Applications of machine learning to support antimicrobial stewardship

Use cases and next steps

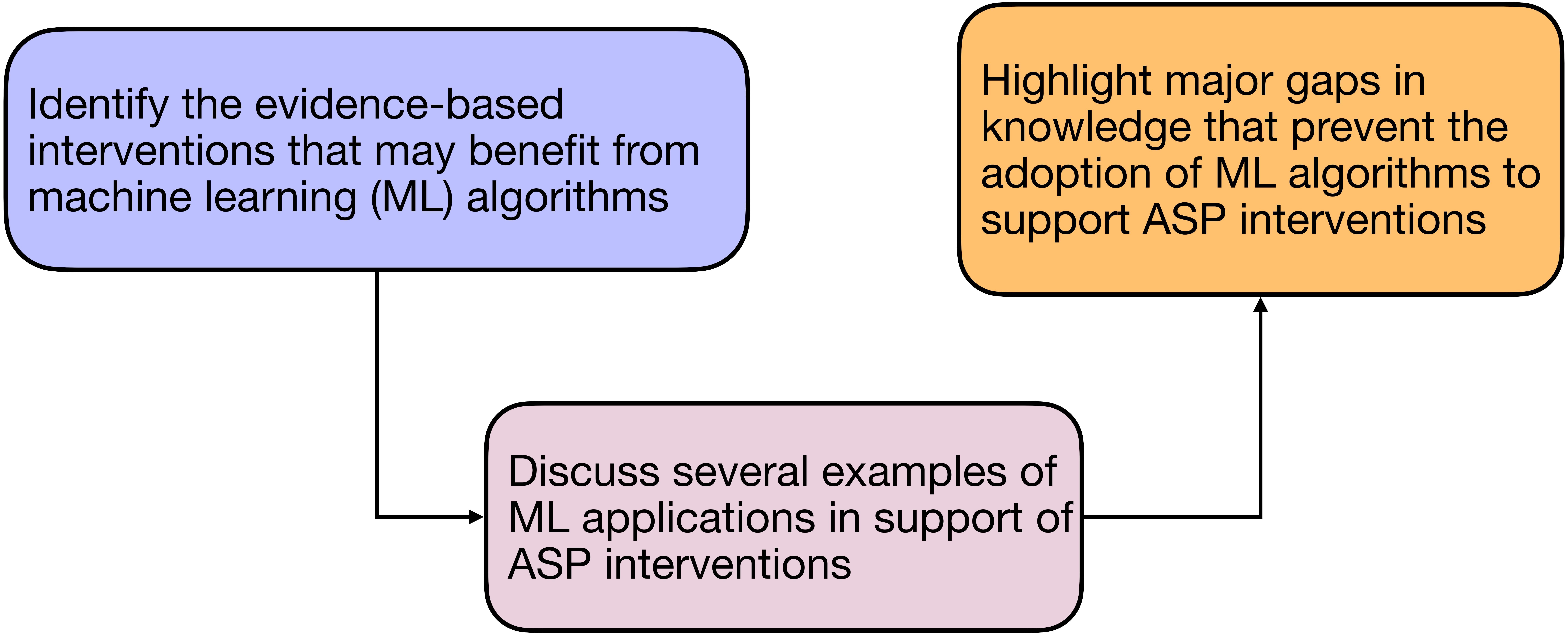
January 16 2025

Sanjat Kanjilal
Department of Population Medicine, HMS and Harvard Pilgrim Healthcare Institute
Division of Infectious Diseases, Brigham & Women's Hospital
@sanjatkanjilal.bsky.social

Disclosures

- Scientific advisor for PhAST Diagnostics
- All relevant financial disclosures have been mitigated

Outline

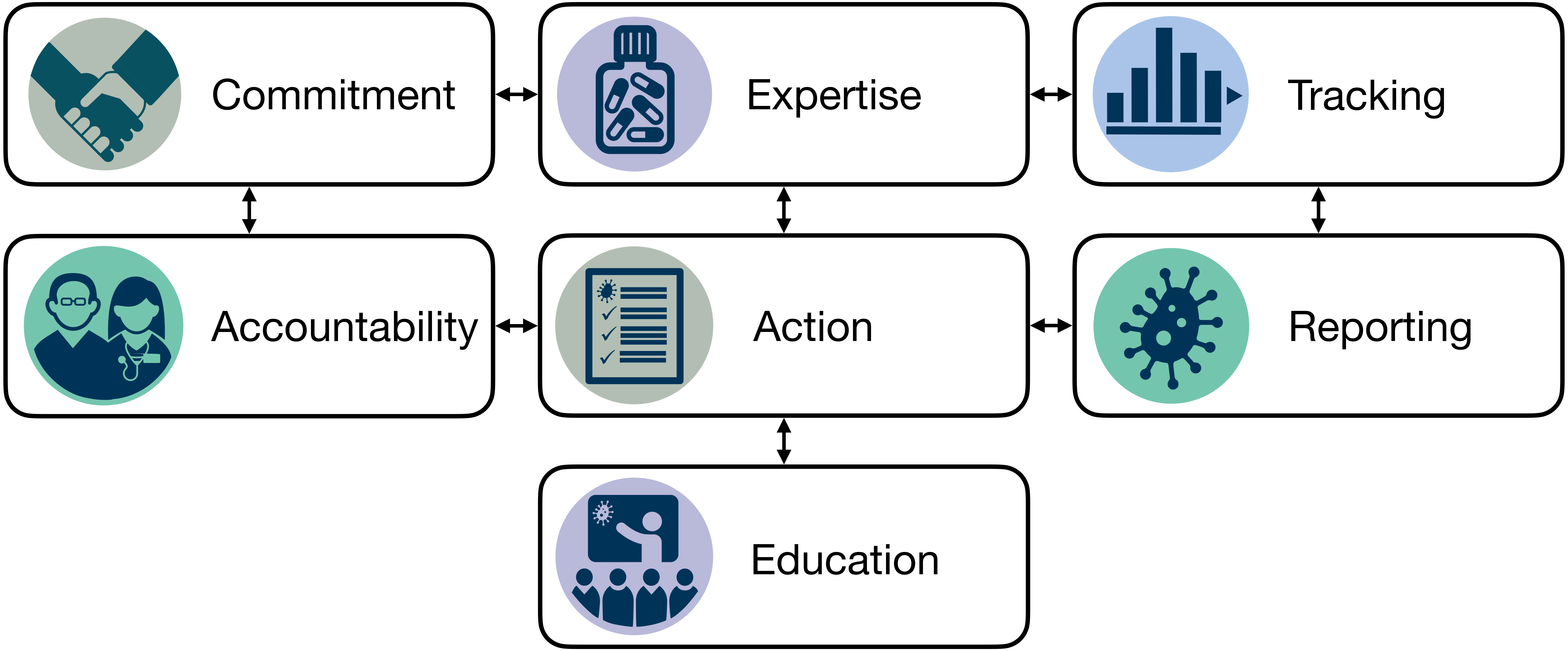

Identify the evidence-based interventions that may benefit from machine learning (ML) algorithms
Discuss several examples of ML applications in support of ASP interventions
Highlight major gaps in knowledge that prevent the adoption of ML algorithms to support ASP interventions

Antimicrobial stewardship

- Antimicrobial stewardship is complex and hard!
  - Must navigate fuzzy science, entrenched behaviors, backwards financial incentives
- Relies fundamentally on the collaboration between stewardship leaders and clinical teams

