.10.20.1" />
<templateId
        root="2.16.840.1.113883.10.20.22.1.1"
        extension="2015-08-01" />
<templateId
        root="2.16.840.1.113883.10.20.22.1.2"
        extension="2015-08-01" />
<id root="2.16.840.1.113883.19.5" extension="46f6aa9d-c38c-4215-833e-19268dadb4ca" assigningAuthorityName="https://github.com/synthetichealth/synthea"/>
<code code="11506-3" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Subsequent evaluation note" />
<title>Health Encounter Summary</title>

{{Effective_Time}}
<confidentialityCode
        code="N"
        codeSystem="2.16.840.1.113883.5.25" />
        <languageCode code="en-US"/>

{{Record_Target}}
{{Author}}
{{Custodian}}
{{Documentation_Of}}
<component>
    <structuredBody>
        <component>
            <!-- Encounters -->
            <section nullFlavor="NI">
              <templateId root="2.16.840.1.113883.10.20.22.2.22"
                          extension="2015-08-01"/>
              <code code="46240-8"
                    codeSystem="2.16.840.1.113883.6.1"
                    codeSystemName="LOINC"
                    displayName="History of encounters"/>
              <title>Encounters</title>
              <text>
                <table border="1" width="100%">
                  <thead>
                    <tr>
                      <th>Start</th>
                      <th>Stop</th>
                      <th>Description</th>
                      <th>Code</th>
                    </tr>
                  </thead>
                  <tbody>

                      {{Encounter_TR}}
                  </tbody>
```

particle

Validation



- HealthIT.gov’s CCDA Scorecard 2.0
    - API Implementation
- CSV of Results for synthetic population generated
    - File score validation per file generated for entire population of synthetic patients

Validation Portion of Synthea_Runner.py and Results

```python

#loop thru output files to validate
patients = []
val_output = []
urllib3.disable_warnings()
for i in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output'):
    if not i.startswith('.'):
        print(i)
        #loop thru folders for patient
        counter = 0
        for j in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output/' + i):
            if not j.startswith('.'):
                print('\tFOLDER:' + j)
                #loop thru files for patient
                for k in os.listdir('../' + condition + '_output_' + str(today) + '/generator_output/' + i +"/" + j):
                    if not k.startswith('.'):
                        print('\t\tFile: ' + k)
                        patients.append(i)
                        folder_name = i
                        data_file = '../' + condition + '_output_' + str(today) + '/generator_output/' + i +"/" + j + '/' + k

                        url = "https://ccda.healthit.gov/scorecard/ccdascorecardservice2"
                        myfile = {"ccdaFile": (k, open(data_file, "rb"))}
                        r = requests.post(url, files = myfile, verify = False).json()
                        val_output.append(list(r.items()))

val_table = pd.DataFrame(data = val_output, columns = ['ErrorMessage','FileName', 'CCDADocumentType', 'Results', 'ReferenceResults', 'ErrorList', 'SchemaErrors'
val_table.insert(0, "Patient", patients, True)

#write validation output to csv
val_table.to_csv('../' + condition + '_output_' + str(today) + '/validation_results.csv')

print('\n' + 'Validation Complete' + '\n')
```

Validation Portion of Synthea_Runner.py and Results

	Patient	ErrorMessage	FileName	CCDADocumentType	Results
0	Era_Crist	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP87702___2021-07-08T2(	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
1	Era_Crist	('errorMessage', None)	('filename', 'Encounter_Summary___PCP87702___2021-07-08T202520.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
2	Era_Crist	('errorMessage', None)	('filename', 'Lab_Summary___PCP87702___2021-07-08T202520.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
3	Era_Crist	('errorMessage', None)	('filename', 'Immunizations_Summary___PCP87702___2021-07-08T202520	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
4	Lettie_Erdman	('errorMessage', None)	('filename', 'Encounter_Summary___PCP49548___2021-07-08T071356.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
5	Lettie_Erdman	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP49548___2021-07-08T07	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
6	Lettie_Erdman	('errorMessage', None)	('filename', 'Refill_Summary___PCP49548___2021-07-08T071356.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
7	Lettie_Erdman	('errorMessage', None)	('filename', 'Lab_Summary___PCP49548___2021-07-08T071356.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
8	Lettie_Erdman	('errorMessage', None)	('filename', 'Encounter_Summary___PCP49548___2021-07-08T071356.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
9	Lettie_Erdman	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP49548___2021-07-08T07	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
10	Lettie_Erdman	('errorMessage', None)	('filename', 'Immunizations_Summary___PCP49548___2021-07-08T071356	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
11	Crystle_Price	('errorMessage', None)	('filename', 'Immunizations_Summary___PCP13802___2021-07-08T105000	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
12	Crystle_Price	('errorMessage', None)	('filename', 'Lab_Summary___PCP13802___2021-07-08T105000.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
13	Crystle_Price	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP13802___2021-07-08T1(	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
14	Crystle_Price	('errorMessage', None)	('filename', 'Encounter_Summary___PCP13802___2021-07-08T105000.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
15	Tammera_MacG	('errorMessage', None)	('filename', 'Immunizations_Summary___PCP115864___2021-07-12T15544	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
16	Tammera_MacG	('errorMessage', None)	('filename', 'Lab_Summary___PCP115864___2021-07-12T155440.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
17	Tammera_MacG	('errorMessage', None)	('filename', 'Encounter_Summary___PCP115864___2021-07-12T155440.xn	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
18	Tammera_MacG	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP115864___2021-07-12T1	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
19	Raymon_Willian	('errorMessage', None)	('filename', 'Lab_Summary___PCP65005___2021-07-10T101916.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
20	Raymon_Willian	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP65005___2021-07-10T1(	('ccdaDocumentType', None)	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
21	Raymon_Willian	('errorMessage', None)	('filename', 'Encounter_Summary___PCP65005___2021-07-10T101916.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
22	Raymon_Willian	('errorMessage', None)	('filename', 'Lab_Summary___PCP65005___2021-07-10T101916.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
23	Raymon_Willian	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP65005___2021-07-10T1(	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
24	Raymon_Willian	('errorMessage', None)	('filename', 'Encounter_Summary___PCP65005___2021-07-10T101916.xml	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
25	Raymon_Willian	('errorMessage', None)	('filename', 'Refill_Summary___PCP65005___2021-07-10T101916.xml')	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
26	Raymon_Willian	('errorMessage', None)	('filename', 'Immunizations_Summary___PCP65005___2021-07-10T101916	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
27	Ara_Prosacco	('errorMessage', None)	('filename', 'Lab_Summary___HALLMARK_HEALTH_MEDICAL_ASSOCIAT	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
28	Ara_Prosacco	('errorMessage', None)	('filename', 'Encounter_Summary___HALLMARK_HEALTH_MEDICAL_ASS	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
29	Ara_Prosacco	('errorMessage', None)	('filename', 'Immunizations_Summary___HALLMARK_HEALTH_MEDICAL_	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
30	Ara_Prosacco	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___HALLMARK_HEALTH_MEDI(	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
31	Octavio_Hyatt	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___CHICOPEE_EYECARE_PC_	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
32	Octavio_Hyatt	('errorMessage', None)	('filename', 'Encounter_Summary___CHICOPEE_EYECARE_PC___2021-0	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
33	Octavio_Hyatt	('errorMessage', None)	('filename', 'Lab_Summary___CHICOPEE_EYECARE_PC___2021-07-09T1	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
34	Octavio_Hyatt	('errorMessage', None)	('filename', 'Immunizations_Summary___CHICOPEE_EYECARE_PC___20:	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
35	Tequila_OReilly	('errorMessage', None)	('filename', 'Continuity_of_Care_Document___PCP118923___2021-07-13T1	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',
36	Tequila_OReilly	('errorMessage', None)	('filename', 'Encounter_Summary___PCP118923___2021-07-13T105356.xn	('ccdaDocumentType', 'CCD')	('results', {'finalGrade': 'C', 'finalNumericalGrade': 75, 'categoryList': [{'categoryName': 'Miscellaneous', 'categoryGrade': 'A+',

Final Output and Usage

- Multiple Point in Time documents generated from Synthea CCDAs
  - Encounter Summary
  - Immunizations Summary
  - Lab Summary
  - Refill Summary
- In-Document Synthetic Provider Notes
- API Interface to Access Information that mirrors the process of National Network APIs

Final Output and Usage

```
In [28]: import requests

         url = 'https://sandbox.scratch.particlehealth.com/api/v1/queries'
         headers = {"Content-Type": "application/json",
         'Authorization': jwt}
         data = {
                     "address_city": "West Bridgewater",
                     "address_lines": [
                       "126 McLaughlin Ferry"
                     ],
                     "address_state": "Massachusetts",
                     "date_of_birth": "1976-03-11",
                     "email": "Arlie@doe.com",
                     "family_name": "Rolfson",
                     "gender": "Male",
                     "given_name": "Arlie",
                     "npi": "1234",
                     "postal_code": "02324",
                     "purpose_of_use": "TREATMENT",
                     "ssn": "123-45-6789",
                     "telephone": "1-234-567-8910"
                   }
         r = requests.post(url, headers=headers, json=data)

         print(r.json())
         query_id = r.json()['id']
```


```
{'id': 'd5dbf549-ae18-4e51-9a04-765f9113062e', 'demographics': {'given_name': 'Arlie', 'family_name': 'Rolfson', 'dat
e_of_birth': '1976-03-11', 'gender': 'MALE', 'ssn': '123-45-6789', 'email': 'Arlie@doe.com', 'address_lines': ['126 M
cLaughlin Ferry'], 'address_state': 'MA', 'address_city': 'West Bridgewater', 'postal_code': '02324', 'hints': None,
'telephone': '(234) 567-8910', 'npi': '1234', 'purpose_of_use': 'TREATMENT'}, 'state': 'PENDING'}
```


