Rapid Actionable Data for Opioid Response in KY (RADOR-KY)

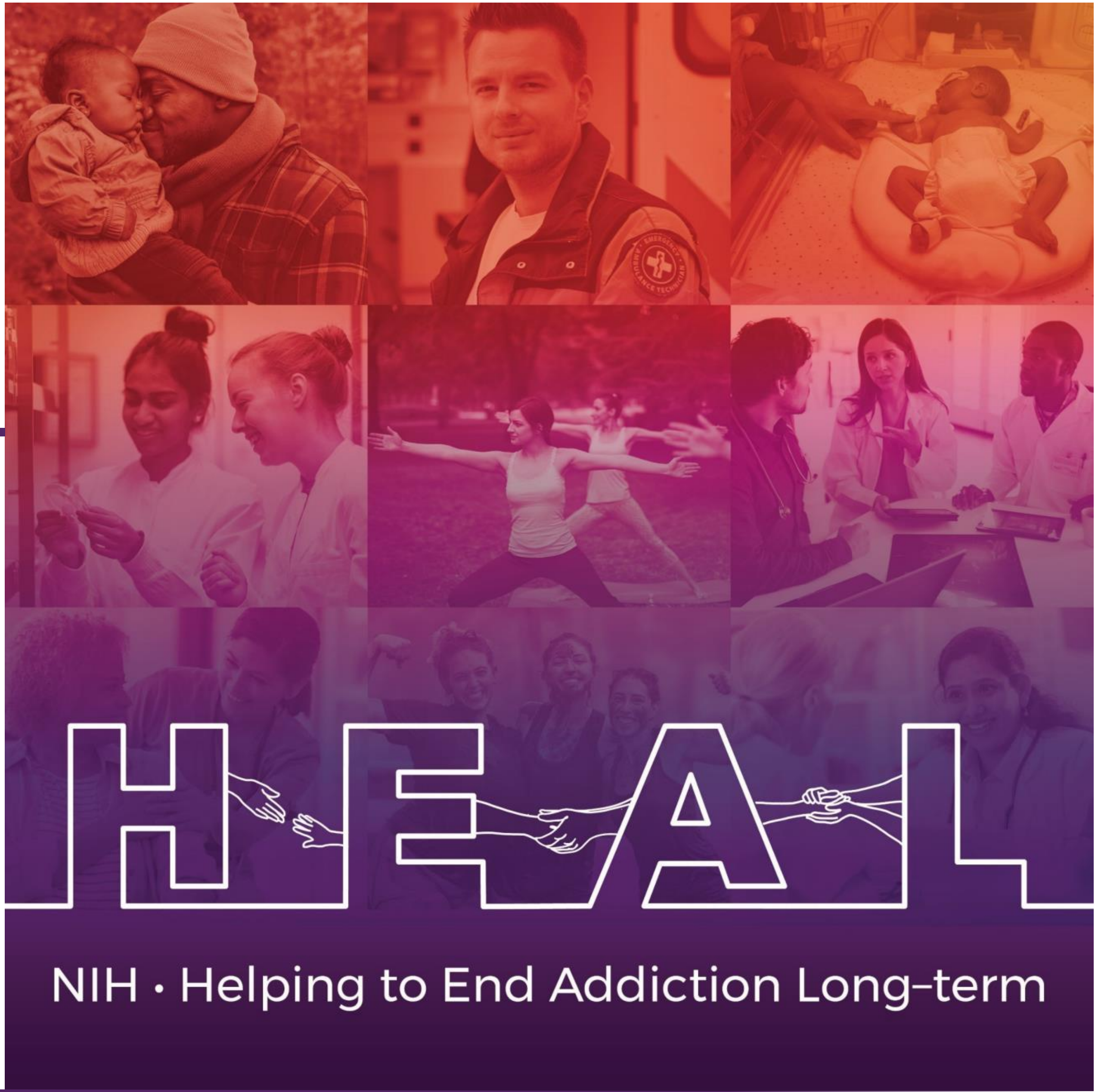
Commonwealth Computational Summit

October 17, 2023

NIH HEAL Initiative: Data and Methods to Address Urgent Needs to Stem the Opioid Epidemic

Peter Rock- RADOR-KY

Aaron Mullen- Institute for Biomedical Informatics



Outline

- RADOR-KY goals
  - Emergency Medical Services (EMS) encounter data for opioid overdose monitoring
  - Supplemental grant – focusing on AI ethics
- Optimized Forecasting
  - Testing different aggregations of space (census blocks/tracts, counties) and different timing (weekly/monthly/quarterly counts)
- Optimized Definition of opioid overdose
  - Proposed definitions vs machine learning model
- Future work

