Breakout Session 8: Track B

Battling Bias in Sepsis Prediction: Towards an Informed Understanding of EMR Data and Its Limitations

Dr. Andre Holder
Assistant Professor, Emory University

Battling Bias in Sepsis Prediction: Towards an Informed Understanding of EMR Data and It's Limitations

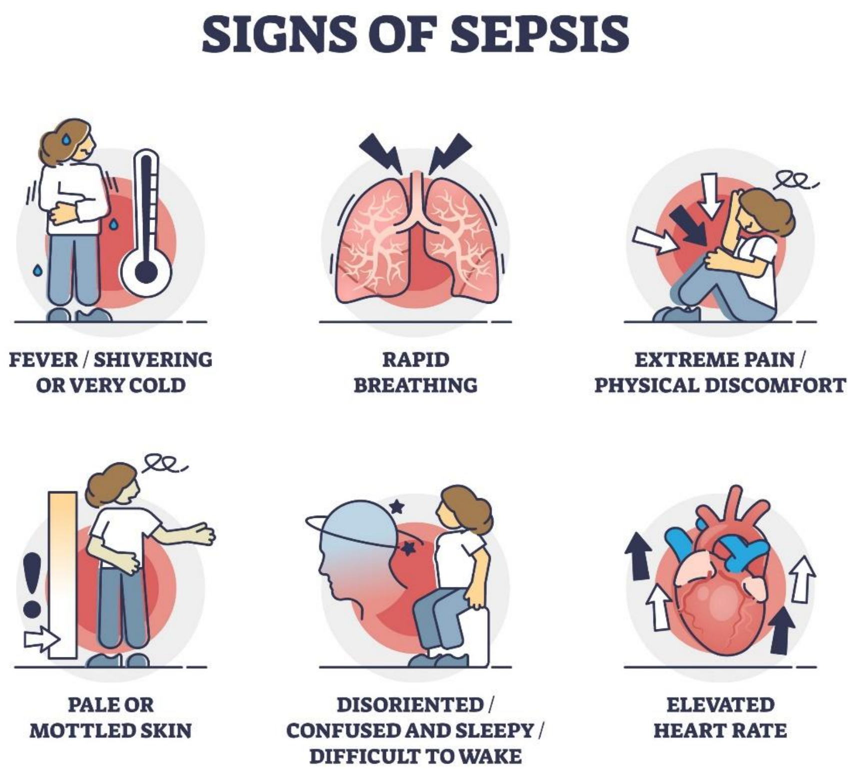
Supplement tile: Ethics and Equity in Developing Artificial Intelligence models for Patients at Risk of Sepsis

Supplement to: Characterizing Patients at Risk for Sepsis Through Big Data

Andre L. Holder, MD, MS, FCCM
Assistant Professor of Medicine
Division of Pulmonary, Allergy, Critical Care & Sleep Medicine
Emory University School of Medicine
Altanta, GA, USA