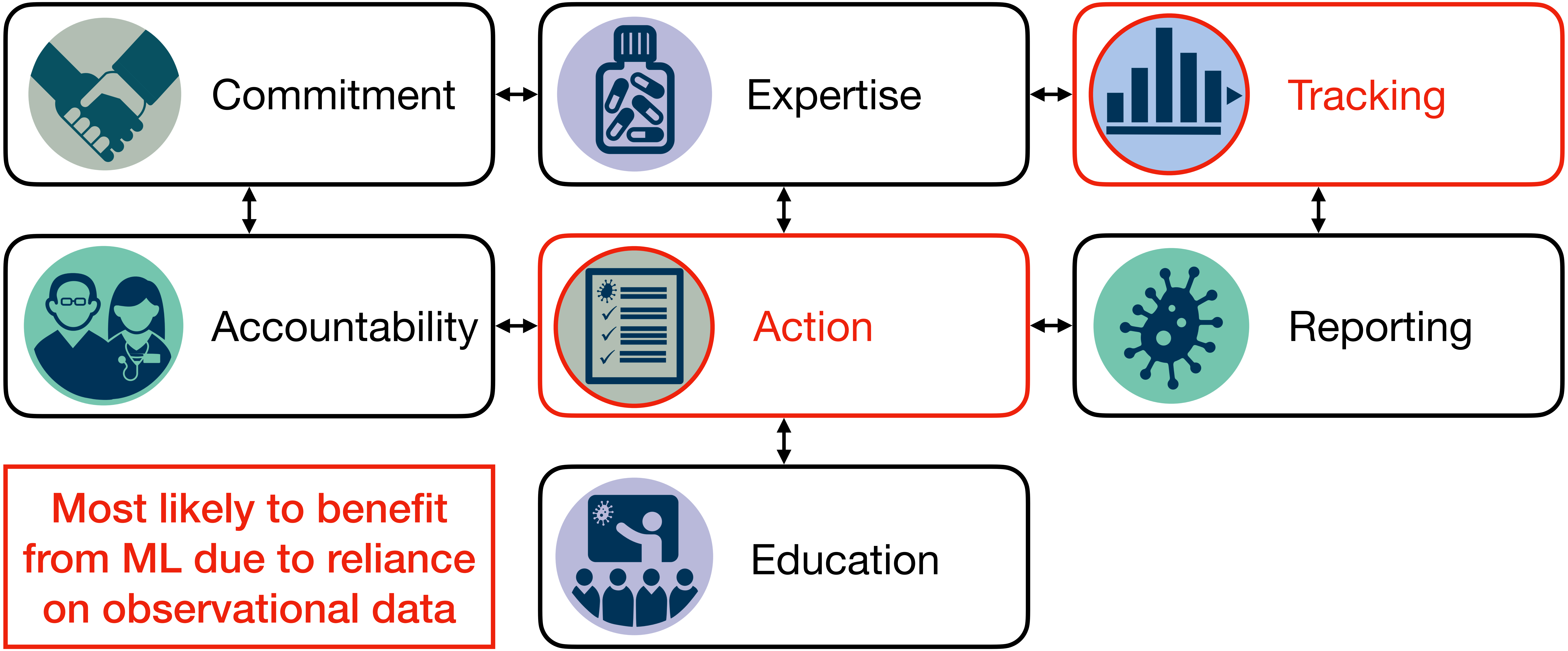
How can ML algorithms support that collaboration?

Core elements



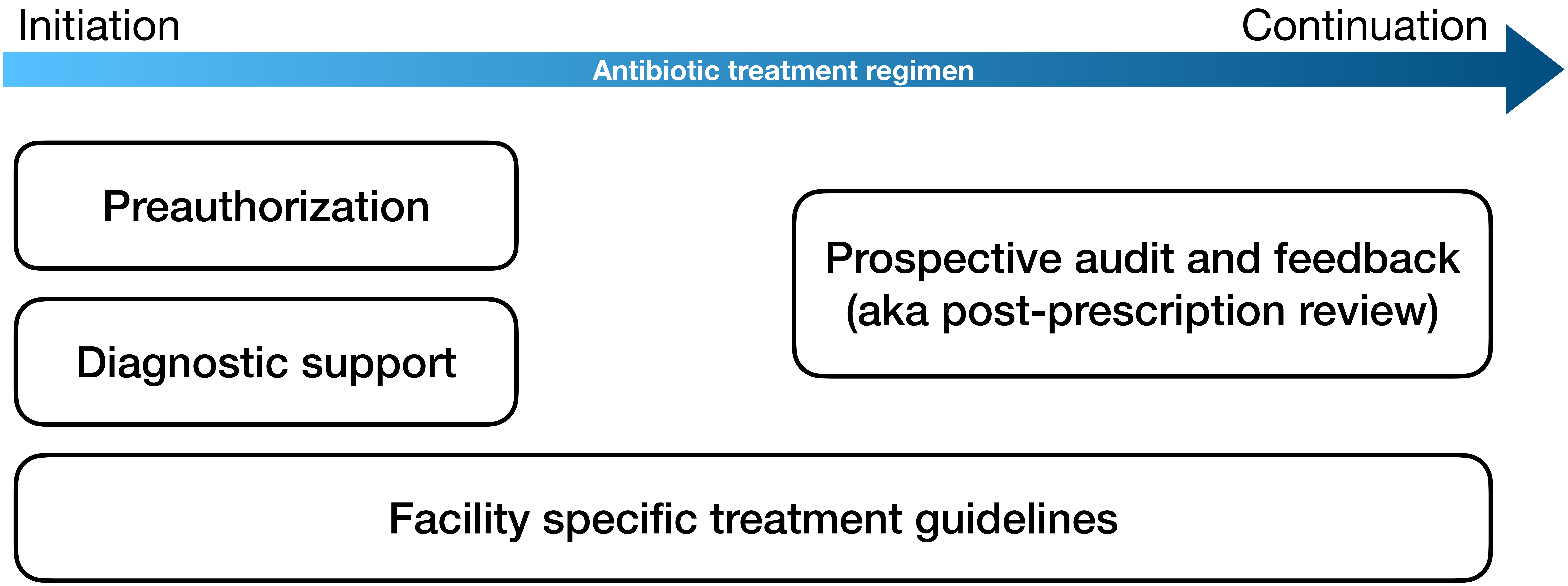
CDC, Core Elements of Hospital Antibiotic Stewardship Programs

Core elements



CDC, Core Elements of Hospital Antibiotic Stewardship Programs

Action



*Incomplete list

Action

Initiation

Continuation

Antibiotic treatment regimen

Preauthorization

Diagnostic support

Prospective audit and feedback (aka post-prescription review)

Facility specific treatment guidelines

Most likely to benefit from ML

*Incomplete list

Tracking

Antibiotic use measures	• Days of therapy • Standardized Antimicrobial Administration Ratio (SAAR)
Outcome measures	• *C. difficile* infections • Antibiotic resistance • Financial impact
Process measures	• Acceptance of feedback • Rates of preauthorization / time to appropriate therapy • Adherence to facility specific guidelines

*Incomplete list

Tracking

Antibiotic use measures

- Days of therapy
- Standardized Antimicrobial Administration Ratio (SAAR)

Outcome measures

- *C. difficile* infections
- Antibiotic resistance
- Financial impact

Process measures

- Acceptance of feedback
- Rates of preauthorization / time to appropriate therapy
- Adherence to facility specific guidelines

Most likely to benefit from ML

*Incomplete list

Modeling goals

Prediction

$$\mathbb{P}(Y \in A \mid X = x)$$

Learns patterns in training data to make predictions on unseen ‘test’ data*

Diagnostic support
Audit & Feedback
SAAR

Causal inference

$$\mathbb{P}(Y \in A \mid \text{set } X = x)$$

Asks what happens to an outcome as a result of a treatment or intervention*

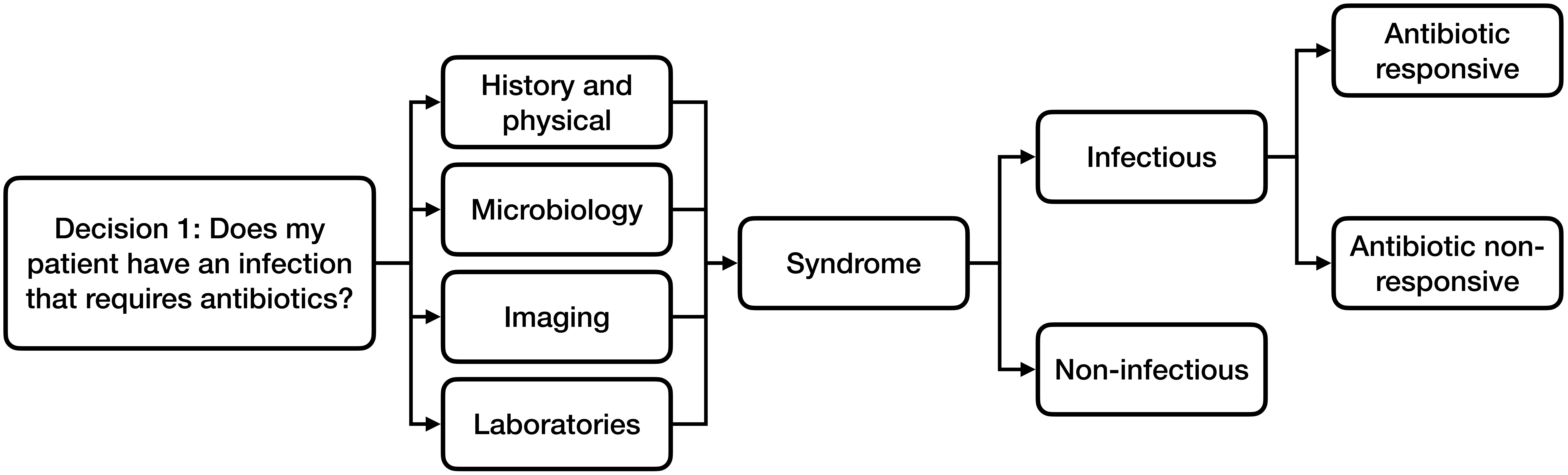
Hospital treatment guidelines
C. difficile rates
Antibiotic resistance rates

* Highly simplified!