RADOR-KY

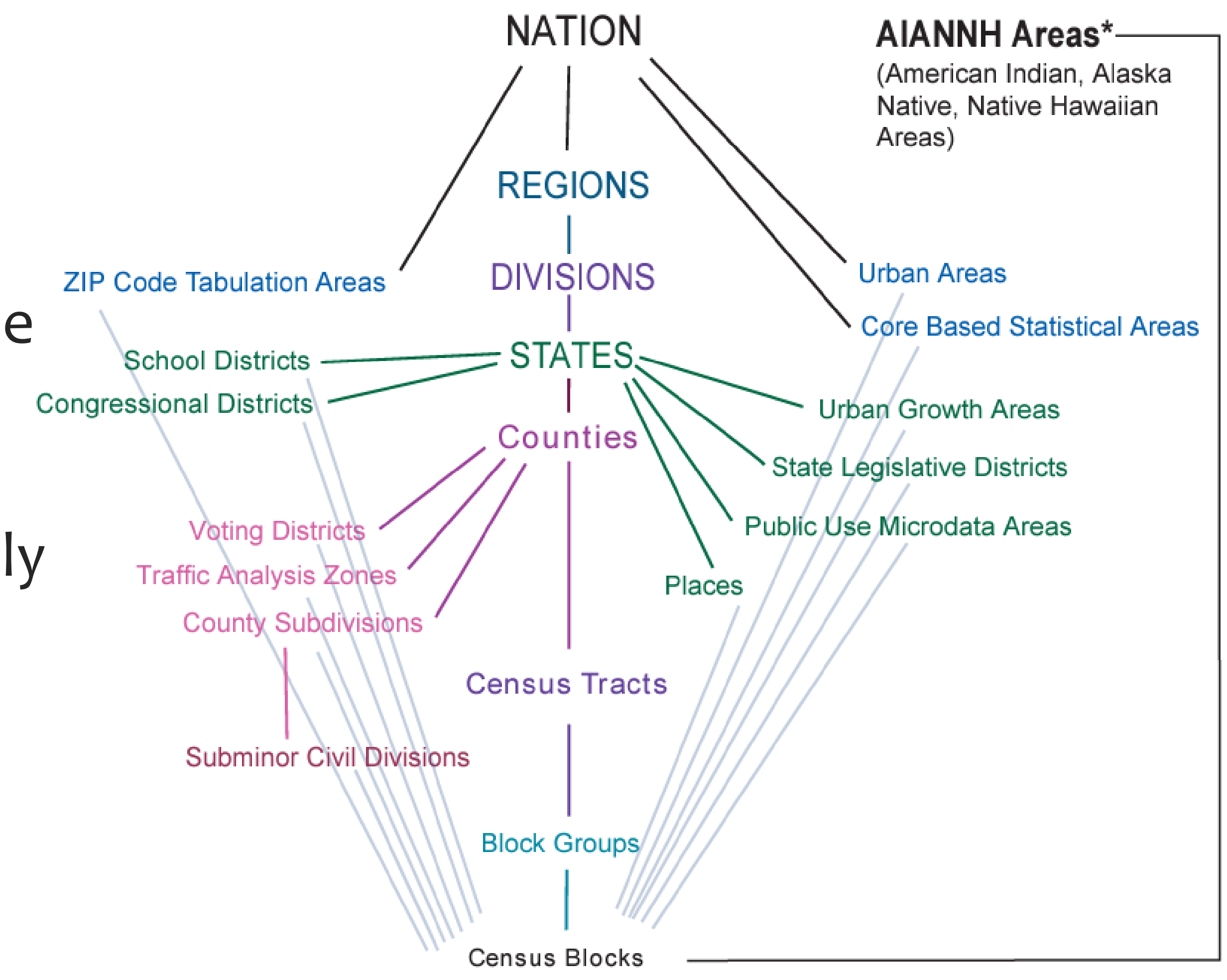
- 5-year NIH NIDA research grant; aiming to bring in a host of federal, state, and local datasets sources to inform stakeholder while improving timeliness of datasets through machine learning and traditional forecasting.

- Today, presenting preliminary work using Emergency Medical Services (EMS) run data for opioid overdose
  - Utilizing over 4.5 millions EMS records for KY statewide coverage from 2017-current.
  - Ability to capture clinical and non-clinical events.
  - Extremely timely compared to traditional datasets.