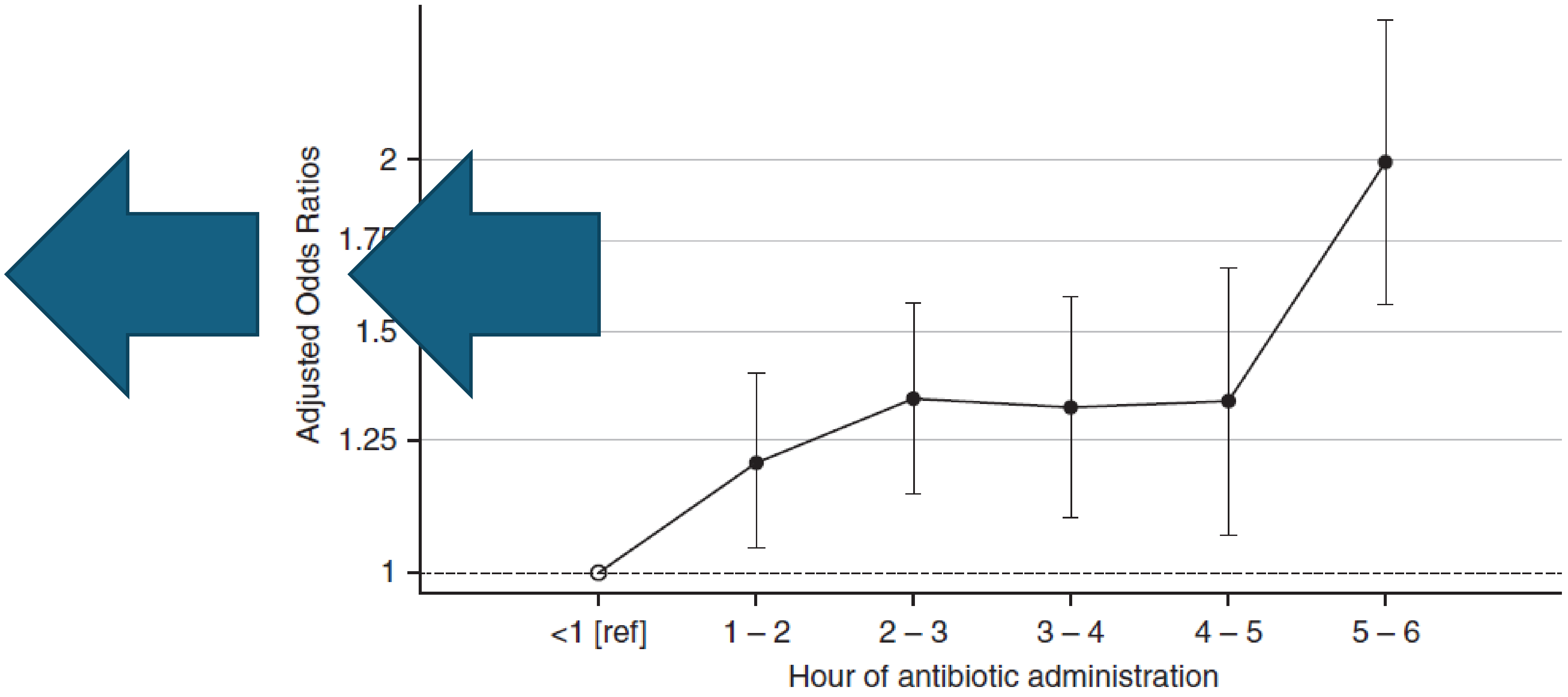
Sepsis is common and deadly

- Most common cause of death in ICUs
- 5.3 million deaths per year globally
- “True” inpatient mortality unchanged



Fleischmann C et al. *Am J Respir Crit Care Med*. 2016;193:259-272.
Angus DC et al. *Crit Care Med*. 2001;29:1303-1310.
Rhee et al. *JAMA*. 2016;318:1241-1249

Earlier antibiotics = better outcomes, so…????



Liu et al, Am J Resp Crit Care Med 2017;196(7):856-63

Bias

We define, for the first time, algorithmic bias in the context of AI and health systems as: "the instances when the application of an algorithm compounds existing inequities in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and adversely impact inequities in health systems."

- Underrepresentation in training
- Measurement bias
  - Pulse oximetry
  - Temporal thermometers
- Implicit bias in:
  - Care (e.g. pain in the ED)
  - Data collection

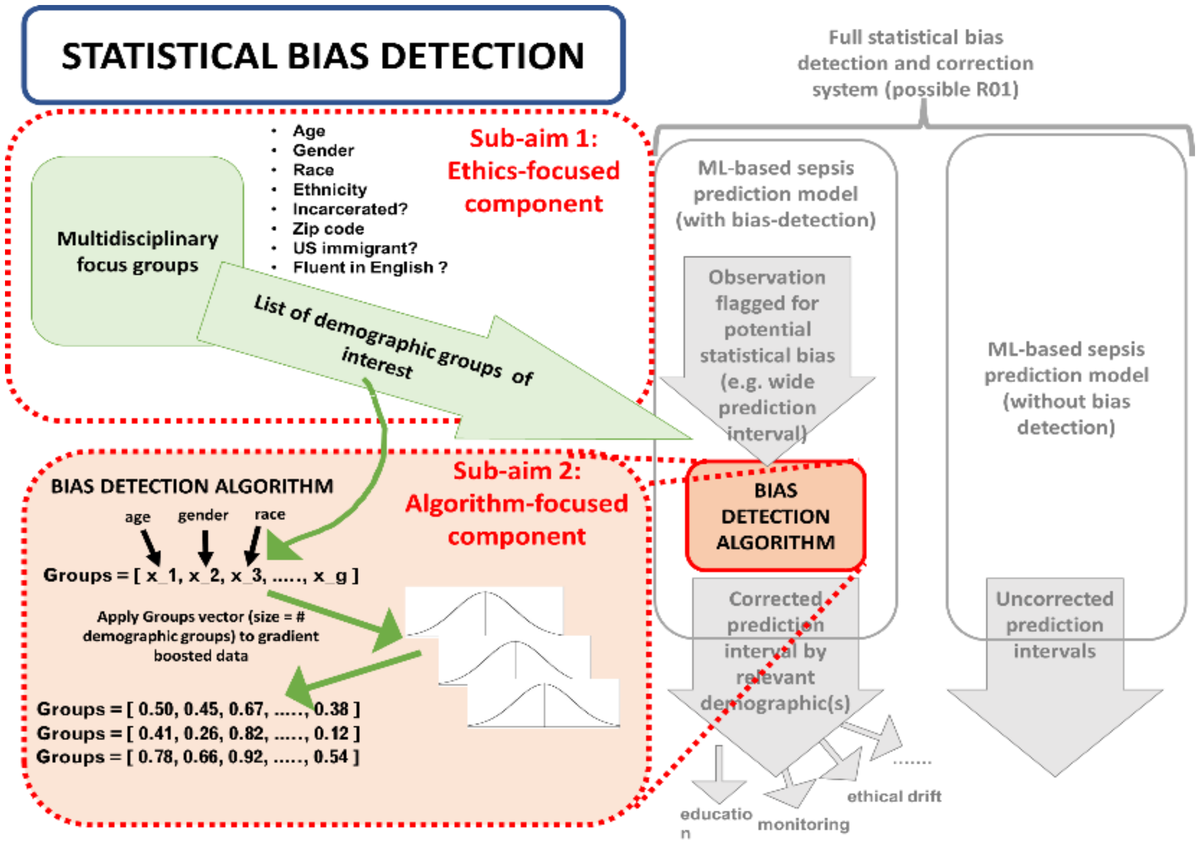
Panch et al. *J Glob Health* 2019;9(2):020318
Charpignon et al, Crit Care Clin 2023;39:751–768
Sjoding et al, N Engl J Med 2020;383:2477-2478
Bhavani et al, JAMA 2022;328(9):885-886
Van der Vegt et al. JAMIA 2023;30(7):1349–1361