```
In [29]: url = 'https://sandbox.scratch.particlehealth.com/api/v1/queries/' + query_id
         headers = {"Content-Type": "application/json",
         'Authorization': jwt}
         r = requests.get(url, headers=headers)

         print(r.json())
```


```
{'id': 'd5dbf549-ae18-4e51-9a04-765f9113062e', 'demographics': {'given_name': 'Arlie', 'family_name': 'Rolfson', 'dat
e_of_birth': '1976-03-11', 'gender': 'MALE', 'ssn': '123-45-6789', 'email': 'Arlie@doe.com', 'address_lines': ['126 M
cLaughlin Ferry'], 'address_state': 'MA', 'address_city': 'West Bridgewater', 'postal_code': '02324', 'hints': None,
'telephone': '(234) 567-8910', 'npi': '1234', 'purpose_of_use': 'TREATMENT'}, 'state': 'COMPLETE', 'files': [{'id':
'1462674f-aac9-4eb2-9f7c-1b9aff9771d9', 'title': 'Encounter_Summary5___PCP10647___2021-07-16T001720.xml', 'type': 'ap
plication/xml', 'url': '/api/v1/files/d5dbf549-ae18-4e51-9a04-765f9113062e/1462674f-aac9-4eb2-9f7c-1b9aff9771d9'},
{'id': '24192099-aa08-45e6-8905-2ab4dd65e917', 'title': 'Lab_Summary___PCP10647___2021-07-16T001720.xml', 'type': 'ap
plication/xml', 'url': '/api/v1/files/d5dbf549-ae18-4e51-9a04-765f9113062e/24192099-aa08-45e6-8905-2ab4dd65e917'},
{'id': '2bf8106a-1aba-48d6-afa8-9d3066509edb', 'title': 'Continuity_of_Care_Document6___PCP10647___2021-07-16T001720.
xml', 'type': 'application/xml', 'url': '/api/v1/files/d5dbf549-ae18-4e51-9a04-765f9113062e/2bf8106a-1aba-48d6-afa8-9
d3066509edb'}, {'id': '34a66d22-3350-4b0a-8447-3e0cdfea2663', 'title': 'Continuity_of_Care_Document___PCP10647___2021
-07-16T001720.xml', 'type': 'application/xml', 'url': '/api/v1/files/d5dbf549-ae18-4e51-9a04-765f9113062e/34a66d22-33
50-4b0a-8447-3e0cdfea2663'}, {'id': '6a468531-7fa2-4520-9b4a-f9592686f2a1', 'title': 'Encounter_Summary7___PCP10647__
_2021-07-16T001720.xml', 'type': 'application/xml', 'url': '/api/v1/files/d5dbf549-ae18-4e51-9a04-765f9113062e/6a4685
```


particle

Final Output and Usage

Encounter_Summary__RELIANT_MEDICAL_... Continuity_of_Care_Document__RELIANT_M...


```xml
    <section nullFlavor="NI">
      <templateId root="2.16.840.1.113883.10.20.22.2.22" extension="2015-08-01"/>
      <code code="46240-8" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="History of encounters"/>
      <title>Encounters</title>
      <text>
        <table border="1" width="100%">
          <thead>
            <tr>
              <th>Start</th>
              <th>Stop</th>
              <th>Description</th>
              <th>Code</th>
            </tr>
          </thead>
          <tbody>
            <tr xmlns="urn:hl7-org:v3" xmlns:sdtc="urn:hl7-org:sdtc" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
              <td>2019-04-08T04:10:10-04:00</td>
              <td>2019-04-09T04:10:10-04:00</td>
              <td ID="encounters-desc-27">Drug rehabilitation and detoxification</td>
              <td ID="encounters-code-27">http://snomed.info/sct 56876005</td>
            </tr>
          </tbody>
        </table>
      </text>
      <entry xmlns="urn:hl7-org:v3" xmlns:sdtc="urn:hl7-org:sdtc" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" typeCode="DRIV">
        <encounter classCode="ENC" moodCode="EVN">
          <templateId root="2.16.840.1.113883.10.20.22.4.49"/>
          <!-- Encounter activity template -->
          <id root="a6192761-0527-115c-6aae-00da9b7f9bdb"/>
          <code code="56876005" codeSystem="2.16.840.1.113883.6.96" displayName="Drug rehabilitation and detoxification">
            <originalText>
              <reference value="#encounters-desc-27"/>
            </originalText>
          </code>
          <text>
            <reference value="#encounters-desc-27"/>
          </text>
          <effectiveTime>
            <low value="20190408041010"/>
            <high value="20190409041010"/>
          </effectiveTime>
        </encounter>
      </entry>
    </section>
  </component>
  <component>
    <!-- Allergies: -->
    <section nullFlavor="NI">
      <templateId root="2.16.840.1.113883.10.20.22.2.6.1" extension="2015-08-01"/>
      <code code="48765-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Allergy List"/>
```

Final Output and Usage