Diagnostic support

Identifying the presence of infection

- Determining the correct syndrome is a critical but challenging first step in antimicrobial stewardship



AHRQ, Four moments of Antibiotic Decision Making

Diagnostic support

Identifying the presence of infection

Decision 1: Does my patient have an infection that requires antibiotics?

Very few peer-reviewed published papers have attempted to differentiate, non-infectious and non-antibiotic responsive infections from treatable bacterial infections using EHR data

↓

Obtaining high quality labelled data identifying ‘antibiotic-responsive infection’ is challenging

AHRQ, Four moments of Antibiotic Decision Making

Diagnostic support

Identifying the presence of infection

Decision 1: Does my patient have an infection that requires antibiotics?

PLOS DIGITAL HEALTH

Can the application of machine learning to electronic health records guide antibiotic prescribing decisions for suspected urinary tract infection in the Emergency Department?

Rockenschaub, PLoS Digital Health, 2023

Used EHR data to predict bacteriuria but not asymptomatic bacteriuria

Journal of Antimicrobial Chemotherapy

Supervised machine learning to support the diagnosis of bacterial infection in the context of COVID-19

Rawson, J Antimicrob Chemother, 2019

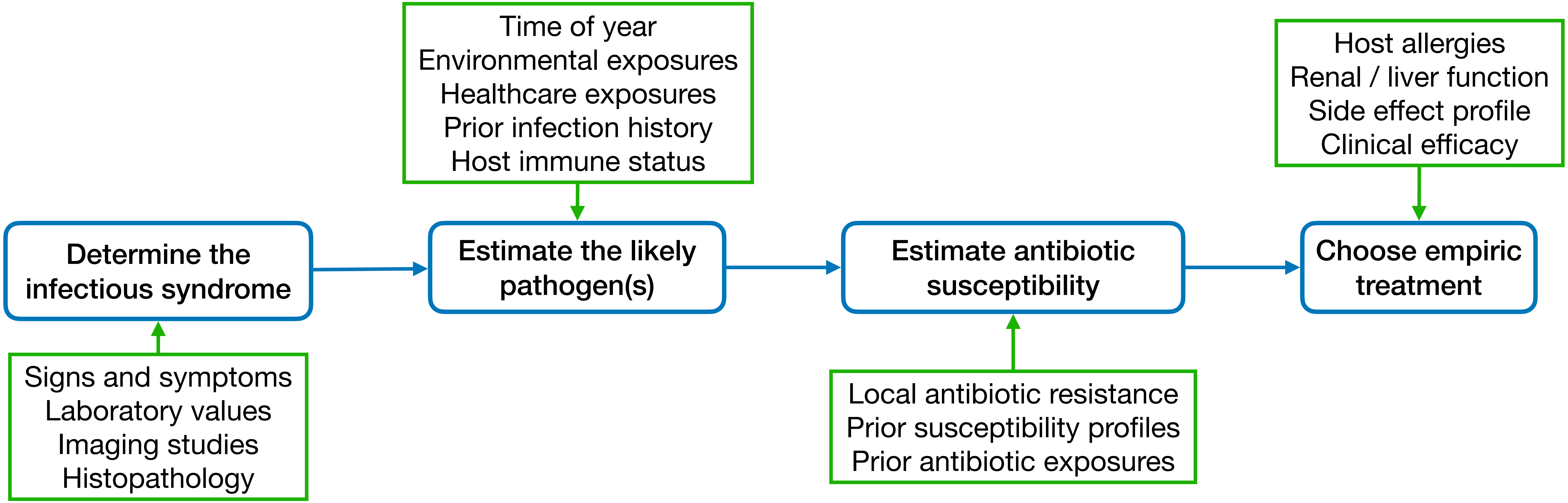
Used EHR data to predict positive bacterial cultures as proxy for infection but does not account for colonization or contamination

AHRQ, Four moments of Antibiotic Decision Making

Diagnostic support

Initiation of empiric therapy

Decision 2: What empiric therapy should I initiate?

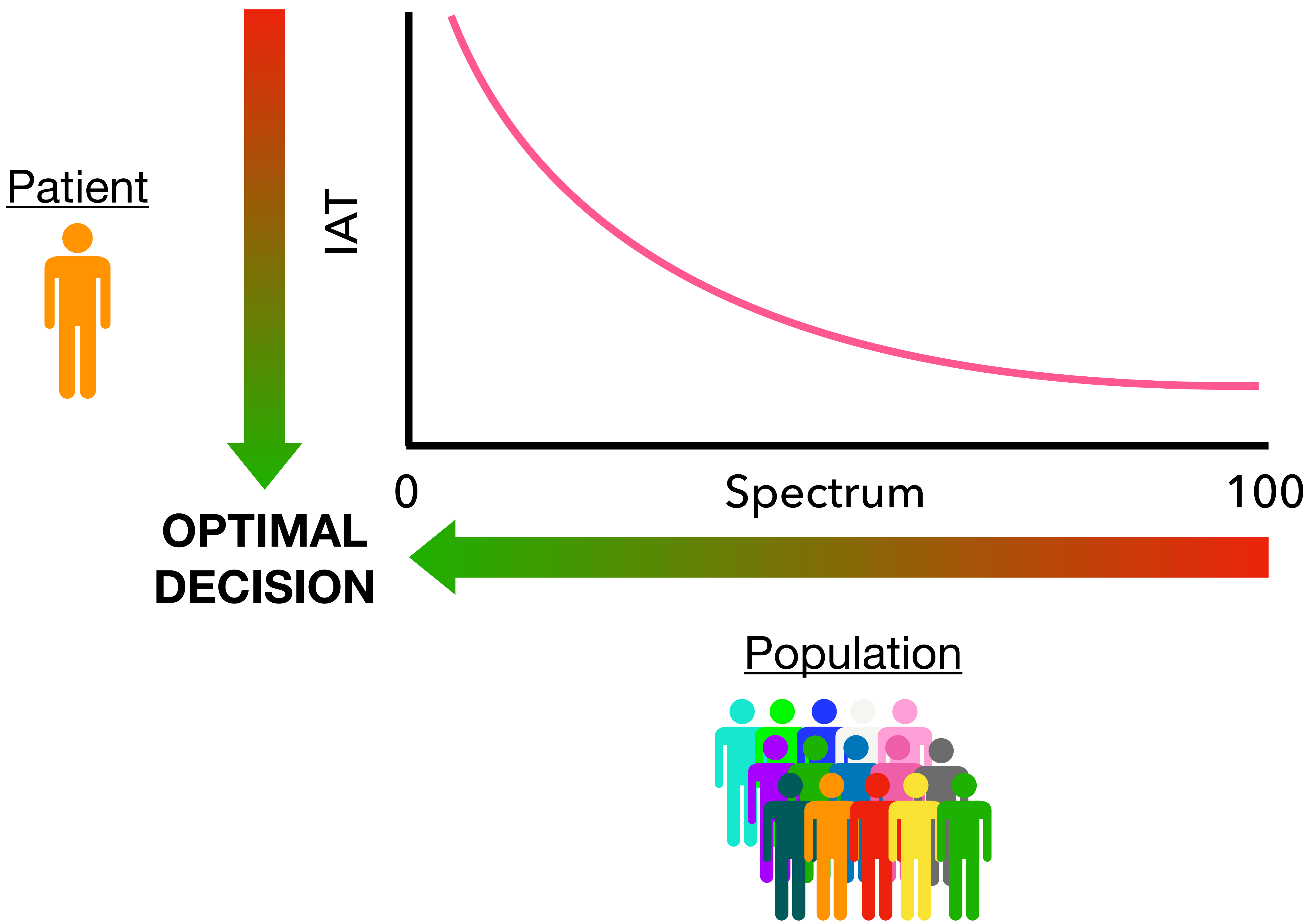


AHRQ, Four moments of Antibiotic Decision Making

Diagnostic support

Initiation of empiric therapy

Decision 2: What empiric therapy should I initiate?