Why this matters (sample structured data)

Patient	Age	Sex	Race	Latino?	BMI	Temp	SpO2	Lactate	WBC	Sepsis?
1	35	M	Black	N	35	38.4	99	???	8.6	Y
2	67	F	White	N	40	37.4	89	1.7	10.6	N
3	48	F	Black	Y	31	36.6	90	???	3.4	N
4	65	M	Asian	N	29	37.3	99	2.5	26.9	Y
5	87	F	White	N	27	39.1	100	1.5	5.0	N
6	67	M	White	Y	36	38.0	98	???	4.7	Y
7	79	M	White	N	29	37.8	96	3.0	9.9	Y
8	58	F	Black	N	40	38.0	91	2.3	12	Y

Project summary

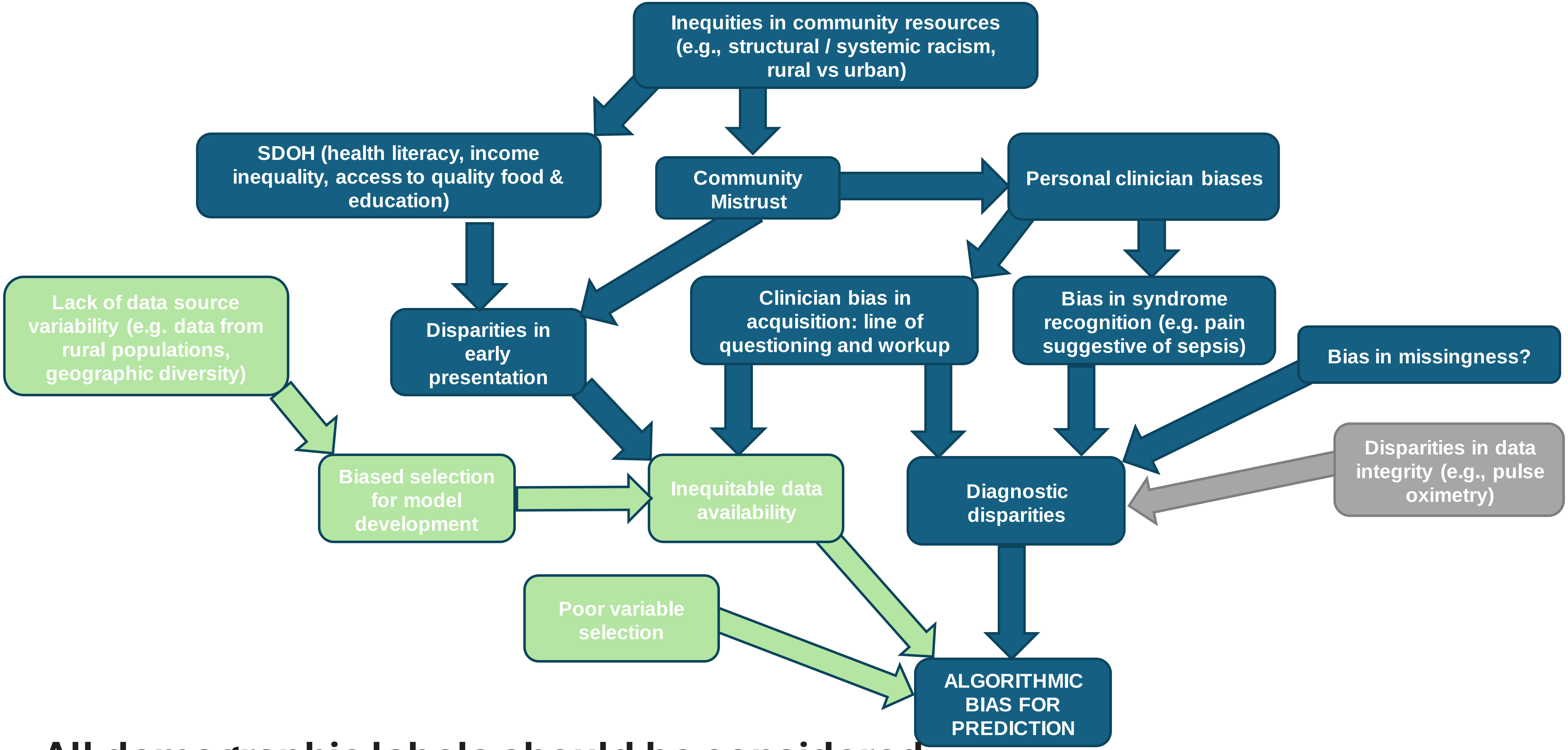
- Aim: Develop a robust, ethically-informed model that will provide quantitative measures of the relative importance of demographic group labels in predicting disease among those susceptible to healthcare disparities
    - **Ethics-focused component:** Convene a multidisciplinary focus group
    - **Algorithm-focused component:** Develop a novel health equity metric