```
            <item>
              <paragraph><br/>2019-04-08
<br/><br/> # Chief Complaint
<br/> No complaints.
<br/>
<br/> # History of Present Illness
<br/> Austin578 is a 46 year-old non-hispanic white male.
<br/>
<br/> # Social History
<br/> Patient is married. Patient has a documented history of opioid addiction. Patient is an active smoker and is an alcoholic. Patient identifies as heterosexual.
<br/>
<br/> Patient comes from a low socioeconomic background. Patient has a high school education. Patient currently has UnitedHealthcare.
<br/>
<br/> # Allergies
<br/> No Known Allergies.
<br/>
<br/> # Medications
<br/> No Active Medications.
<br/>
<br/> # Assessment and Plan
<br/>
<br/>
<br/> ## Plan
<br/>
<br/>
<br/>
<br/>
<br/><br/>
                                                              </paragraph>
            </item>
```

Discussion

- Progression from original Synthea (Single CCDA)
  - Why realistic data helps developers innovate
- Validation Component
- Delivering a specific sub-population of patients
  - Impact on innovation and research (Opioid or Complex-Care)
- Synthea solution publically available on our GitHub Repository:
  - https://github.com/ParticleHealth/Particle-Health-Sandbox-Environment
- Leveraged Open-Source tools and libraries
- Visit our website to use our sandbox environment
  - Pre-loaded with Synthea Data modified with our solution
  - https://www.particlehealth.com

particle

Lessons Learned and Future Work

- Lessons Learned
  - Synthea is a powerful tool for generating synthetic CCDA and equivalent FHIR data
  - There are many potential opportunities to develop on top of Synthea to improve upon it and the use cases it can deliver
  - Real Clinical Information is highly variable and comes in different shapes and sizes
  - Quality development environments are important for enabling innovation
  - Policy greatly impacts the future direction of health information technology
- Future Work
  - Expanding Point In Time Document Types
    - ie. Discharge summaries
  - Generating other types of data seen in clinical practice
    - ie. Allergies

Thank you!

Questions?

go@particlehealth.com

particle

Empirical inference of Underlying Condition Probabilities Using Synthea-Generated Synthetic Health Data

Team TeMa

Dr. Michael D. Teter
miketeter@yahoo.com

Dr. Christopher E. Marks
cemarks@alum.mit.edu

Challenge Category: II (Novel Uses of Synthea Generated Data)

Background

Problem Motivation

- Simulation is often used to investigate complicated phenomena for which analytic determination of outcome probabilities is intractable.
- Synthea is built in a way that makes it well-suited for this purpose.
  - It inputs conditional probabilities that can be validated.
  - Its outputs are the result of tailorable combinations of these input probabilities.

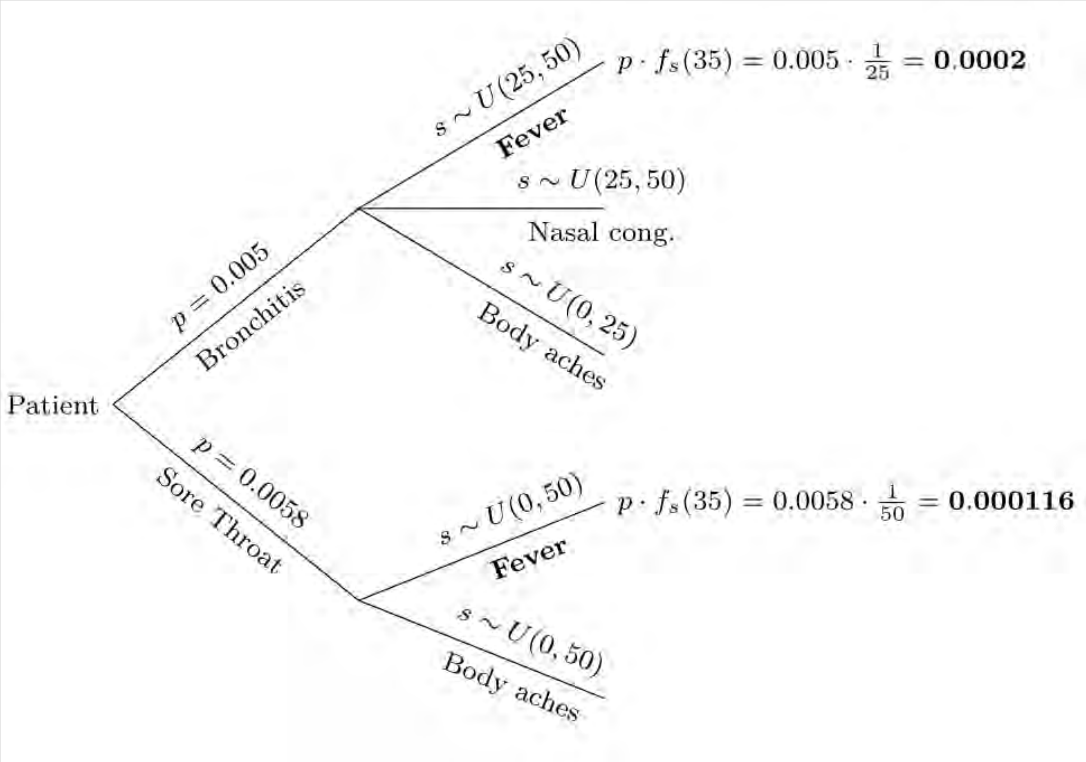
Our Task

- Use Synthea-generated data to investigate relationships between a patient's pathology and a given set of symptoms and severities.

Our Methods (1 of 2)

Empirical Bayes

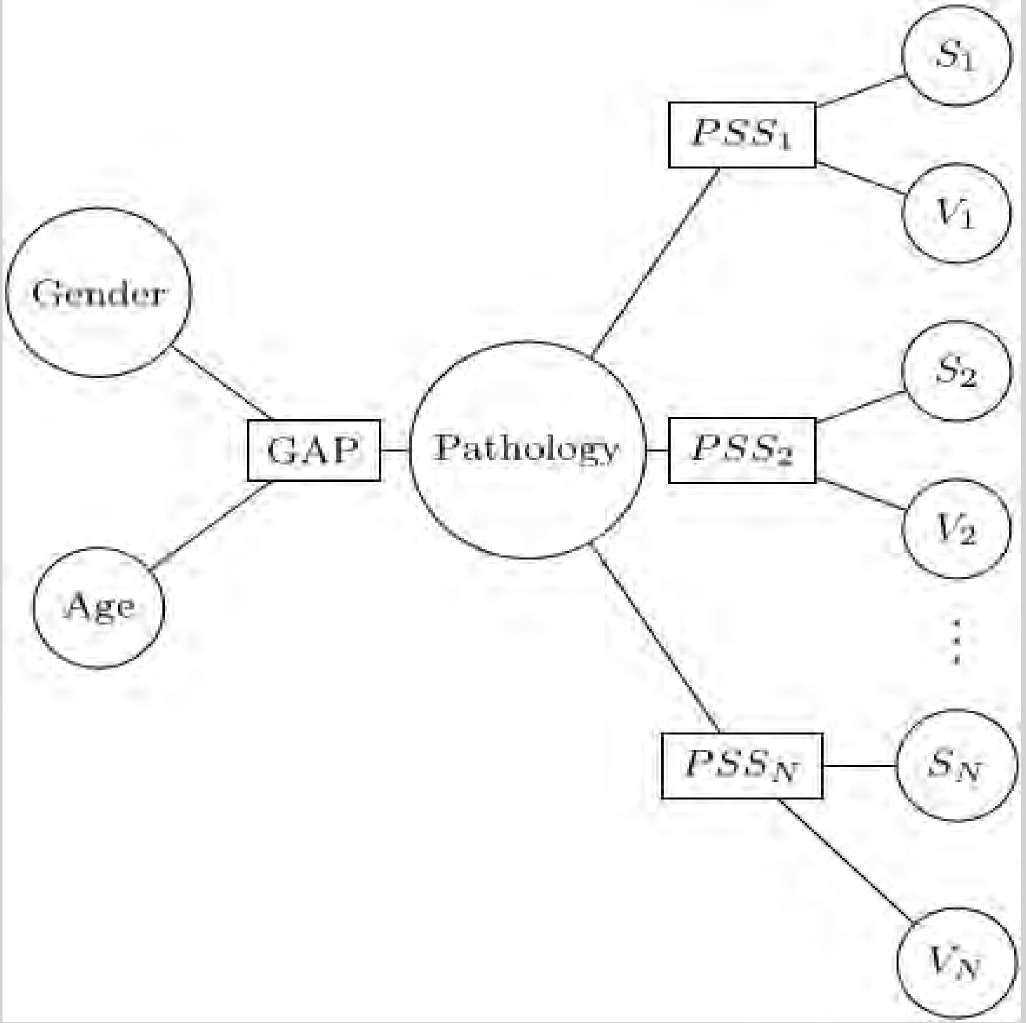
- Canonical Bayesian analysis
- Using empirical distributions in Synthea data, no need to enumerate a complicated tree.



Our Methods (2 of 2)

Bayesian Network

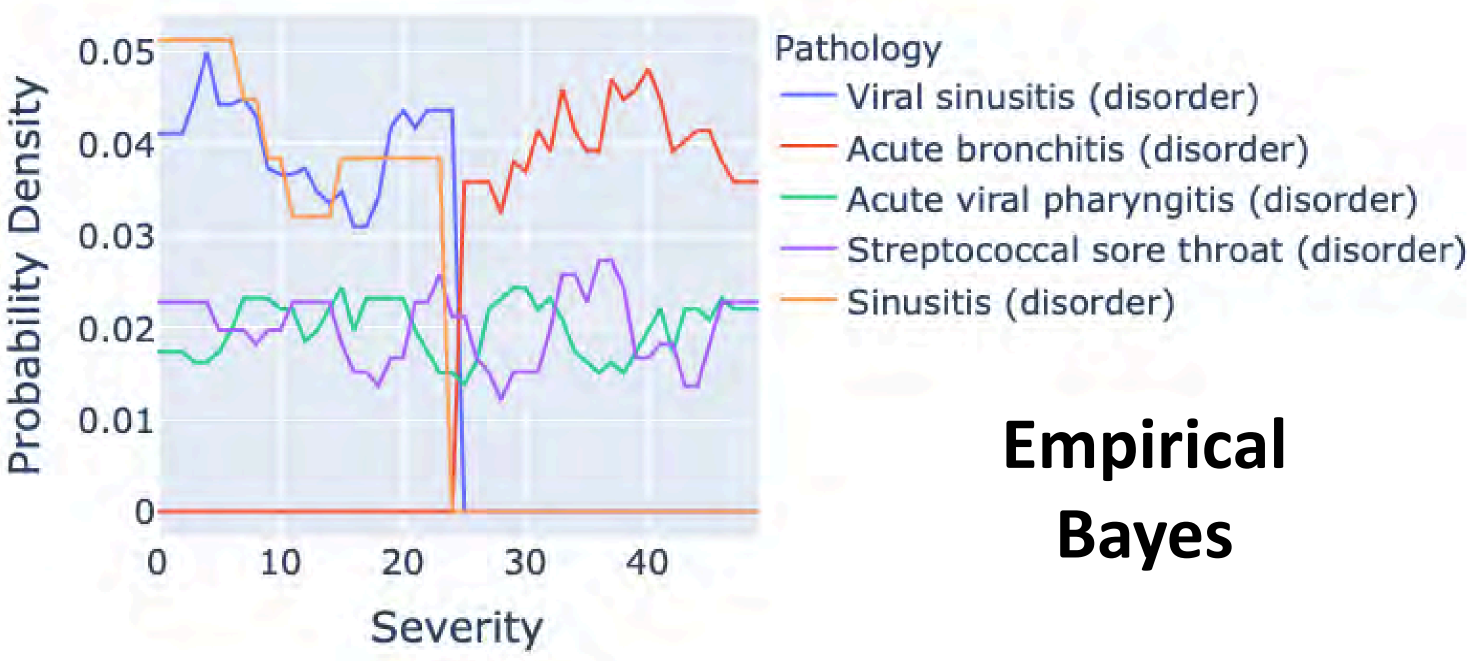
- Graph-based machine learning method.
- We decide which variables are related.
- More versatile than the strictly empirical model.



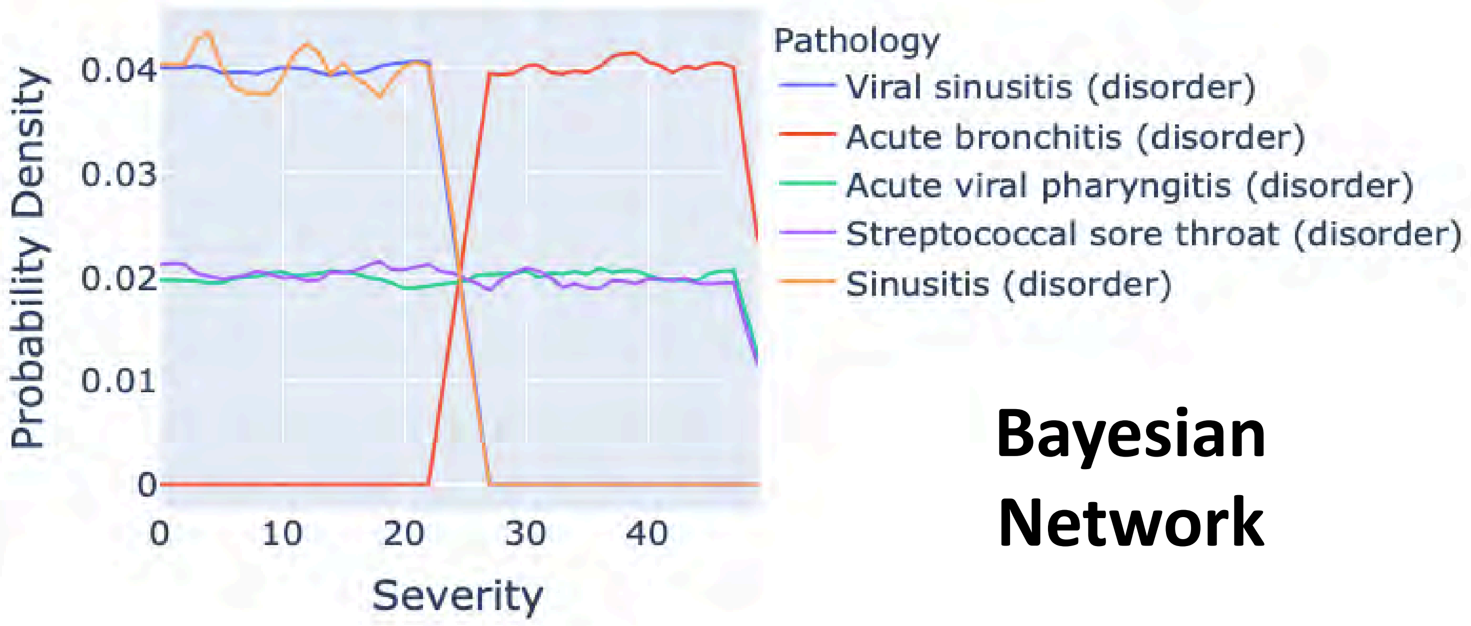
Example Outputs

<u>Patient</u>

- 5 years old
- Female
- Fever



Empirical Bayes



Bayesian Network

Validation

- Internal Validation through code testing
- External Validation through comparison with existing tools.
- A useful method for validating Synthea!

External validation: comparison with WebMD

WebMD	Synthea Bayes
Bacterial Pneumonia	Viral sinusitis (disorder)
Middle Ear Infection	Acute bronchitis (disorder)
Viral Pneumonia	Acute viral pharyngitis (disorder)
Influenza (Flu) Child	Streptococcal sore throat (disorder)
Strep Throat	Sinusitis (disorder)

Summary & Next Steps