- Supplemental award focusing on differences in proposed definitions, incorporating machine learning classification; with a focus on subgroup performances.

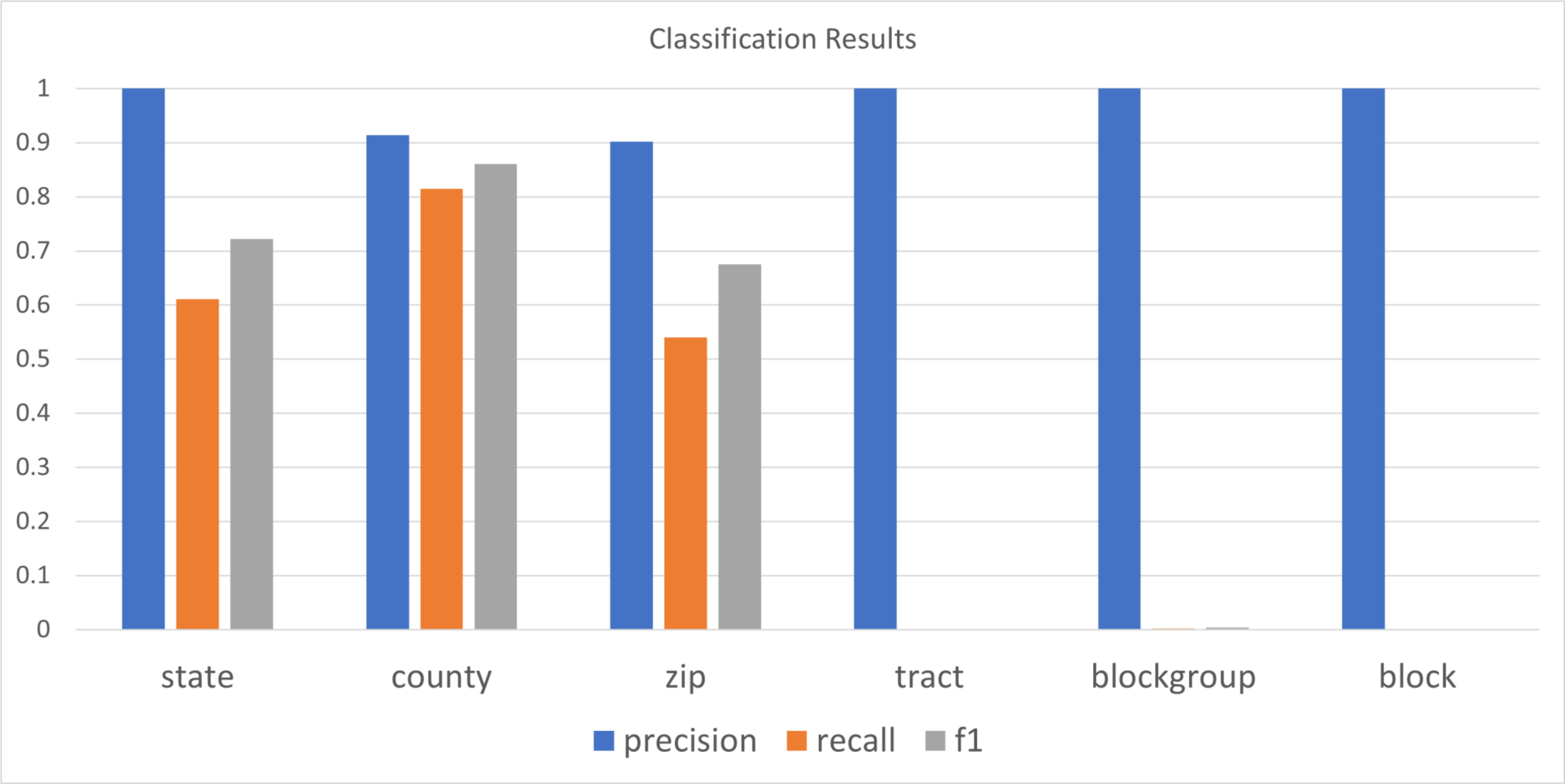
Forecasting

- Different levels of scale to analyze data
  - Geography: census block --> state
  - Time: weekly -> monthly -> yearly
- As data gets more specific, it gets more sparse
  - Standard forecasting techniques may not work effectively
  - Instead of predicting specific values, simply predict whether an incident occurred at a given location and timestep or not
  - Or predict how a county ranks in total incidents compared to other counties