Antimicrobial stewardship must balance the tension between the interests of the patient and the interest of the population

AHRQ, Four moments of Antibiotic Decision Making

Diagnostic support

Initiation of empiric therapy

- We looked at converting probabilities to decisions and found we could optimize the population ↔ patient conflict

SCIENCE TRANSLATIONAL MEDICINE

A decision algorithm to promote outpatient antimicrobial stewardship for uncomplicated urinary tract infection

Kanjilal, Sci Trans Med, 2020

71%

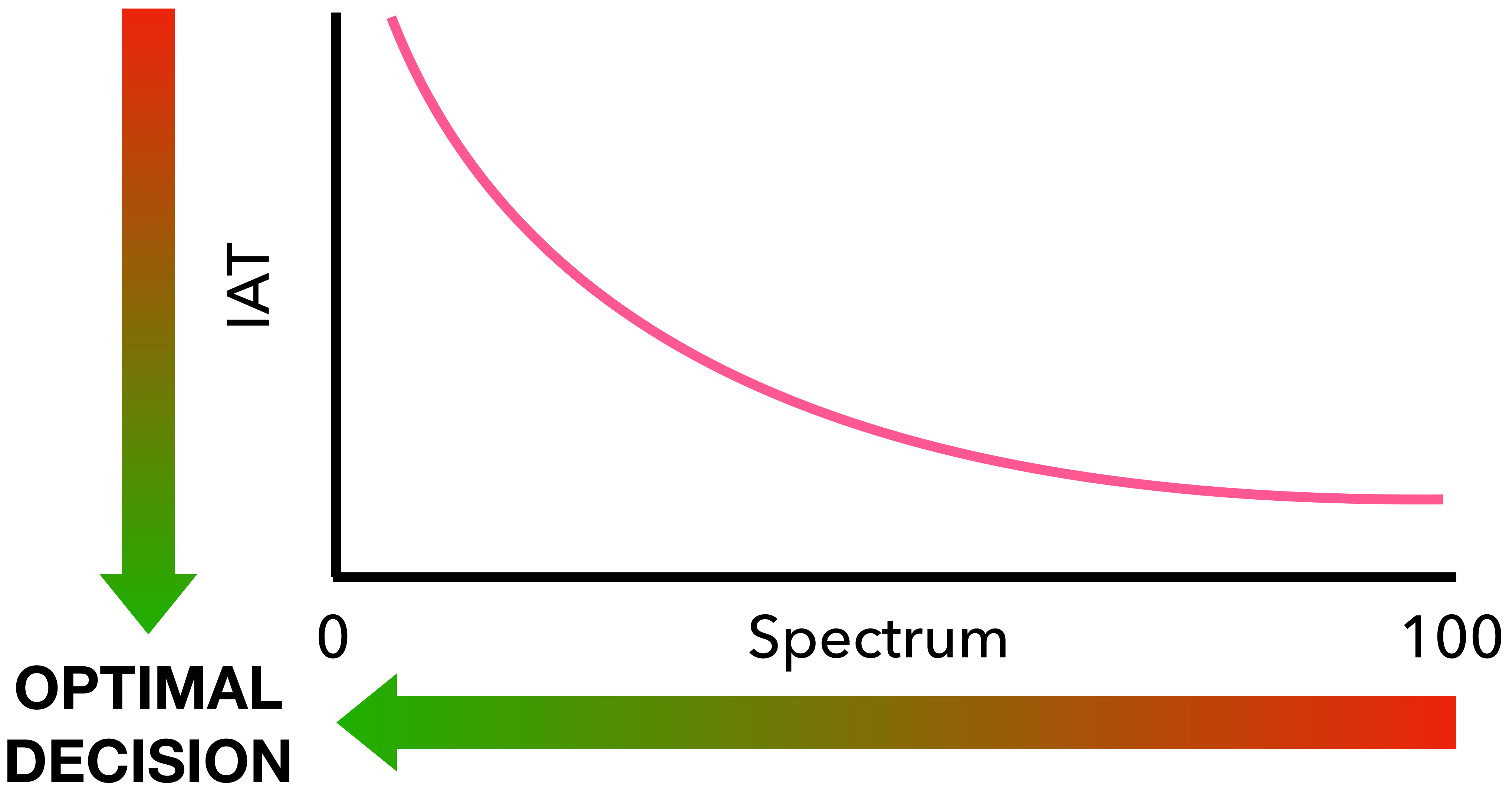
REDUCTION IN USE OF BROAD SPECTRUM ANTIBIOTICS

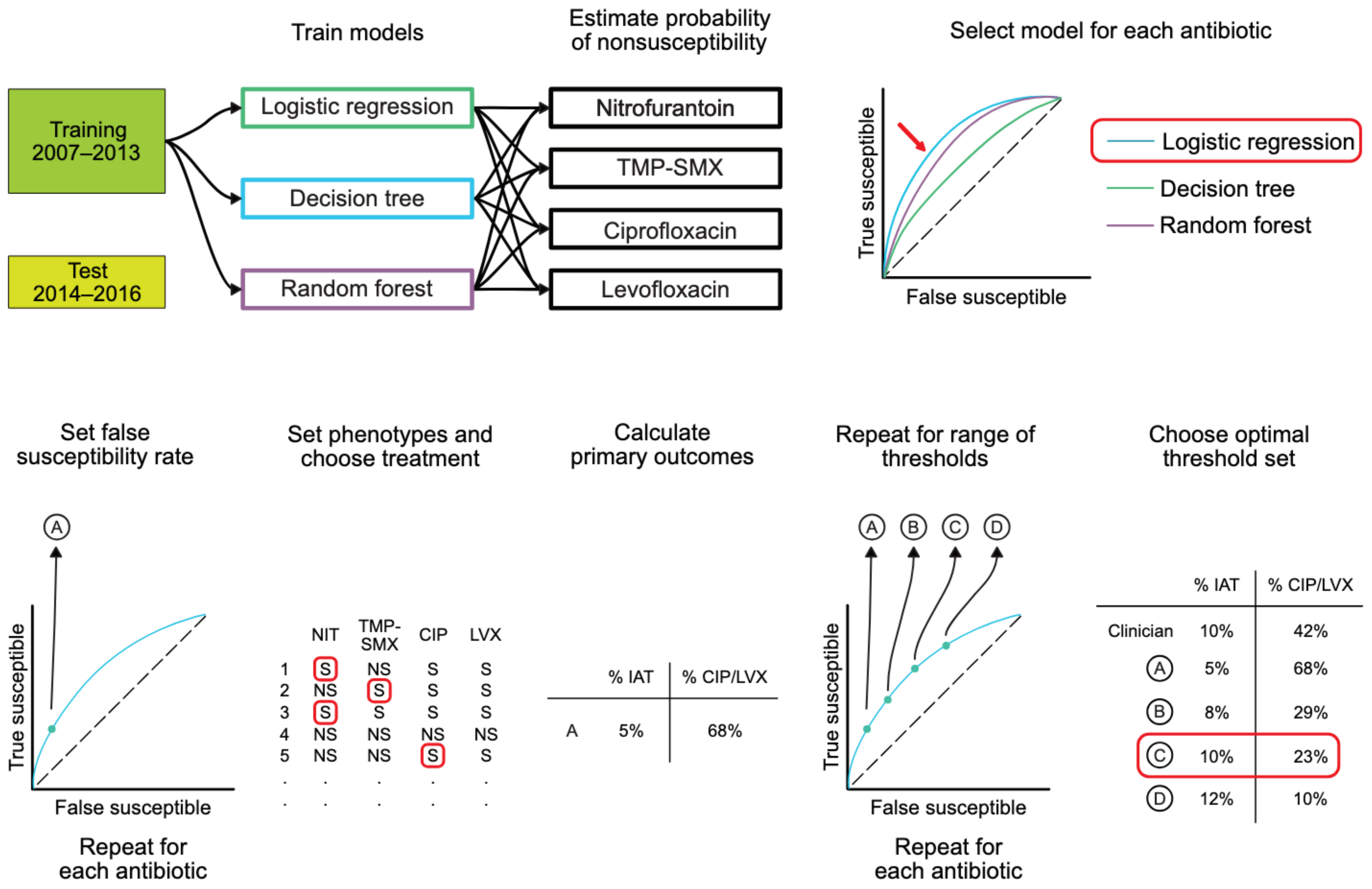
16%