- Our approach can be extended to account for demographics, encounter types, patient location, patient history, etc.
- Compare results to known distributions as a way of **validating Synthea.**
- Identify areas where Synthea can be improved.
  - Standardization!
- Look for real-world applications.
  - Rare pathologies?

SPATIOTEMPORAL BIG DATA ANALYSIS OF THE OPIOID EPIDEMIC IN ILLINOIS

- Office of the National Coordinator for Health Information Technology (ONC) Synthetic Health Data Challenge
- Category II: Novel Uses of Synthea™ Generated Synthetic Data

- Arash Jalali, MPH, MSHI
- Sean Huang, MD
- Karl Kochendorfer, MD, FAAFP

UI HEALTH

- Comprehensive care, education, and research to train health care leaders and foster healthy communities in Illinois and beyond.
- 465 bed tertiary care hospital, 21 outpatient clinics, 11 federally qualified Mile Square Health Center locations
- Campuses in Chicago, Peoria, Quad Cities, Rockford, Springfield, and Urbana
- 7 Health Science colleges: Applied Health Sciences, Dentistry, Medicine, Nursing, Pharmacy, School of Public Health, Jane Addams College of Social Work

1 in 3	Only	7	$243M	1 in 4	Over 8,000
IL Physicians trained at UIC	Public, Research 1 University in Chicago	Health Sciences Colleges at UIC	in total annual health sciences research funding	IL Social Workers Trained at UIC	IL Nurses trained at UIC

Blair Turner, Wilnise Jasmin, Isabel Chung, Ponni Arunkumar, Mark Kiely, Steven Aks, Nikhil Prachand, Allison Arwady. Opioid Overdose Surveillance Report—Chicago 2019. City of Chicago, March 2021.

OPIOID CRISIS: CHICAGO STATISTICS

- 2019: 855 people died from opioid overdoses (from 793 previous year)
- 2018 -> 2019: Opioid-related overdose death rate increased by 10.1%
- CFD EMS team responded to average 29 responses per day (increase in 25.4%)

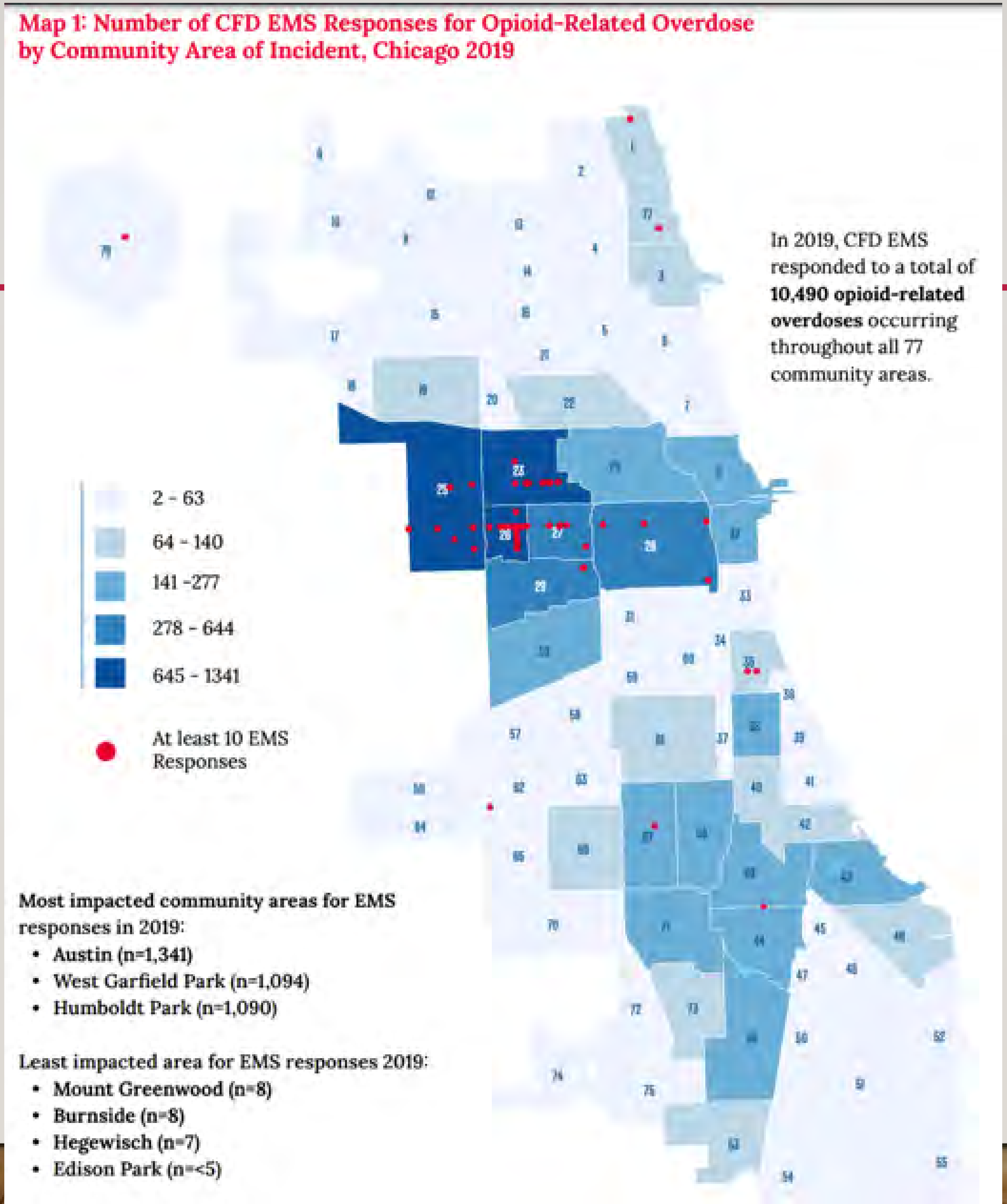
Figure 1. Opioid-Related EMS Runs by Quarter, 2013-2019

Data Source: Chicago Fire Department

OPIOID CRISIS: CHICAGO STATISTICS

- Men
- Aged 45-64 years old
- Non-Hispanic African-Americans
- Use of combination of other opioids and illicit drugs
  - Cocaine
- High economic hardship
  - Education
  - Income levels, Unemployment
  - Crowded Housing

Blair Turner, Wilnise Jasmin, Isabel Chung, Ponni Arunkumar, Mark Kiely, Steven Aks, Nikhil Prachand, Allison Arwady. Opioid Overdose Surveillance Report—Chicago 2019. City of Chicago, March 2021.



WORKPLACE INJURIES

- Especially with prescription pain relievers
- NIOSH, John Howard:
  - Potential for addiction may be preceded by injuries that happen in the workplace, with the consequences affecting both an individual’s working life as well as their home life
- Exposure to opiate powders -> hazardous environment for healthcare workers

AMA 2021 OVERDOSE EPIDEMIC REPORT

2021 OVERDOSE EPIDEMIC REPORT

Physicians' actions to help end the nation's drug-related overdose and death epidemic —and what still needs to be done.

"develop and implement systems to collect timely, adequate and standardized data to identify at-risk populations, and implement public health interventions that directly address removing structural and racial inequities."

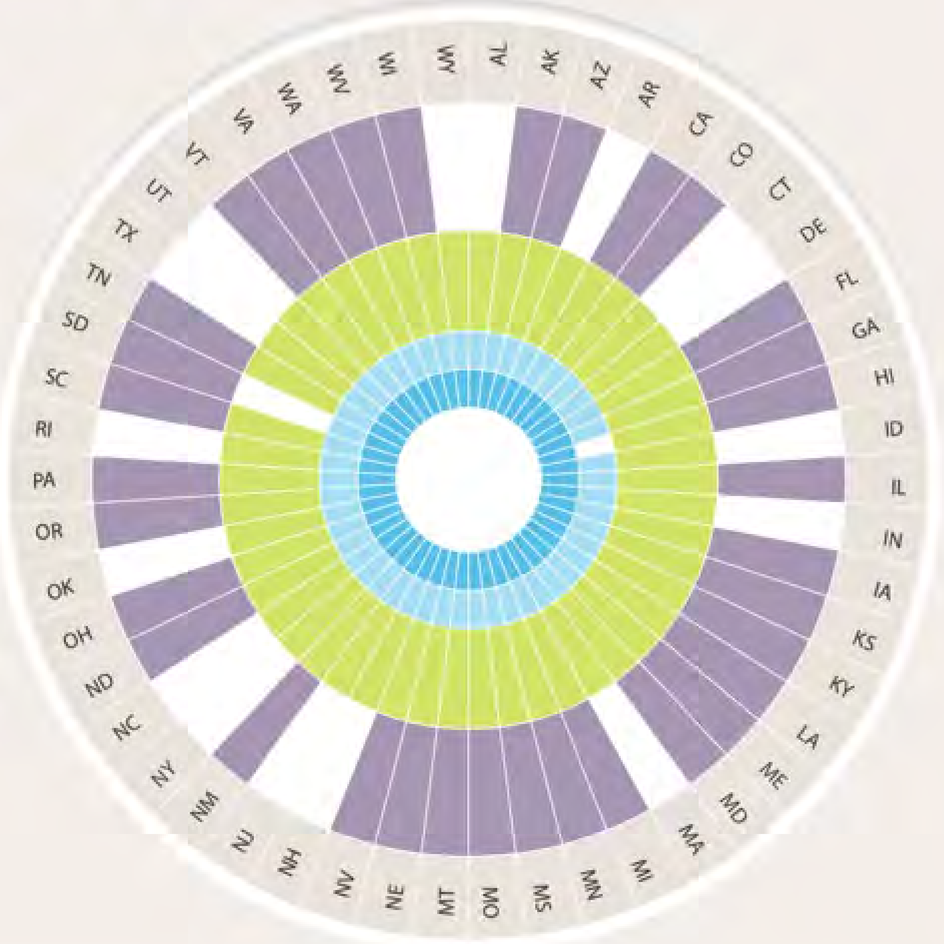
AMA (2021). 2021 OVERDOSE EPIDEMIC REPORT: Physicians' actions to help end the nation's drug-related overdose and death epidemic—and what still needs to be done. Retrieved from https://end-overdose-epidemic.org/wp-content/uploads/2021/09/AMA-2021-Overdose-Epidemic-Report_92021.pdf

While data is critical to improving outcomes, current data is:

…incomplete
…not standardized for comparison
…not timely
…widely variable from location to location

Difficulties remain in accessing high quality, timely, comprehensive and standardized data. While metrics are generally available for drug-related overdoses, data for non-fatal overdoses and other key indicators are not widely collected or standardized across states and communities. These data gaps greatly hinder understanding of local situations and advancing prevention, treatment and harm reduction efforts.