Focus group details

- Focus group goals
    1. Identify the groupings at risk
    2. Identify causes of *unequal sepsis care* that might also contribute to *inequitable prediction*
    3. Discuss algorithmic choices that could exacerbate inequalities
    4. Understand difference between perceived and actual risk of inequitable prediction
- 14 participants (4 clinicians, 3 data scientists, 2 ethicists, 5 advocates)
- Three sessions in 2023: 1/12, 1/26, & 1/30

Social bias effects on critical care prediction



- **All demographic labels should be considered**

Subgroup Performance Assessment, Detection & Evaluation (SPADE)

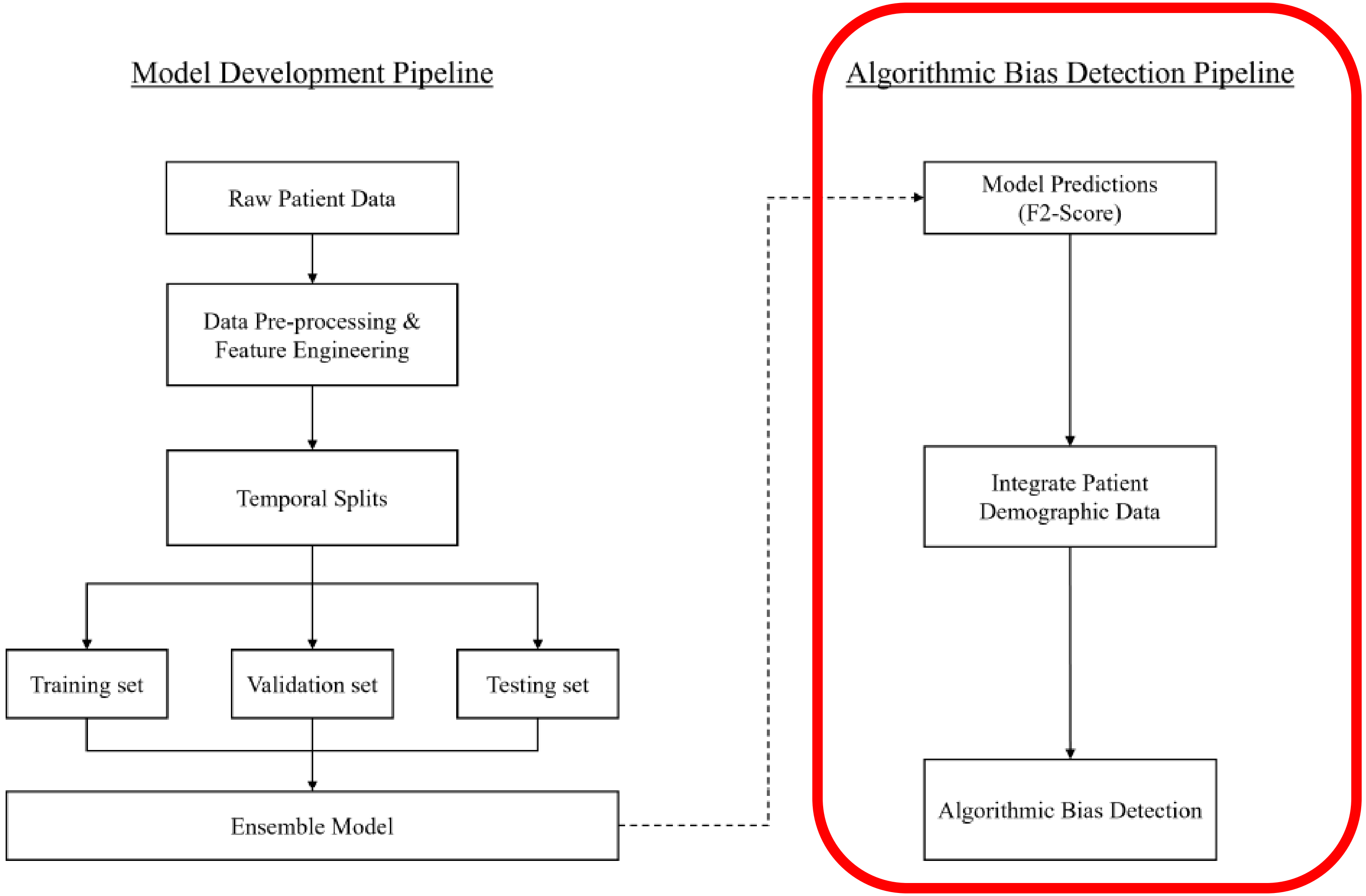


Figure 8: Algorithmic bias detection pipeline.