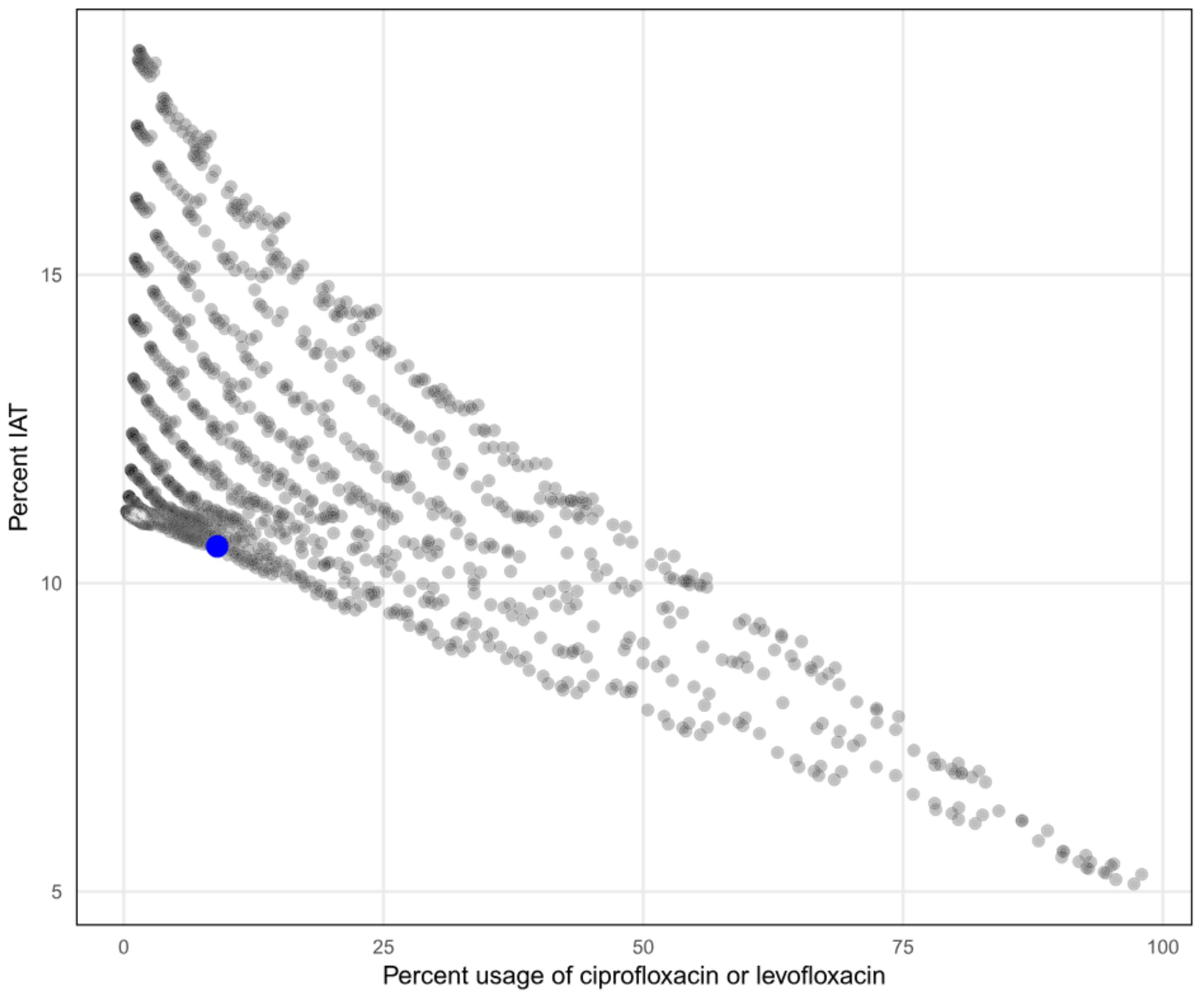
REDUCTION IN INAPPROPRIATE ANTIBIOTIC THERAPY

Diagnostic support

Initiation of empiric therapy







Kanjilal, Sci Trans Med, 2020

Tracking

Standardized Antimicrobial Administration Ratio (SAAR)

- The SAAR is a benchmarking tool to help hospitals monitor their antibiotic usage rates
- Depends heavily on facility-specific case mix
- NHSN uses 7 variables for adjustment

Ward type	Teaching status
Facility type	% of beds that are ICU
Number of beds	ICU beds
Mean LOS	

Patient characteristics not directly accounted for

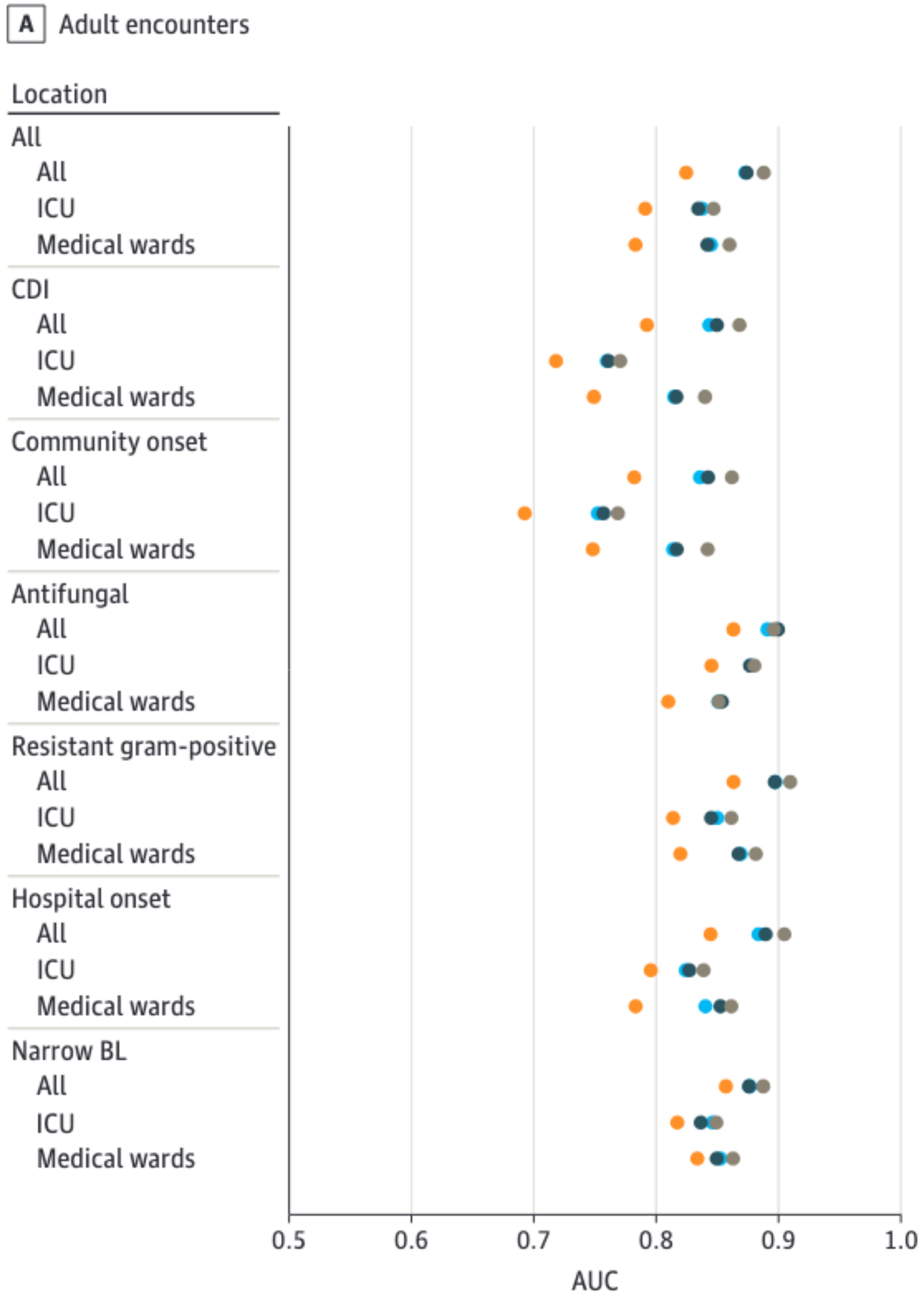
Tracking

Standardized Antimicrobial Administration Ratio (SAAR)