Inadequate data collection prevents effective public health interventions to reduce overdose and death.



Data categories

- Prescriptions
- PDMP
- Fatal overdoses
- Non-fatal overdoses
- No data

AMA (2021). 2021 OVERDOSE EPIDEMIC REPORT: Physicians' actions to help end the nation's drug-related overdose and death epidemic—and what still needs to be done. Retrieved from https://end-overdose-epidemic.org/wp-content/uploads/2021/09/AMA-2021-Overdose-Epidemic-Report_92021.pdf

Final Report of the Health Information Technology Advisory Committee's Public Health Data Systems Task Force 2021

Submitted to the Office of the National Coordinator for Health IT on July 14, 2021

Public Health Data Systems Task Force (2021). Final Report of the Health Information Technology Advisory Committee's Public Health Data Systems Task Force 2021. Retrieved from https://www.healthit.gov/sites/default/files/page/2021-08/2021-07-14_PHDS_TF_2021_HITAC%20Recommendations%20Report_Signed_508_0.pdf

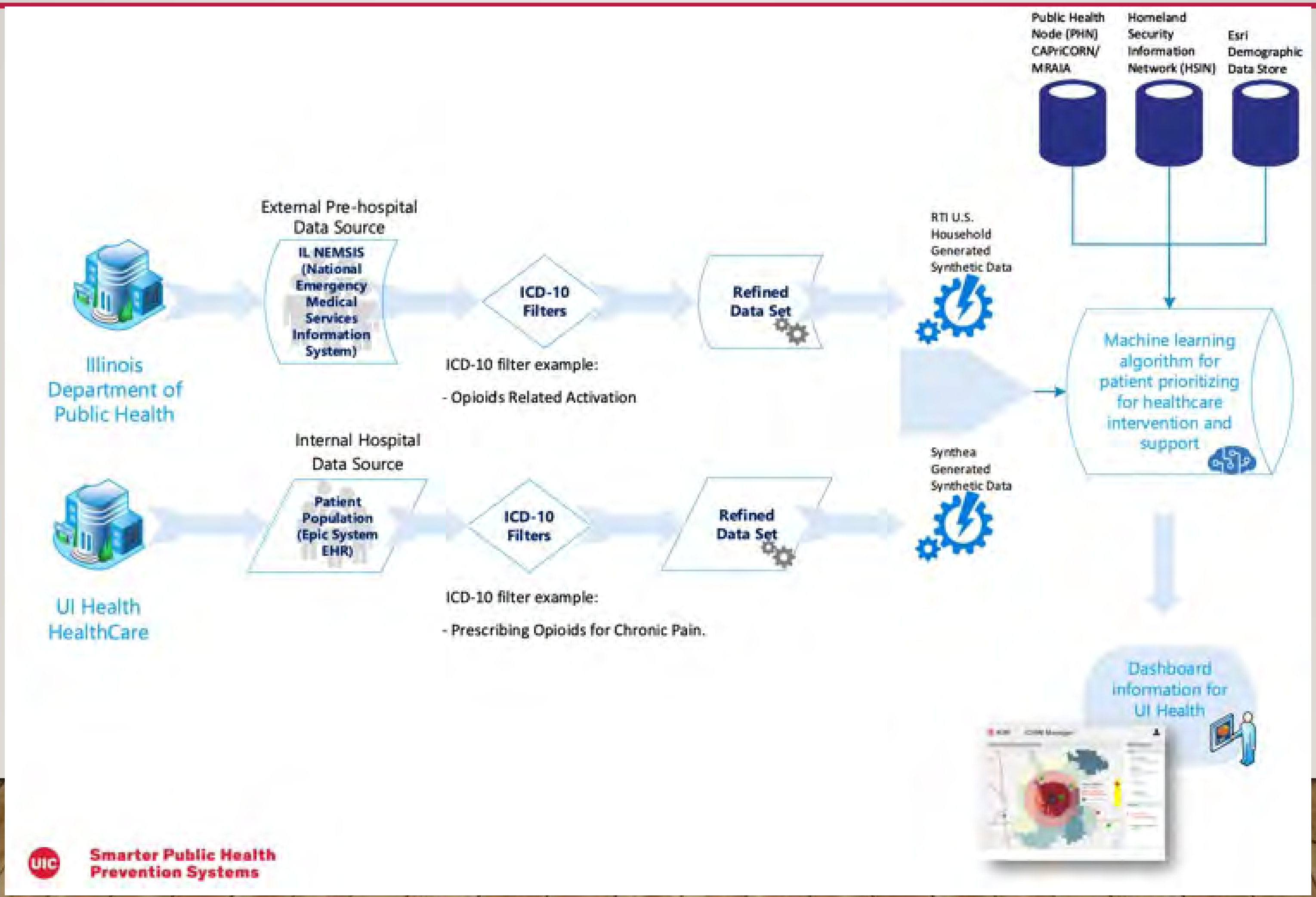
HITAC RECOMMENDATIONS ON PUBLIC HEALTH DATA SYSTEMS

- Improving interoperability
  - NEMSIS
  - Cloud computing
- Synthetic syndromic surveillance to assist "traditionally under-resourced areas to support creation of a public health system able to support health equity and health disparities"
  - Help facilitate interoperability, geolocation
  - Merging with census and other SDOH data
- Explore traditional and non-traditional data sources to assist with early identification of early clusters/outbreaks of disease incidence

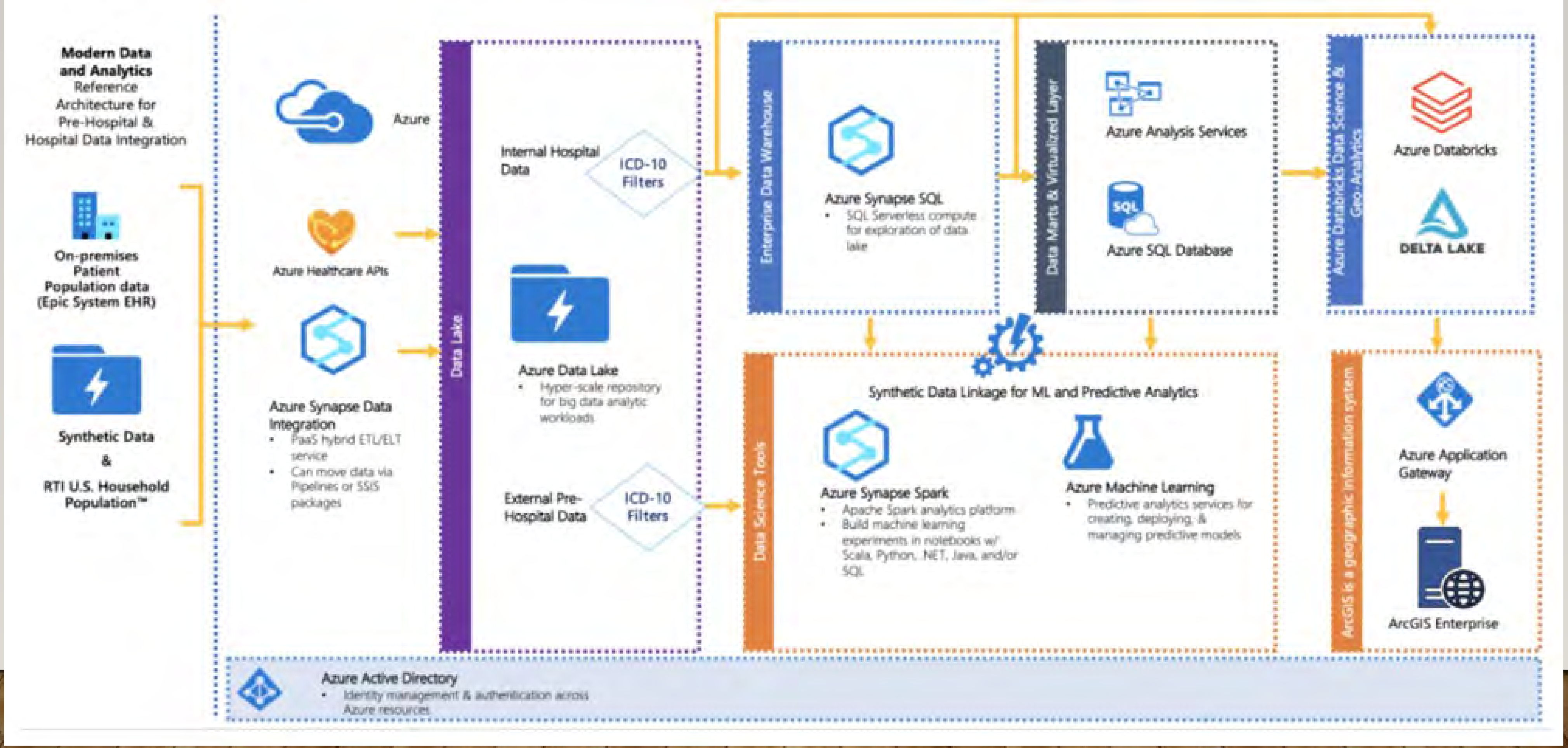
SPATIOTEMPORAL BIG DATA ANALYSIS OF OPIOID EPIDEMIC IN ILLINOIS

- Spatiotemporal distribution of EMS 911 calls and ambulance dispatches related to drug overdoses
- Obtain Chicago EMS data – store in Azure Data Lake. Scripts execute over Azure Cloud
- Opiate cases identified and geospatial information extracted
- Analysis on ArcGIS enterprise
- Enrich opiate cases with Esri demographic and census data of surrounding neighborhoods
- Machine learning to understand features to predict opiate use

SMARTER PUBLIC HEALTH PREVENTION SYSTEM (SPHPS) SYNTHETIC DATA INTEGRATION OF PRE-HOSPITAL TO HOSPITAL DATA



AZURE MODERN ANALYTICS ARCHITECTURE FOR SYNTHETIC SYNDROMIC SURVEILLANCE



CHICAGO EMS DATA

- National Emergency Medical Services Information System (NEMSIS)
- Identify opiate cases on SQL Server
- Provider's Primary Impression
- Primary Symptom

NEMSIS
Powered by NHTSA's Office of EMS

https://nemsis.org/

eSituation.11 - Provider's Primary Impression

Definition

The EMS personnel's impression of the patient's primary problem or most significant condition which led to the management given to the patient (treatments, medications, or procedures).

National Element	Yes	Pertinent Negatives (PN)	No
State Element	Yes	NOT Values	Yes
Version 2 Element	E09_15	Is Nillable	Yes
Usage	Required	Recurrence	1 : 1

Associated Performance Measure Initiatives

Airway Cardiac Arrest Pediatric STEMI Stroke Trauma

Attributes

NOT Values (NV)
7701001 - Not Applicable 7701003 - Not Recorded

Constraints

Pattern
```
(R[0-6][0-9](\.[0-9]{1,4})?|(R73\.9)|(R99))|([A-QSTZ][0-9][0-9A-Z])(\.[0-9A-Z]{1,4})?)
```

Data Element Comment

Code list is represented in ICD-10-CM. Reference the NEMSIS Suggested Lists at:
https://nemsis.org/technical-resources/version-3/version-3-resources/

ICD-10-CM
Website - http://uts.nlm.nih.gov
Product - UMLS Metathesaurus

eSituation.09 - Primary Symptom

Definition

The primary sign and symptom present in the patient or observed by EMS personnel.

National Element	Yes	Pertinent Negatives (PN)	No
State Element	Yes	NOT Values	Yes
Version 2 Element	E09_13	Is Nillable	Yes
Usage	Required	Recurrence	1 : 1

Associated Performance Measure Initiatives

Airway Cardiac Arrest Pediatric STEMI Stroke Trauma

Attributes

NOT Values (NV)
7701001 - Not Applicable 7701003 - Not Recorded

Constraints

Pattern
```
(R[0-6][0-9](\.[0-9]{1,4})?|(R73\.9)|(R99))|([A-QSTZ][0-9][0-9A-Z])(\.[0-9A-Z]{1,4})?)
```

Data Element Comment