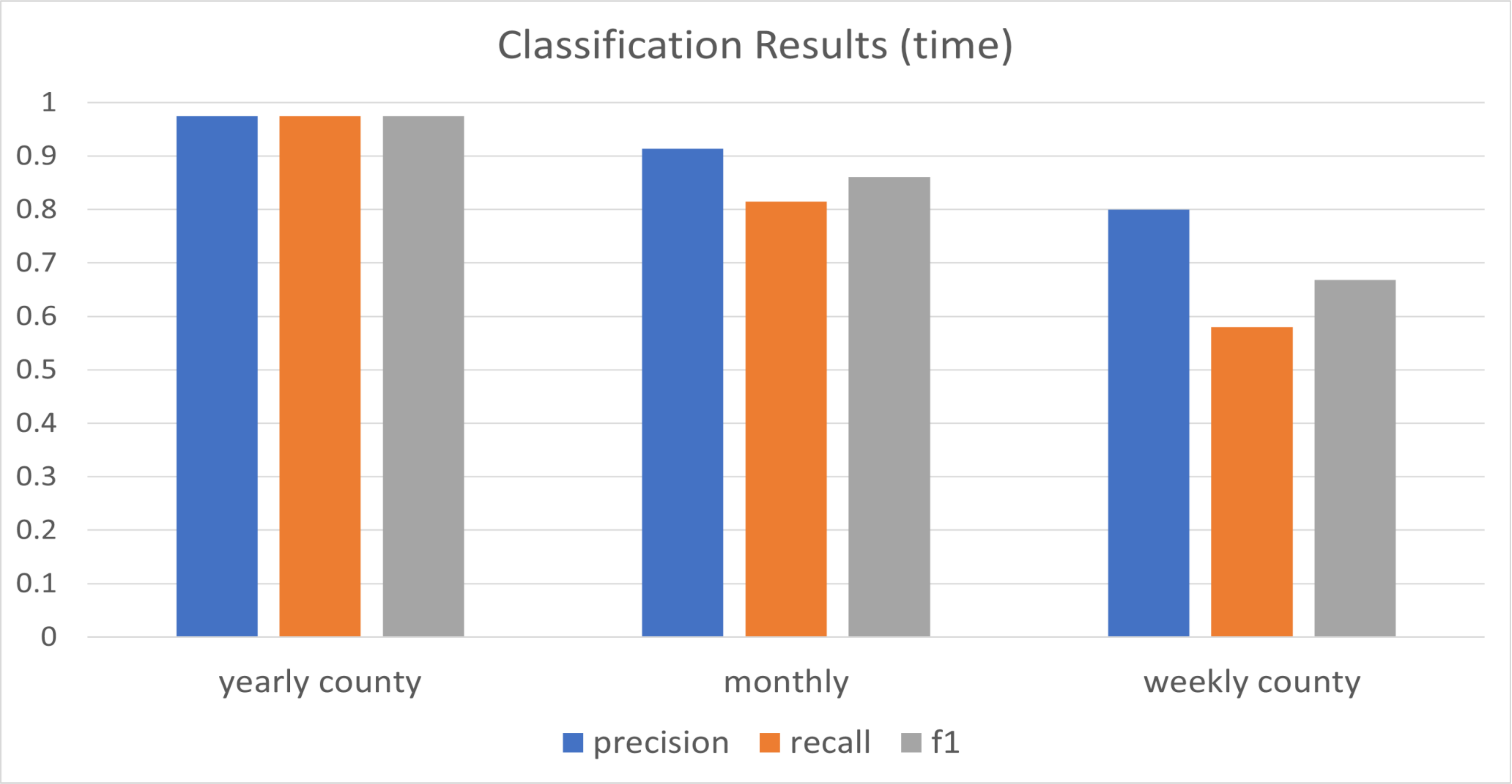


Results- Geography

- Data gets too sparse from tract level on
  - Model essentially only predicts zero
- County level is consistently best