Sepsis model development pipeline

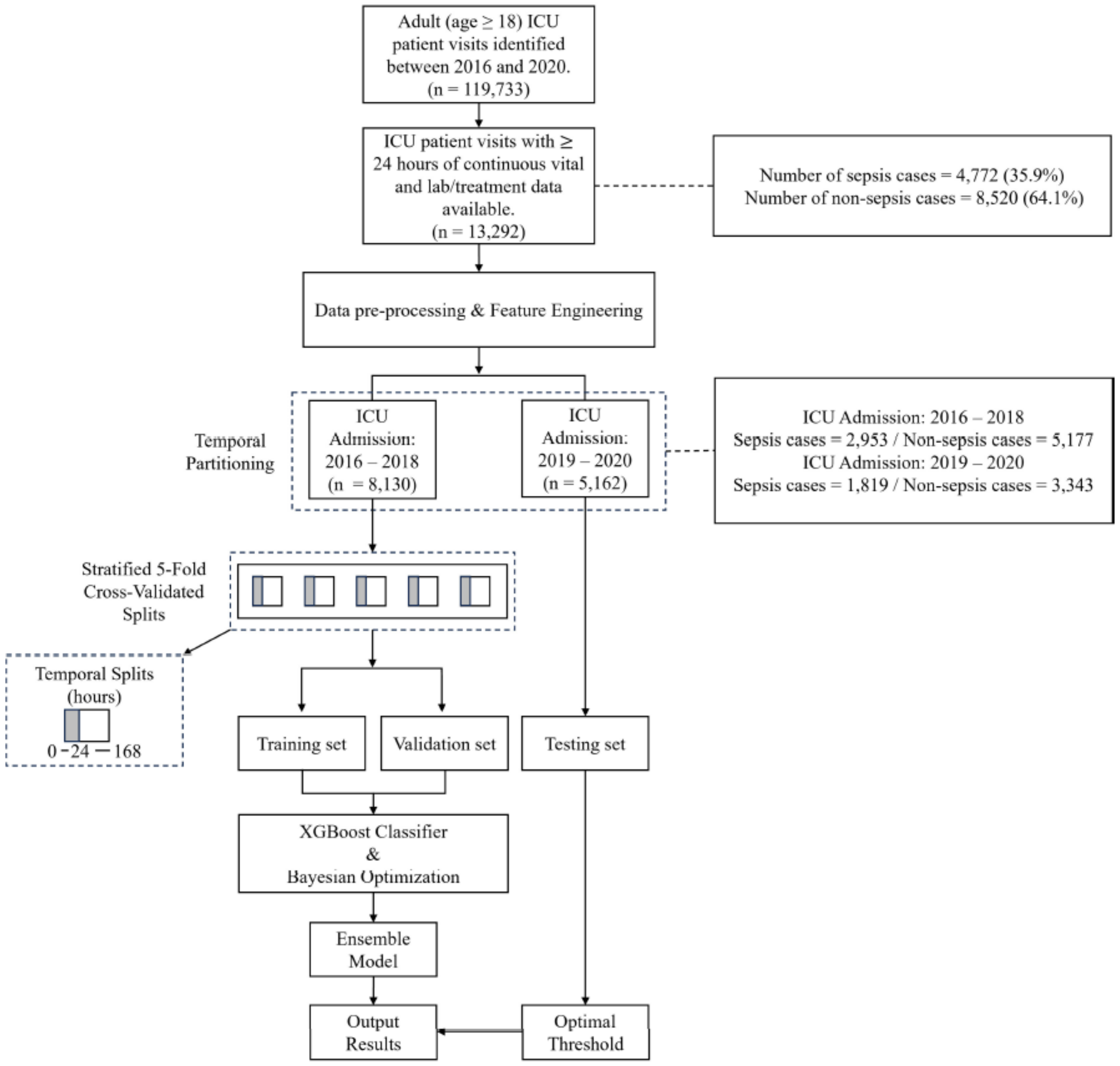
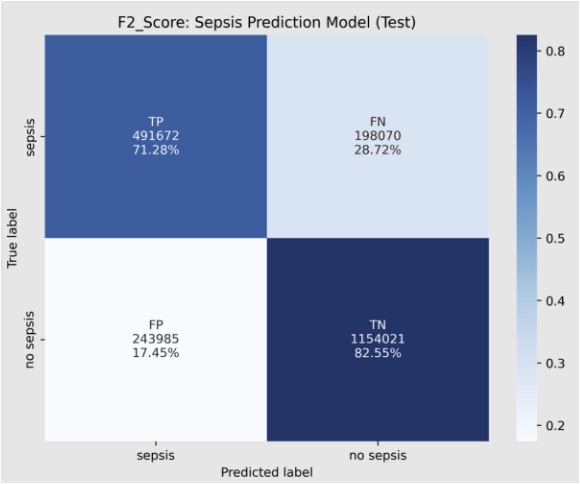
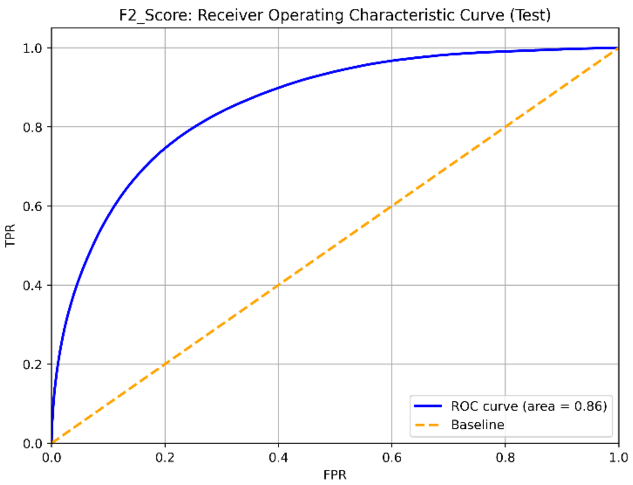