eSituation.02 (Possible Injury), eSituation.09 (Primary Symptom), eSituation.07 (Chief Complaint Anatomic Location), and eSituation.08 (Chief Complaint Organ System) are grouped together to form the EMS Reason for Encounter.

Code list is represented in ICD-10-CM Diagnosis Codes. Reference the NEMSIS Suggested Lists at:
https://nemsis.org/technical-resources/version-3/version-3-resources/

ICD-10-CM
Website - http://uts.nlm.nih.gov
Product - UMLS Metathesaurus

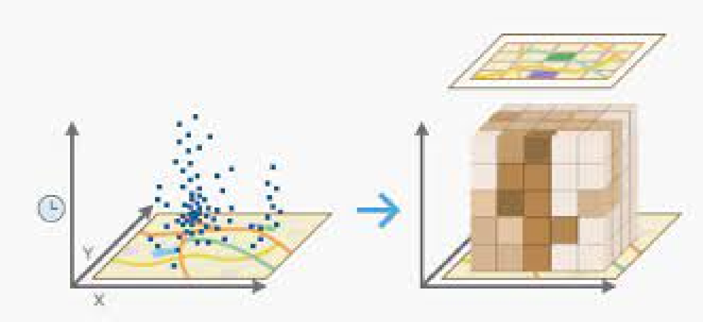
CHICAGO EMS DATA

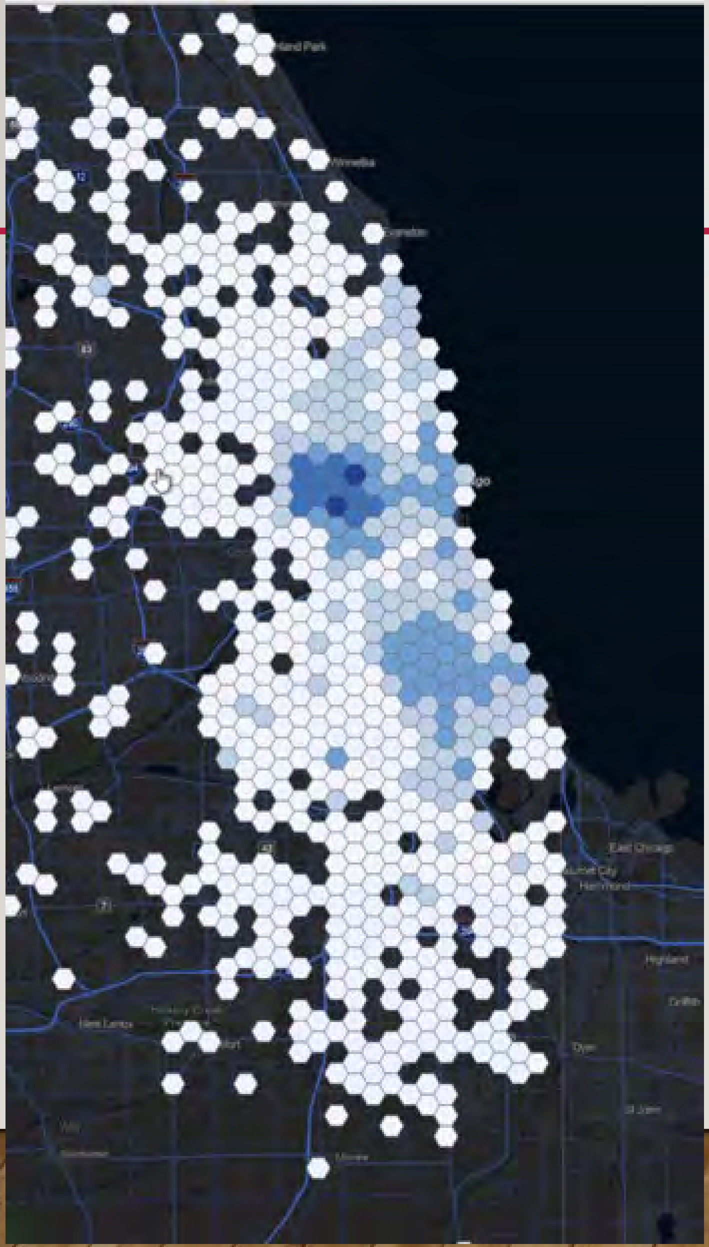
- Provider's Primary Impression and/or Primary Symptom = 'Opioid related disorders', 'Opioid use, unspecified'. Or ICD-10 codes in F11, T40 categories or Z79.891

ProvidersPrimaryImp...	PrimarySymptom	WorkRelatedIllnessInj...
Opioid related disord...	Gait-Limp/Difficulty...	No
Opioid related disord...	Vomiting	No
Opioid related disord...	Apnea	No
Opioid related disord...	Abnormal breathing...	No
F11	Altered mental status	No
Opioid related disord...	Anxiety or Worries	No
Opioid use, unspecifi...	Abnormal breathing...	No
Opioid use, unspecifi...	Slowness/poor respo...	No
Opioid use, unspecifi...	Altered mental status	No
Opioid use, unspecifi...	Stupor or Semicoma	No
Opioid related disord...	Altered mental status	No
Opioid use, unspecifi...	Altered mental status	No
Opioid use, unspecifi...	Altered mental status	No
Opioid use, unspecifi...	Altered mental status	No
Opioid related disord...	Abnormal breathing...	No
Opioid use, unspecifi...	Slowness/poor respo...	No
F11	Not Recorded	No
F11	Coma, unspecified	No
Opioid use, unspecifi...	Slowness/poor respo...	No
Opioid use, unspecifi...	Coma, unspecified	No

ARCGIS ENTERPRISE

- Geospatial data management, data visualization, analytics, geospatial forecasting of opiates and overdoses
- Space time cubes for all overdoses and opioid cases
- Create predictive models using time series forecasting tools
- Curve fit forecast models
  - Forecast future future values using curve fitting
- Exponential smoothing forecast model – predicts values by decomposing time series at each location into seasonal and trend components





Winnetka
Evanston
East Chicago
Hammond
Highland
Griffith
Dyer

Pop-up
OPIOID_CurveFitForecast (1)
10219
OPIOID_CurveFitForecast - 10219
Forecast Method: Linear
COUNT
Feb 2019
Apr 2019
Jun 2019
Aug 2019
Oct 2019
Dec 2019
Feb 2020
Apr 2020
Original Value
Fitted Value
Forecasted Value
1 of 1
1,141,698.92E 1,902,141.90N ftUS

Pop-up
OVERDOSE_CurveFitForecast (1)
10889
OVERDOSE_CurveFitForecast - 10889
Forecast Method: S-shaped (Gompertz)
COUNT
Feb 2019
Apr 2019
Jun 2019
Aug 2019
Oct 2019
Dec 2019
Feb 2020
Apr 2020
Original Value
Fitted Value
Forecasted Value
1 of 1
1,148,756.22E 1,902,477.87N ftUS

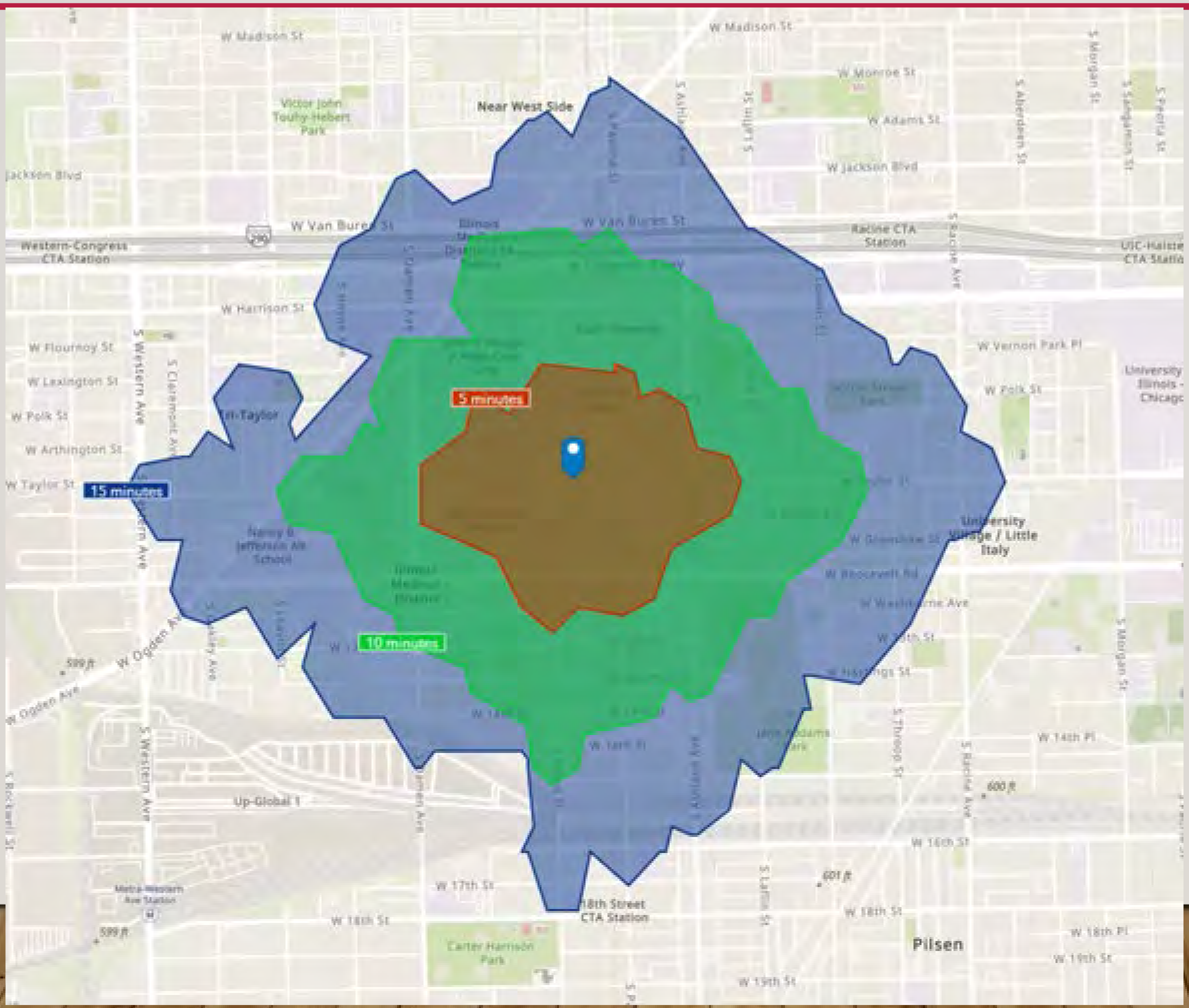
lethod	Forecast Equation
	Xt = a*t+b; a=-0.009804, b=0.137255
al	Xt = k + a*exp(b*t); k=0.084258, a=-0.000747, b=0.316302
	Xt = a*t^2+b*t+c; a=0.001548, b=-0.039474, c=0.238390
	Xt = a*t^2+b*t+c; a=-0.004902, b=0.083333, c=-0.117647
	Xt = a*t^2+b*t+c; a=-0.002580, b=0.046182, c=-0.083591
	Xt = k+a*exp(-b*exp(-c*t)); k=0.000000, a=0.224779, b=52.858909, c=0.6076.
	Xt = a*t^2+b*t+c; a=-0.002580, b=0.046182, c=-0.083591
	Xt = a*t^2+b*t+c; a=-0.001935, b=0.038313, c=-0.077399
	Xt = a*t+b; a=-0.007353, b=0.117647
	Xt = a*t+b; a=0.017157, b=-0.078431
	Xt = a*t^2+b*t+c; a=0.008256, b=-0.073271, c=0.094943
	Xt = k+a*exp(-b*exp(-c*t)); k=0.000010, a=0.557301, b=91.371438, c=0.3877.
	Xt = a*t+b; a=-0.007353, b=0.117647
	Xt = a*t+b; a=0.014706, b=-0.058824
	Xt = a*t^2+b*t+c; a=-0.006579, b=0.107714, c=-0.106295
	Xt = a*t^2+b*t+c; a=-0.005934, b=0.099845, c=-0.158927
	Xt = k+a*exp(-b*exp(-c*t)); k=0.000000, a=0.194351, b=108.978494, c=0.859.
	Xt = a*t^2+b*t+c; a=0.004902, b=-0.063725, c=0.254902
al	Xt = k + a*exp(b*t); k=0.084258, a=-0.000747, b=0.316302
	Xt = a*t+b; a=0.009804, b=-0.019608

EXPONENTIAL SMOOTHING FORECAST MODEL

- Geospatial information -> upload into ESRI
- Demographic data store that allows for geocoding and automated enrichment

- USA 2020 demographic data
- USA 2010 Census Demographic Data
- USA 2014/2018 American Community Survey (ACS) Demographic Data,