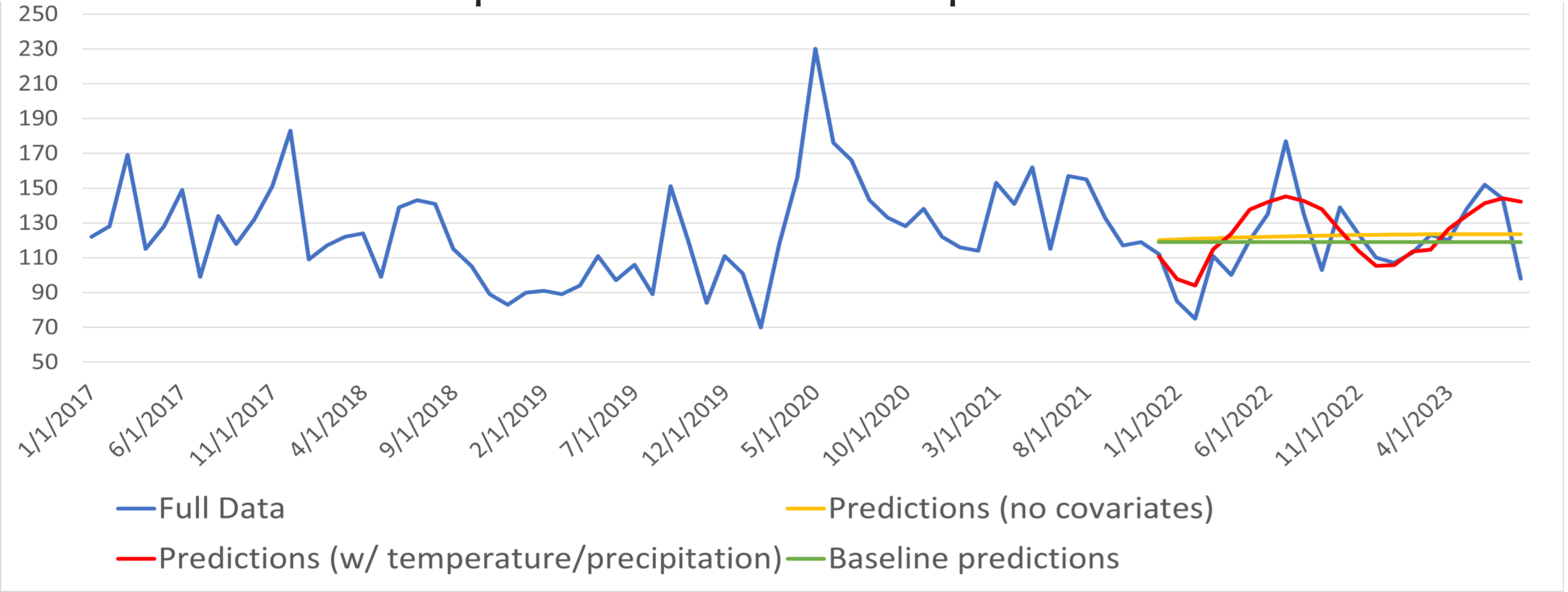
Results- Time



- Yearly performs best, but may be too general for use
  - Now just mostly 1's instead of 0's