Figure 5: Data pre-processing and model development diagram.

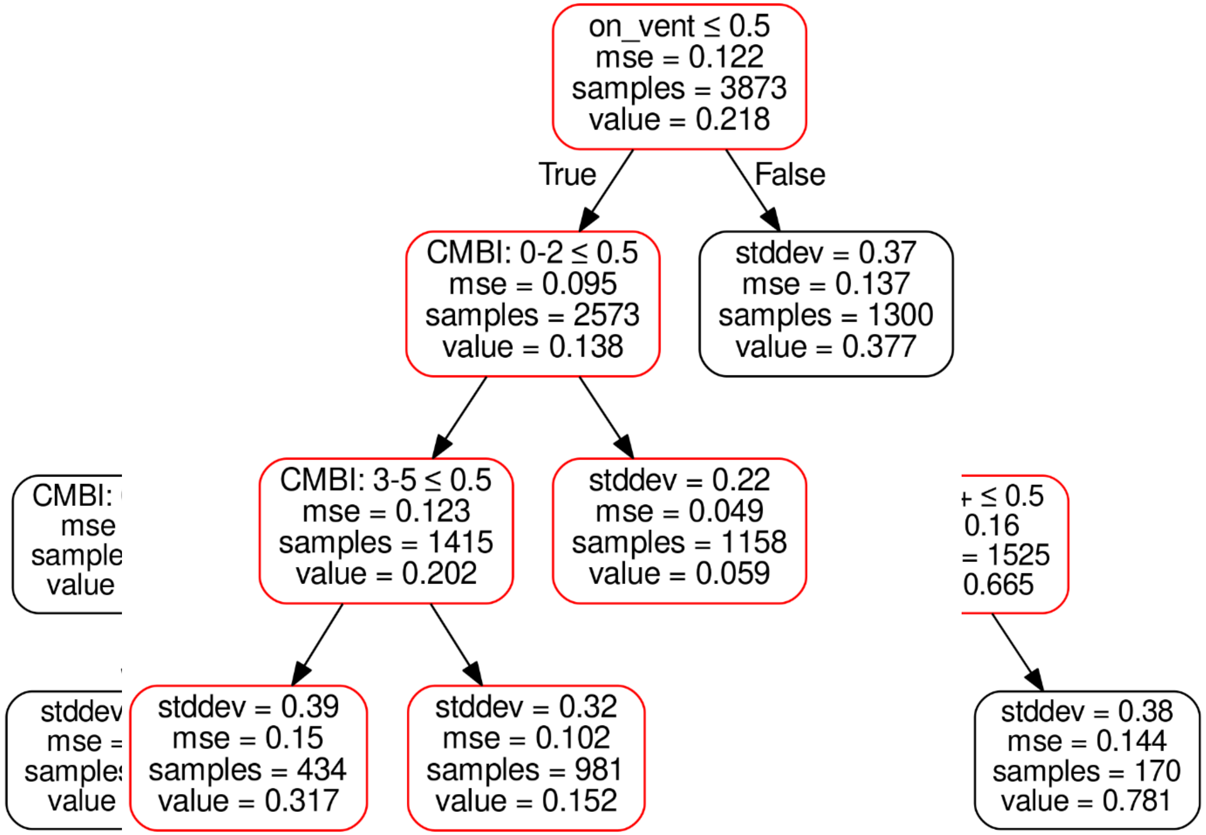
- Model: XGBoost
  - Bayesian optimization
  - Tree-structured Parzen Estimator (TPE) approach