- USA 2020 Consumer Expenditure data
- USA 2020 Tapestry Segmentation Data.
- Its geography information is updated to 2020/2021.

5-MINUTE WALK TIME



Smarter Public Health Prevention Systems

2010 Census Profile

1740 W Taylor St, Chicago, Illinois, 60612 5
1740 W Taylor St, Chicago, Illinois, 60612
Walk Time: 5 minute radius

Prepared by Esri

	2000	2010	2000-2010 Annual Rate
Population	2,564	654	-12.77%
Households	72	60	-1.81%
Housing Units	85	67	-2.35%

Population by Race	**Number**	**Percent**
Total	654	100.0%
Population Reporting One Race	622	95.1%
White	251	38.4%
Black	145	22.2%
American Indian	2	0.3%
Asian	195	29.8%
Pacific Islander	1	0.2%
Some Other Race	28	4.3%
Population Reporting Two or More Races	32	4.9%
Total Hispanic Population	79	12.1%
Population by Sex		
Male	302	46.2%
Female	352	53.8%
Population by Age		
Total	655	100.0%
Age 0 - 4	9	1.4%
Age 5 - 9	12	1.8%
Age 10 - 14	11	1.7%
Age 15 - 19	181	27.6%
Age 20 - 24	232	35.4%
Age 25 - 29	94	14.4%
Age 30 - 34	26	4.0%
Age 35 - 39	22	3.4%
Age 40 - 44	13	2.0%
Age 45 - 49	7	1.1%
Age 50 - 54	8	1.2%
Age 55 - 59	8	1.2%
Age 60 - 64	4	0.6%
Age 65 - 69	7	1.1%
Age 70 - 74	2	0.3%
Age 75 - 79	7	1.1%
Age 80 - 84	5	0.8%
Age 85+	6	0.9%
Age 18+	613	93.7%
Age 65+	27	4.1%

2010 Census Profile

1740 W Taylor St, Chicago, Illinois, 60612 5
1740 W Taylor St, Chicago, Illinois, 60612
Walk Time: 5 minute radius

Prepared by Esri

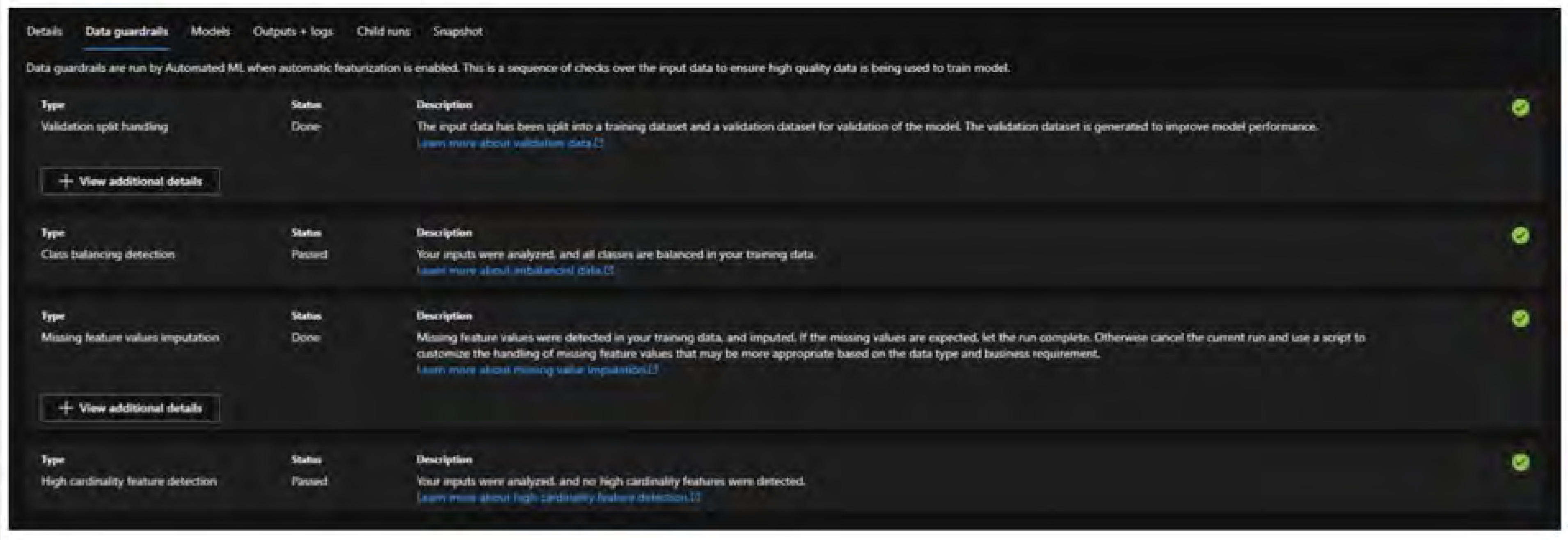
Households by Type		
Total	60	100.0%
Households with 1 Person	20	33.3%
Households with 2+ People	40	66.7%
Family Households	27	45.0%
Husband-wife Families	9	15.0%
With Own Children	4	6.7%
Other Family (No Spouse Present)	18	30.0%
With Own Children	10	16.7%
Nonfamily Households	13	21.7%
All Households with Children	16	26.7%
Multigenerational Households	2	3.3%
Unmarried Partner Households	4	6.7%
Male-female	4	6.7%
Same-sex	0	0.0%
Average Household Size	4.48	
Family Households by Size		
Total	28	100.0%
2 People	11	39.3%
3 People	7	25.0%
4 People	5	17.9%
5 People	3	10.7%
6 People	1	3.6%
7+ People	1	3.6%
Average Family Size	4.26	
Nonfamily Households by Size		
Total	33	100.0%
1 Person	20	60.6%
2 People	6	18.2%
3 People	4	12.1%
4 People	2	6.1%
5 People	1	3.0%
6 People	0	0.0%
7+ People	0	0.0%
Average Nonfamily Size	4.48	
Population by Relationship and Household Type		
Total	654	100.0%
In Households	269	41.1%
In Family Households	120	18.3%
Householder	38	5.8%
Spouse	16	2.4%
Child	49	7.5%
Other relative	13	2.0%
Nonrelative	5	0.8%
In Nonfamily Households	148	22.6%
In Group Quarters	385	58.9%
Institutionalized Population	0	0.0%
Noninstitutionalized Population	385	58.9%

Azure Machine Learning

- After data cleaning, variables entered into Azure Machine Learning to run predictions
- Many AutoML experiments created
- Target: opioid activations by EMS. Total overdose case activations by EMS
- Classification machine learning models

Azure Machine Learning

- Variables cleaned and filtered. Removed based on overfitting and imbalanced data

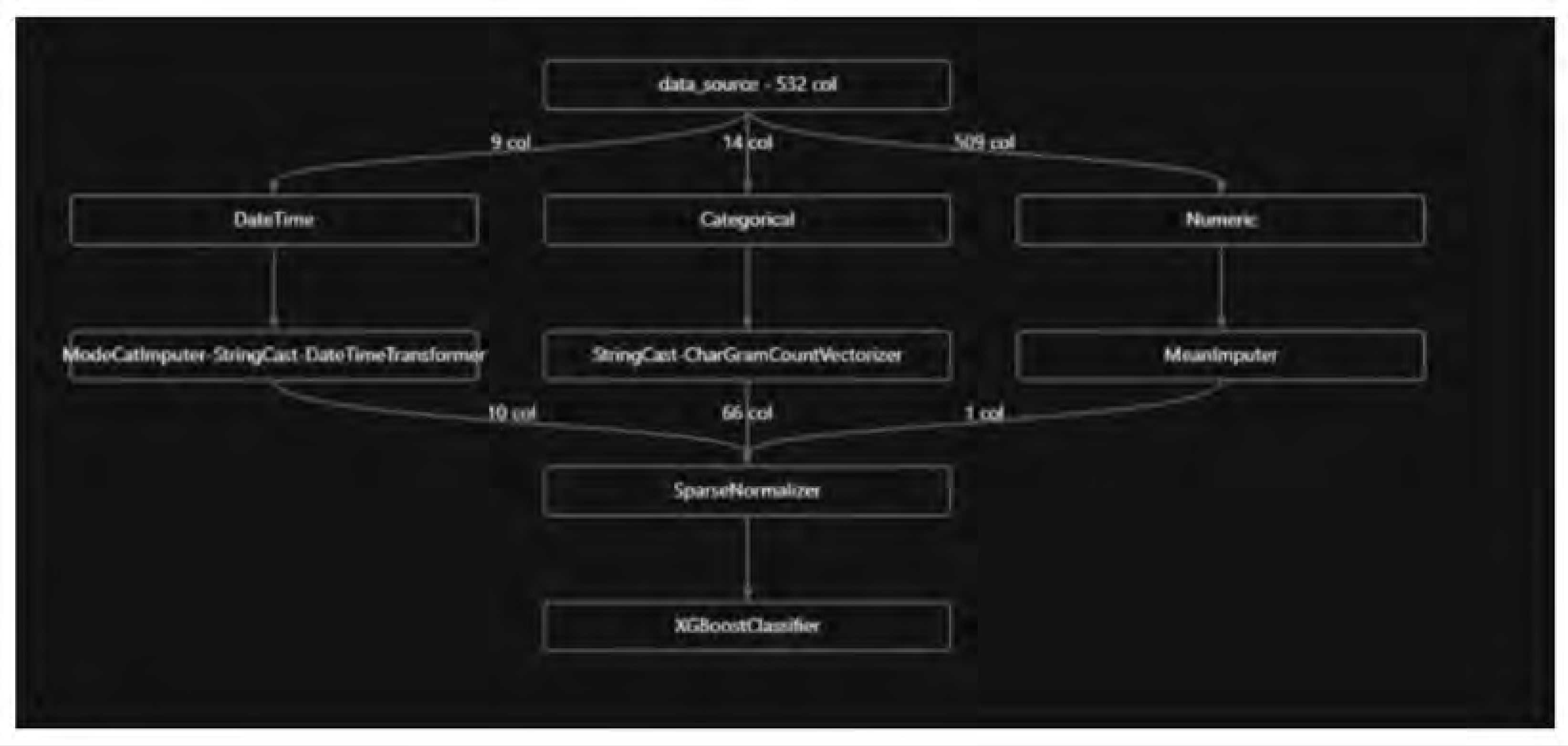


RESULTS

- Second to last runs

Algorithm name	Explained	Accuracy ↓
MaxAbsScaler, LogisticRegression		0.82468
MaxAbsScaler, LightGBM		0.82405
MaxAbsScaler, LightGBM		0.82358
SparseNormalizer, XGBoostClassifier		0.82296
StandardScalerWrapper, LightGBM		0.82171
SparseNormalizer, XGBoostClassifier		0.82140
MaxAbsScaler, LightGBM		0.82124
MaxAbsScaler, GradientBoosting		0.82093
MaxAbsScaler, GradientBoosting		0.82093
StandardScalerWrapper, RandomForest		0.81813
SparseNormalizer, LightGBM		0.81563

SPARSENORMALIZER, XGBOOST CLASSIFIER



RESULTS

- Final runs

Algorithm name	Explained	Accuracy ↓
VotingEnsemble	View explanation	0.84012
StackEnsemble		0.83637
SparseNormalizer, XGBoostClassifier		0.83388
SparseNormalizer, XGBoostClassifier		0.83388
SparseNormalizer, XGBoostClassifier		0.83388
SparseNormalizer, XGBoostClassifier		0.83341
SparseNormalizer, XGBoostClassifier		0.83326
SparseNormalizer, XGBoostClassifier		0.83326
SparseNormalizer, XGBoostClassifier		0.83326
MaxAbsScaler, LightGBM		0.83232
SparseNormalizer, XGBoostClassifier		0.83216

CONSIDERATIONS WITH SYNTHEA