JAMA Network | Open

Development of a Machine Learning Model Using Electronic Health Record Data to Identify Antibiotic Use Among Hospitalized Patients

Moehring, JAMA Network Open, 2021



- Models that used patient-level features had good predictive power for overall antibiotic exposure and days of therapy (AUC > 0.8)
- Sophisticated models had better performance than simpler models

Though with an eventual plateau in performance

Action

Facility-specific treatment guidelines

- Development of facility-specific antibiotic use policies relies on integrating
  - Research literature
    - Randomized controlled trials
    - Observational studies
    - Pre-clinical studies
  - Cumulative antibiogram data
  - Local case mix & practice patterns

- Limited in number
- Heterogeneity in treatment effects

- Varying ability to control for confounding and selection bias

- Difficult to generalize to clinical settings

- Highly biased
- Subjective interpretation

- Highly variable
- Occasionally irrational

Action

Facility-specific treatment guidelines

- There is intense interest in leveraging EHR data to understand the impact of treatments or interventions on outcomes
- But there are many limitations

Data

Observation bias
Non-random missingness
Noisy

Study design

Selection bias
Immortal time bias
Confounding by indication

Action

Facility-specific treatment guidelines

- Many methodologic advances in the past 15 years to improve analysis of observational data for causal inference

Target trial emulation

Hernán, Am J Epi, 2016

- Design a hypothetical trial that answers the clinical question of interest
- Apply parameters to observational dataset

Components

- Eligibility criteria
- Treatment strategies
- Assignment procedures
- Follow-up period
- Outcome
- Causal contrast
- Analysis plan

Action

Facility-specific treatment guidelines

- Many methodologic advances in the past 15 years to improve analysis of observational data for causal inference

Doubly robust methods

Rose, Am J Epi, 2014

- Targeted maximum likelihood estimation
- Specifies 2 ML models
  - Propensity for treatment
  - Outcome

Only one of these needs to be consistent to have a consistent estimator

Highly simplified workflow

Model expected outcomes

↓

Estimate propensity for treatment

↓

Update initial expected outcome model

↓

Compute average treatment effect (ATE)

↓

Calculate error bounds

Re-evaluation of guidelines for UTI

JAMA Network Open

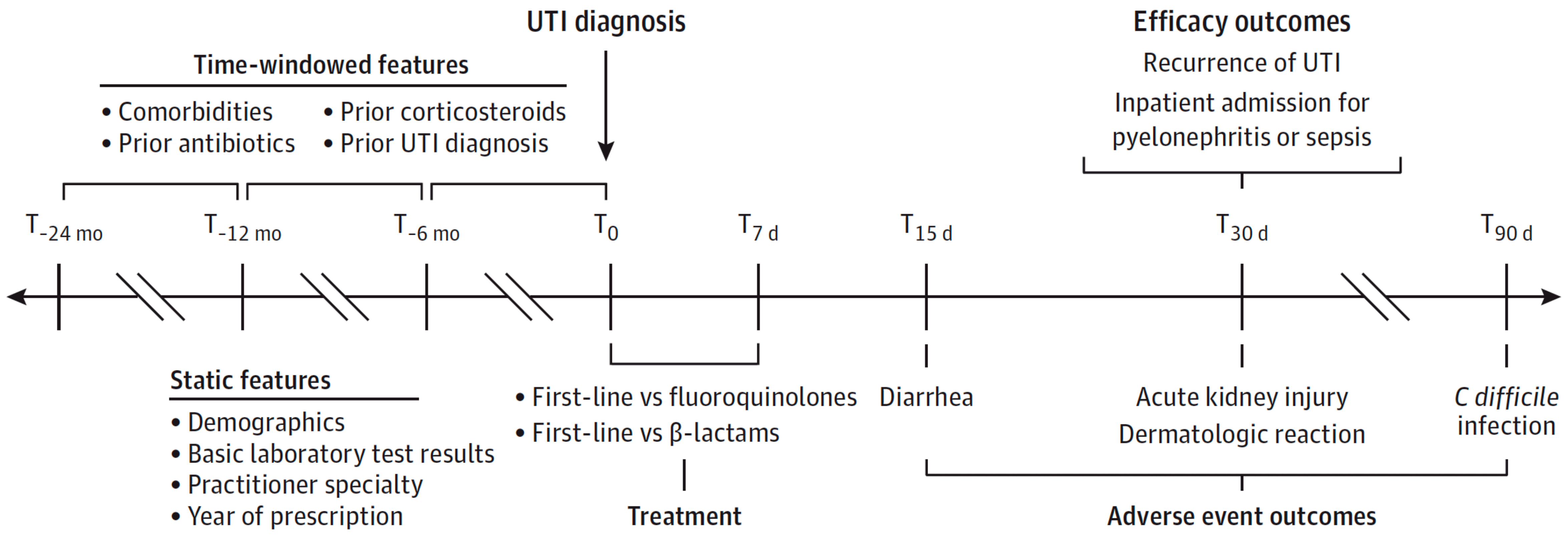
Use of Machine Learning to Assess the Management of Uncomplicated Urinary Tract Infection