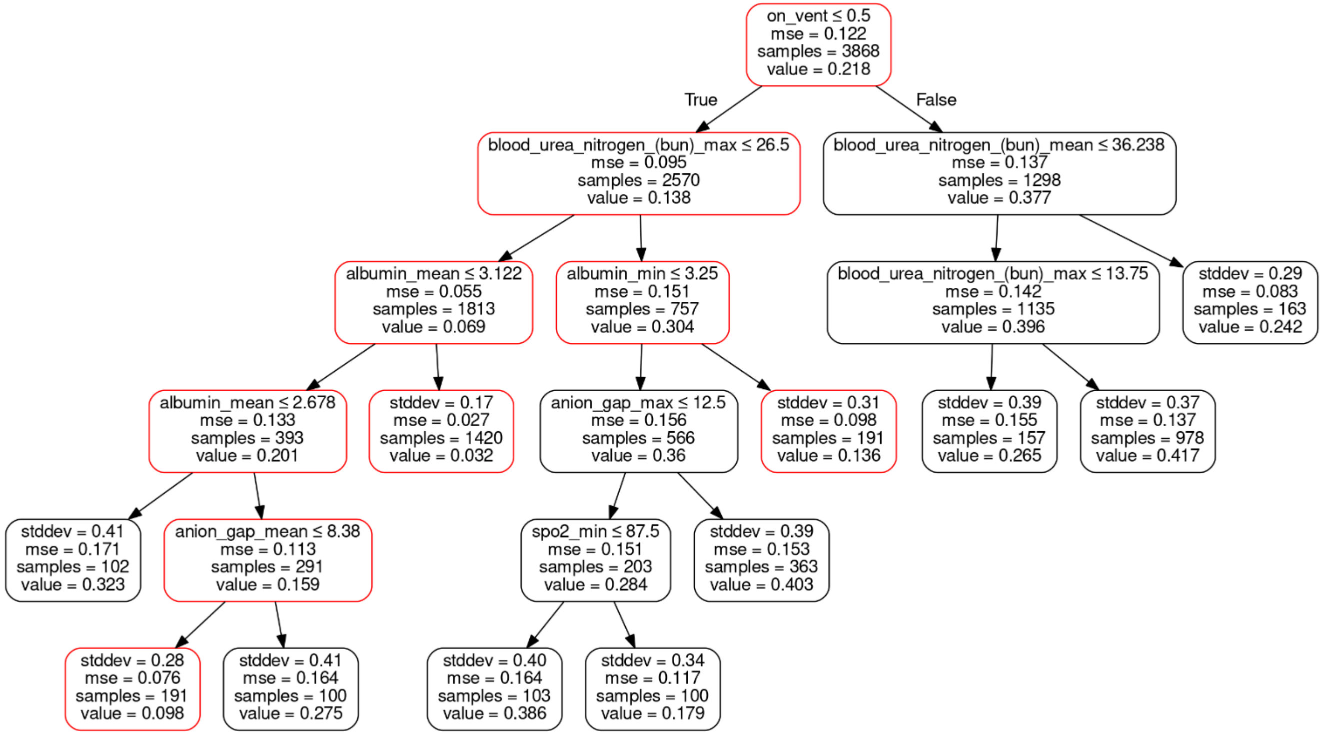
SPADE identifies bias by differences from mean performance within the cohort

- CART analysis
  - Test data only (2019-2020)
- 8 input (discriminating) features:
  - Race, age, gender, incarceration status, distance to hospital (based on home zip code), homelessness, insurance type, Elixhauser comorbidity index
- Optimization based on primary sepsis model (e.g., accuracy)
- Adjusting SPADE optimization changes the output



Advantages over other approaches

- Algorithm agnostic
- No *a priori* assumptions
- Captures intersectionality
- Not just limited to use on sensitive labels like race



Challenges

- Ethics-focused component
  - How should we define bias?
  - Participants confused about the ask
  - No qualitative analysis background
  - Fitting focus group results into existing ethical frameworks
- Algorithm-focused component
  - Operationalizing bias
    - How? What metrics?
  - Working within the limits of some labels (e.g., SES, incarceration)