```
C:\Synthea\synthea>.\run_synthea.bat -m "onc_opioids" -p 255000 Illinois Chicago
```

```
254984 -- Elva122 Langworth352 (25 y/o F) Chicago, Illinois
254982 -- Drucilla444 Paucek755 (29 y/o F) Chicago, Illinois
254991 -- Coleen678 Sauer652 (17 y/o F) Chicago, Illinois
254988 -- Tracey100 Gottlieb798 (30 y/o M) Chicago, Illinois
254985 -- Luigi346 Schmeler639 (39 y/o M) Chicago, Illinois
254983 -- Kevin729 Hahn503 (54 y/o M) Chicago, Illinois
254989 -- Long300 Hammes673 (41 y/o M) Chicago, Illinois DECEASED
```

		CITY	STATE	COUNTY	ZIP	LAT	LON	HE
9	rchard Apt...	Chicago	Illinois	DuPage County	60018	41.681	-87.616	
10		Chicago	Illinois	DuPage County	60646	41.774	-87.806	
11	Suite 29	Chicago	Illinois	DuPage County	60616	41.904	-87.779	
12	arade	Chicago	Illinois	DuPage County	60068	41.959	-87.617	
13	Harbor	Chicago	Illinois	DuPage County	60647	41.866	-87.642	
14		Chicago	Illinois	DuPage County	60634	41.944	-87.769	
15	ay	Chicago	Illinois	DuPage County	60640	41.640	-87.578	
16	ding	Chicago	Illinois	DuPage County	60621	41.734	-87.666	
17	Unit 29	Chicago	Illinois	DuPage County	60610	41.667	-87.629	
18	m	Chicago	Illinois	DuPage County	60661	42.001	-87.689	
19		Chicago	Illinois	DuPage County	60176	41.772	-87.562	
20		Chicago	Illinois	DuPage County	60652	41.986	-87.657	
21	unction	Chicago	Illinois	DuPage County	60610	41.972	-87.785	
22	n	Chicago	Illinois	DuPage County	60617	41.715	-87.615	
23	Suite 55	Chicago	Illinois	DuPage County	60614	41.809	-87.609	
24	leadow	Chicago	Illinois	DuPage County	60630	41.799	-87.696	
25	rade	Chicago	Illinois	DuPage County	60604	41.919	-87.672	
26	Gate	Chicago	Illinois	DuPage County	60610	41.816	-87.776	
27	r Unit 36	Chicago	Illinois	DuPage County	60617	41.842	-87.762	
28	hroughwa...	Chicago	Illinois	DuPage County	60654	42.029	-87.621	

PROPOSED METHOD WITH SYNTHEA

- Similar 5 minute walk times
- Enrich synthetic data using USA 2020 demographic data, USA 2010 Census Demographic Data, USA 2014/2018 American Community Survey (ACS) Demographic Data, USA 2020 Consumer Expenditure data, and USA 2020 Tapestry Segmentation Data.
- Azure Machine Learning for similar classification techniques.

EXPERIENCE WITH SYNTHEA

- Machine learning results not as robust as with real 911 data (NEMSIS).
- Location data not reflective of true demographics of Chicago
- Consider integration of other U.S. Synthetic Household population data (Ex. RTI) into Synthea workflow

PARTNERS

UI Health | UIC

IDPH
ILLINOIS DEPARTMENT OF PUBLIC HEALTH

NEMSIS
BETTER DATA. BETTER CARE.

Microsoft

MATTER

RTI INTERNATIONAL

OSF HEALTHCARE

REFERENCES

- AMA (2021). 2021 OVERDOSE EPIDEMIC REPORT: Physicians' actions to help end the nation's drug-related overdose and death epidemic—and what still needs to be done. Retrieved from https://end-overdose-epidemic.org/wp-content/uploads/2021/09/AMA-2021-Overdose-Epidemic-Report_92021.pdf
- Blair Turner, Wilnise Jasmin, Isabel Chung, Ponni Arunkumar, Mark Kiely, Steven Aks, Nikhil Prachand, Allison Arwady. Opioid Overdose Surveillance Report—Chicago 2019. City of Chicago, March 2021.
- DEA Intelligence Report. (2017). The Opioid Threat in the Chicago Field Division (Report No. DEA- CHI-DIR-023-17). DEA United States Drug Enforcement Administration. https://www.dea.gov/documents/2017/2017-06/2017-06-01/opioid-threat-chicago-field-division
- Public Health Data Systems Task Force (2021). Final Report of the Health Information Technology Advisory Committee's Public Health Data Systems Task Force 2021. Retrieved from https://www.healthit.gov/sites/default/files/page/2021-08/2021-07-14_PHDS_TF_2021_HITAC%20Recommendations%20Report_Signed_508_0.pdf
- "RTI U.S. Synthetic Household Population TM" *RTI International.* https://www.rti.org/impact/rti-us- synthetic-household-population%E2%84%A2 Accessed July 2021.
- The National Institute for Occupational Safety and Health (NIOSH) (2020, April 13) *Opioids in the Workplace.* Centers for Disease Control and Prevention. https://www.cdc.gov/niosh/topics/opioids/
- Xiodan Zhou (2020, July 20) Time Series Forecasting 101 – Part 4. Forecast and visualize with Exponential Smoothing. ESRI ArcGIS Blog. https://www.esri.com/arcgis-blog/products/arcgis- pro/analytics/time-series-forecasting-101-part-4-forecast-and-visualize-with-exponential-smoothing/

The Office of the National Coordinator for Health Information Technology

Thank You

Phone: 202-690-7151

Health IT Feedback Form: https://www.healthit.gov/form/healthit-feedback-form

Twitter: @onc_healthIT

LinkedIn: Search “Office of the National Coordinator for Health Information Technology”

HealthIT.gov

Subscribe to our weekly eblast at healthit.gov for the latest updates!