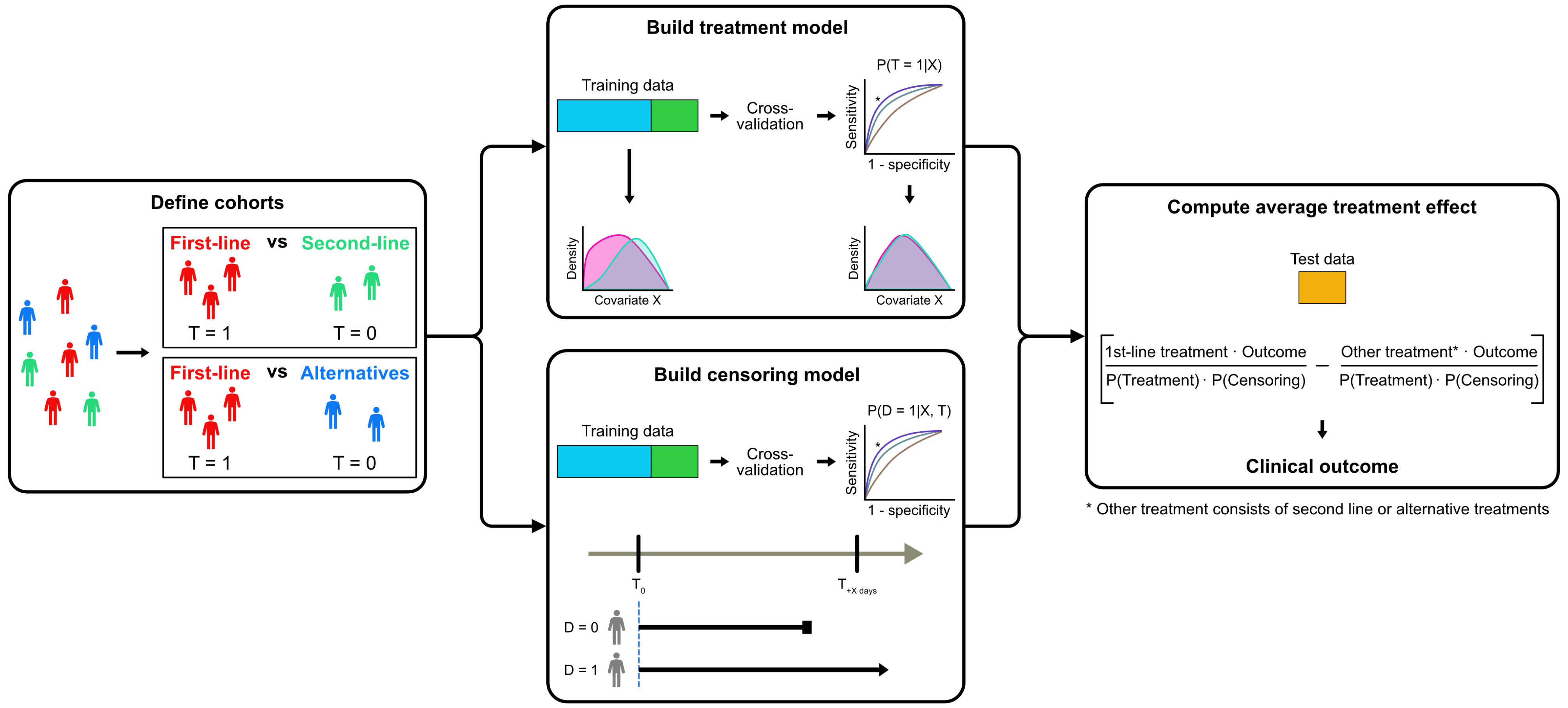
Jones, JAMA Network Open, forthcoming

- Things have changed since the IDSA released their guidelines for treatment of uncomplicated UTI in 2011
- We looked at a large claims database formatted into the OMOP common data model to see whether the recommendations still hold
- Used target trial emulation* combined with ML to adjust for confounding by indication and informative censoring

*with some minor deviations

Re-evaluation of guidelines for UTI





- Things have changed since the IDSA released their guidelines for treatment of uncomplicated UTI in 2011
- We looked at a large claims database formatted into the OMOP common data model to see whether the recommendations still hold
- Used target trial emulation* combined with ML to adjust for confounding by indication and informative censoring

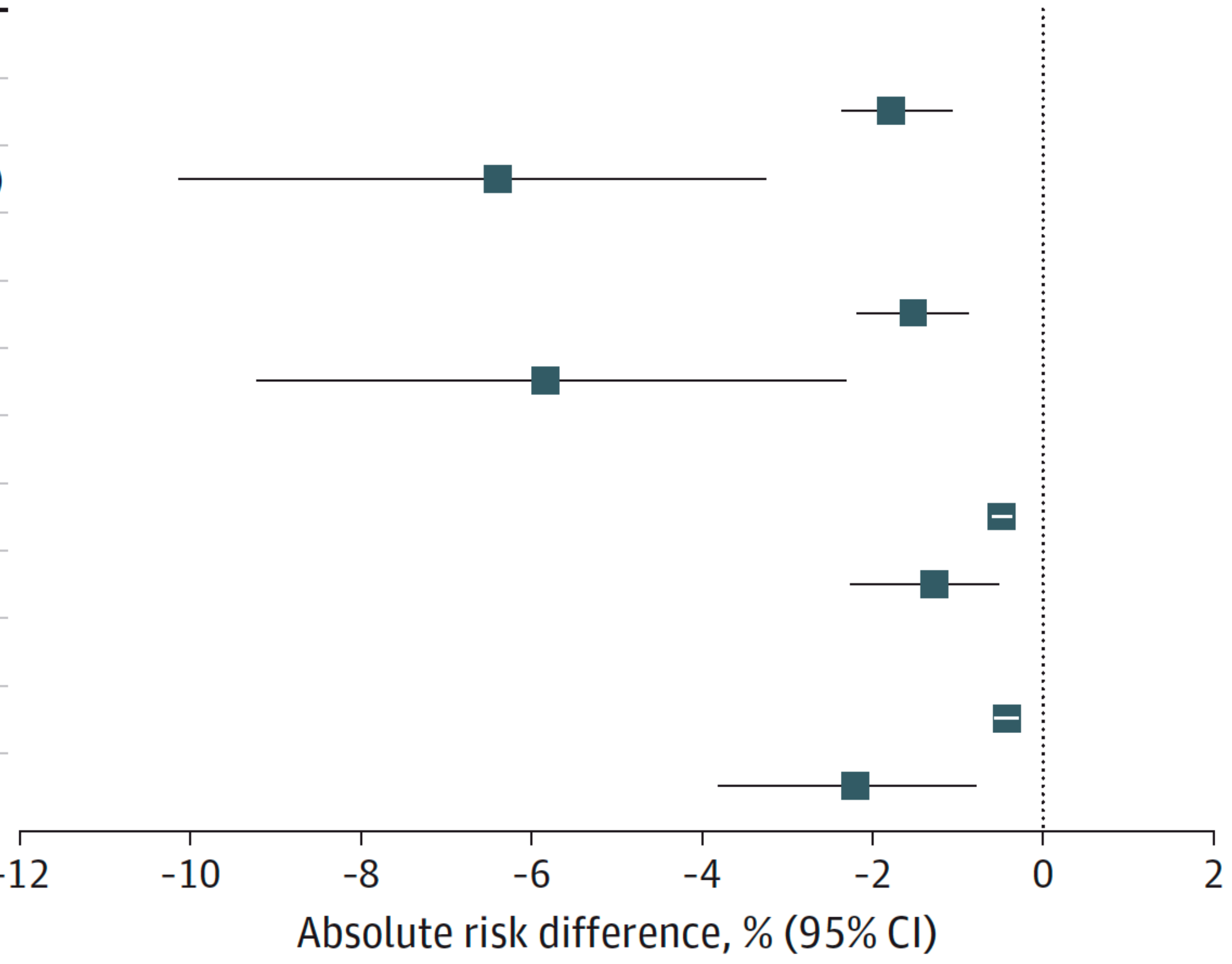
Jones, JAMA Network Open, forthcoming

*with some minor deviations

Re-evaluation of guidelines for UTI

A Revisits

Comparison	Absolute risk difference, % (95% CI)
Any	
First-line vs fluoroquinolones	-1.78 (-2.37 to -1.06)
First-line vs β-lactams	-6.40 (-10.14 to -3.24)
UTI	
First-line vs fluoroquinolones	-1.52 (-2.18 to -0.88)
First-line vs β-lactams	-5.83 (-9.22 to -2.31)
Pyelonephritis	
First-line vs fluoroquinolones	-0.49 (-0.60 to -0.37)
First-line vs β-lactams	-1.28 (-2.26 to -0.52)
Sepsis	
First-line vs fluoroquinolones	-0.43 (-0.57 to -0.30)
First-line vs β-lactams	-2.20 (-3.81 to -0.79)