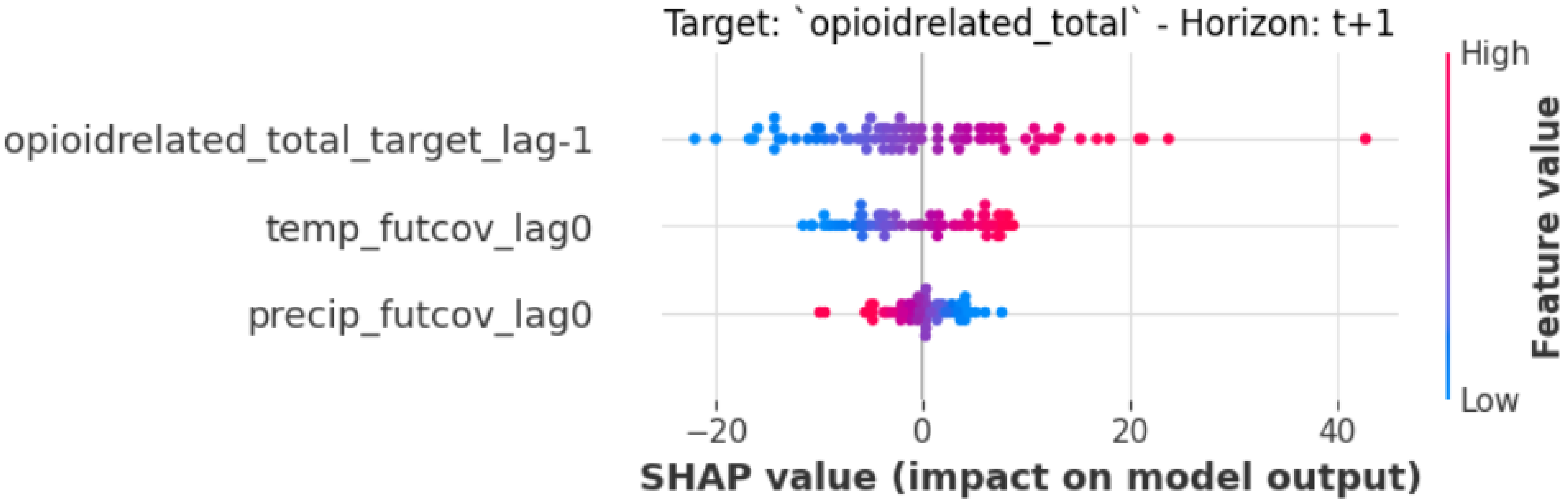
Covariates

- Problem: model predicts the same value for each time step
- Covariates can help make model more specific and accurate



SHAP Scores

- Analyze what covariates are most important to a model's predictions
- Scores demonstrate impact of each feature and whether it is positive/negative



Supplemental – Definition Evaluation

- EMS for opioid overdose is relatively new, and there is no consensus definitions.
  - Completeness of structured fields may be limited, with a reliance on patient care narratives for case ascertainment.
- Several organizations and jurisdictions have proposed various definitions that involve combinations of structured and unstructured data.

- We created a dataset of classification-difficulty-weighted sample of EMS records; labeled as opioid overdose related- by EMS paramedics.

- The following table presents our initial results comparing popular EMS definitions, as well as some early results from “simple” ML classification methods

Examined Definitions

- Definitions, ranked by perceived stringency:

More stringent definitions

1. **New York definition** - Based solely on naloxone-related structured fields.
2. **National Emergency Medical Services Information System (NEMSIS) definition** - Uses various structured fields, akin to the New York approach. Notably, NEMSIS does not access descriptive narratives.
3. **Rhode Island** - Combines structured data with narrative keyword scans.
4. **Massachusetts** - Broadens the scope, considering even non-acute opioid-related events.
5. **CSTE** - Uses a wide range of keywords and structured data points. Its approach is more inclusive.

Definition Performance

More stringent definitions

Defintion	Sensitivity/ Recall	Specificity	Accuracy	PPV/ Precision	F-score
New York	41.0%	**98.7%**	81.6%	**92.8%**	0.569
NEMSIS	44.1%	95.0%	79.9%	78.8%	0.565
Rhode Island	80.7%	86.3%	84.7%	71.3%	**0.757**
CSTE Def	82.9%	73.3%	76.2%	56.7%	0.673
Massachusetts	**94.3%**	66.4%	74.7%	54.2%	0.689
ML – NLP models					
Untuned RoBERTa	42.0%	90.8%	76.4%	65.9%	0.513
Random Forest	89.1%	95.4%	**93.5%**	89.0%	**0.891**

Machine Learning Techniques

Next steps

- RADOR-KY is early in the project.
  - Onboard many different sources of data (Toxicology, Police drug seizure data, PDMP, Justice, Vital Statistics, and more).
  - Preliminary work is already establishing the foundations for EMS overdose forecasting capabilities.

- Supplemental likewise is only ~2 weeks into the project with much more detailed evaluation and ML techniques to be incorporated.
  - Success in definitions will be incorporated into the RADOR-KY full project; as well as shared externally.

Acknowledgements

- This presentation was supported by Cooperative Agreement NU17CE924971, funded by the Centers for Disease Control and Prevention, Overdose Data to Action. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the Centers for Disease Control and Prevention.
- Research reported in this presentation was supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health under award number R01DA057605 . The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
- Research reported in this presentation was performed in collaboration with the Center for Applied Artificial Intelligence (CAAI) and the Institute for Biomedical Informatics (IBI). The content is solely the responsibility of the authors and does not necessarily represent the official views of the CAAI or IBI.

Questions and Contacts

Presenter

Peter Rock, MPH

RADOR-KY project Director

Director of Biomedical Data Science

Aaron Mullen

Institute for Biomedical Informatics

PI Contacts

Jeffery Talbert, PhD, FAMIA
Professor and University Research Professor
Division Chief for Biomedical Informatics
Director, Institute for Biomedical Informatics
Associate Director
Center for Clinical and Translational Science

Svetla Slavova, PhD
Interim Associate Dean for Research
College of Public Health
Associate Professor, Biostatistics
University Research Professor
University of Kentucky