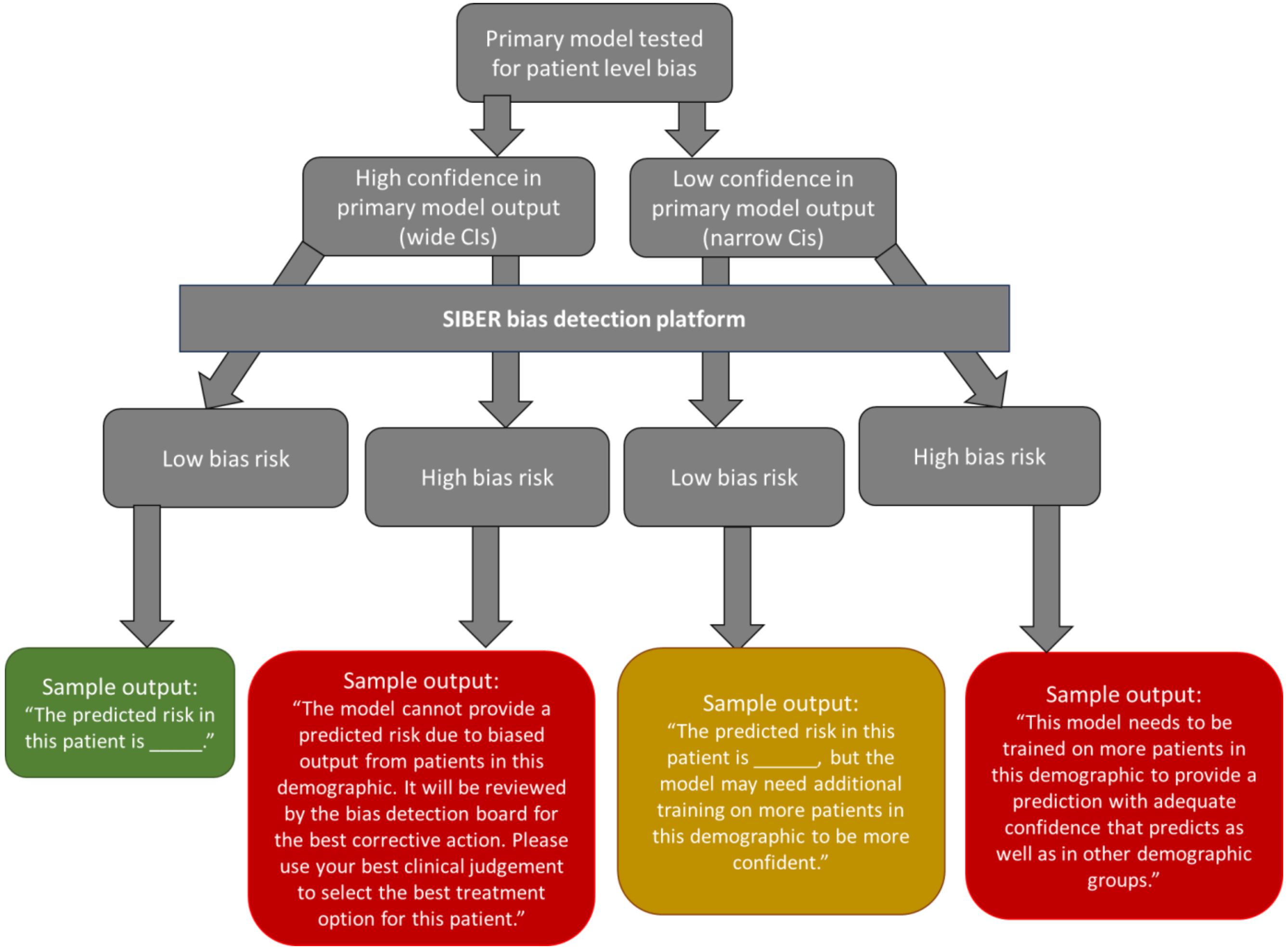
		Levels of Influence*			
		Individual	Interpersonal	Community	Societal
Domains of Influence (Over the Lifecourse)	Biological	Biological Vulnerability and Mechanisms	Caregiver–Child Interaction Family Microbiome	Community Illness Exposure Herd Immunity	Sanitation Immunization Pathogen Exposure
	Behavioral	Health Behaviors Coping Strategies	Family Functioning School/Work Functioning	Community Functioning	Policies and Laws
	Physical/Built Environment	Personal Environment	Household Environment School/Work Environment	Community Environment Community Resources	Societal Structure
	Digital Environment	Digital Literacy, Digital Self-Efficacy, Technology Access, Attitudes Towards Use	Implicit Tech Bias, Interdependence (e.g. shared devices), Patient-Tech-Clinician Relationship	Community Infrastructure, Healthcare Infrastructure, Community Tech Norms, Community Partners	Tech Policy, Data Standards, Design Standards, Social Norms & Ideologies, Algorithmic Bias
	Sociocultural Environment	Sociodemographics Limited English Cultural Identity Response to Discrimination	Social Networks Family/Peer Norms Interpersonal Discrimination	Community Norms Local Structural Discrimination	Social Norms Societal Structural Discrimination
	Health Care System	Insurance Coverage Health Literacy Treatment Preferences	Patient–Clinician Relationship Medical Decision-Making	Availability of Services Safety Net Services	Quality of Care Health Care Policies
Health Outcomes		Individual Health	Family/ Organizational Health	Community Health	Population Health

Fig. 1 Framework for digital health equity. National Institute on Minority Health and Health Disparities Research Framework Expanded for Digital Health Equity.

Richardson et al, npj Digital Medicine (2022) 5:119

Future research

- Defining causes of bias (lack of data source variability? Measurement bias?)
- Implementation into an existing AI infrastructure



https://www.hsph.harvard.edu/ecpe/how-to-prevent-algorithmic-bias-in-health-care/

Summary

- We should probably assess bias in a very inclusive list of sociodemographic and comorbidity labels, but know their limits
- A post hoc, model-agnostic approach to identifying bias within certain patient subgroups is feasible
- SPADE has advantages over *a priori* decisions of bias detection
- Output can vary for the same model based on multiple factors
- This approach will need to be prospectively and externally validated, and operationalized in an actionable way to improve equity in sepsis prediction