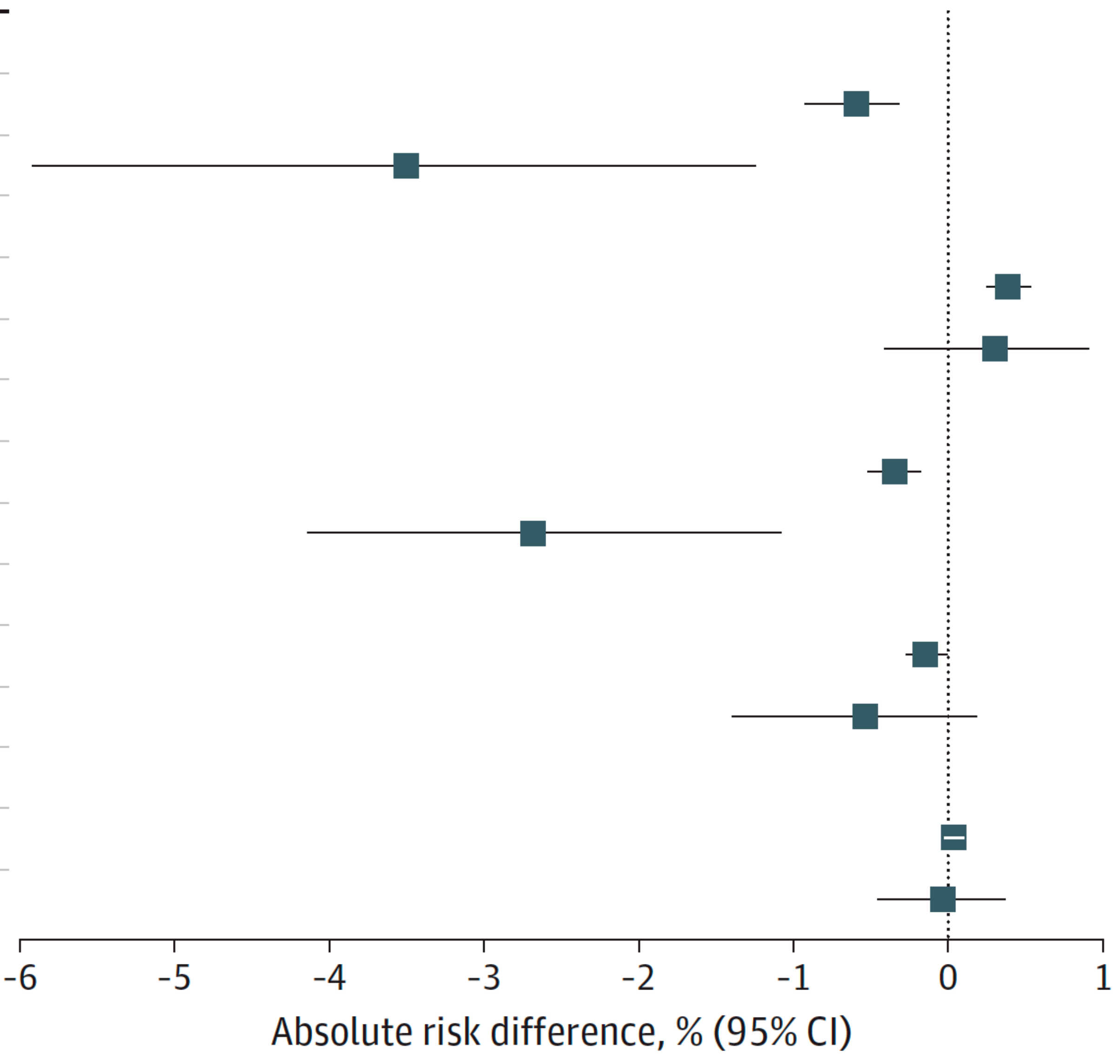
- 1st-line treatments performed well vs 2nd-line treatments and better than alternatives (code for β-lactams) with respect to primary outcomes (ie revisit within a month)

Jones, JAMA Network Open, forthcoming

Re-evaluation of guidelines for UTI

B Treatment-related adverse events

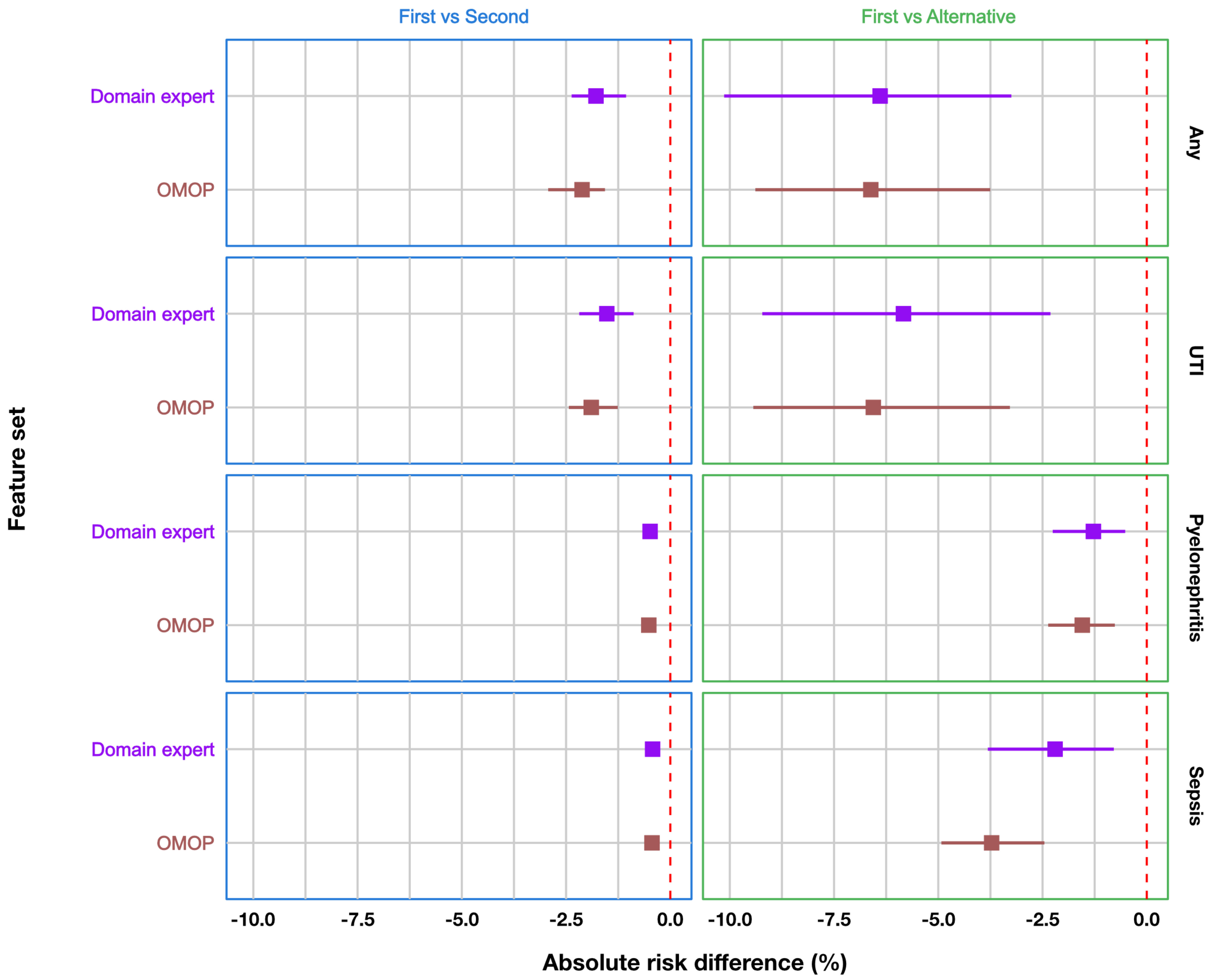
Comparison	Absolute risk difference, % (95% CI)
Any	
First-line vs fluoroquinolones	-0.60 (-0.93 to -0.32)
First-line vs β-lactams	-3.50 (-5.92 to -1.24)
Skin	
First-line vs fluoroquinolones	0.38 (0.25 to 0.53)
First-line vs β-lactams	0.30 (-0.42 to 0.91)
AKI	
First-line vs fluoroquinolones	-0.35 (-0.52 to -0.18)
First-line vs β-lactams	-2.68 (-4.14 to -1.08)
Diarrhea	
First-line vs fluoroquinolones	-0.15 (-0.27 to 0.00)
First-line vs β-lactams	-0.54 (-1.40 to 0.18)
C difficile	
First-line vs fluoroquinolones	0.04 (-0.03 to 0.10)
First-line vs β-lactams	-0.04 (-0.46 to 0.37)



- Adverse event (AE) rates were the same, if not slightly better, for first-line antibiotics vs second-line
- 1st-lines better than alternative treatments due ↓ risk of AKI
- The one exception is a slight increase in skin-related AEs (ie rash)

Jones, JAMA Network Open, forthcoming

Re-evaluation of guidelines for UTI



- Sensitivity analysis compared models derived using domain-expert knowledge to an automated feature extraction package (`OMOP-learn`)
- Results were similar!

Jones, JAMA Network Open, forthcoming

Future research agenda

Improving models

- Build multi center datasets
- Natural language processing
- Overcome existing and emerging technical challenges

Preparing for deployment

- Build data pipelines for continuous model training
- Infrastructure for model oversight
- Implementation science to optimize delivery of information to ASP leadership and end users

Large language models for improved risk stratification

Human-AI interaction studies

Data shift
Algorithmic fairness
Feed-forward loops

Pragmatic trials

Federated learning & common data models

EHR-agnostic decision support tools

Learning to defer
Risk calibration

Thank you

DEPARTMENT OF POPULATION MEDICINE

HARVARD MEDICAL SCHOOL

Harvard Pilgrim Health Care Institute

Michael Klompas

Chanu Rhee

Ted Pak

David Sontag

Helen Zhou

Michael Oberst

Sooraj Boominathan

Sontag lab (+ alumni)