Lessons Learned Implementing Machine Learning in Workflows

ROADMAP

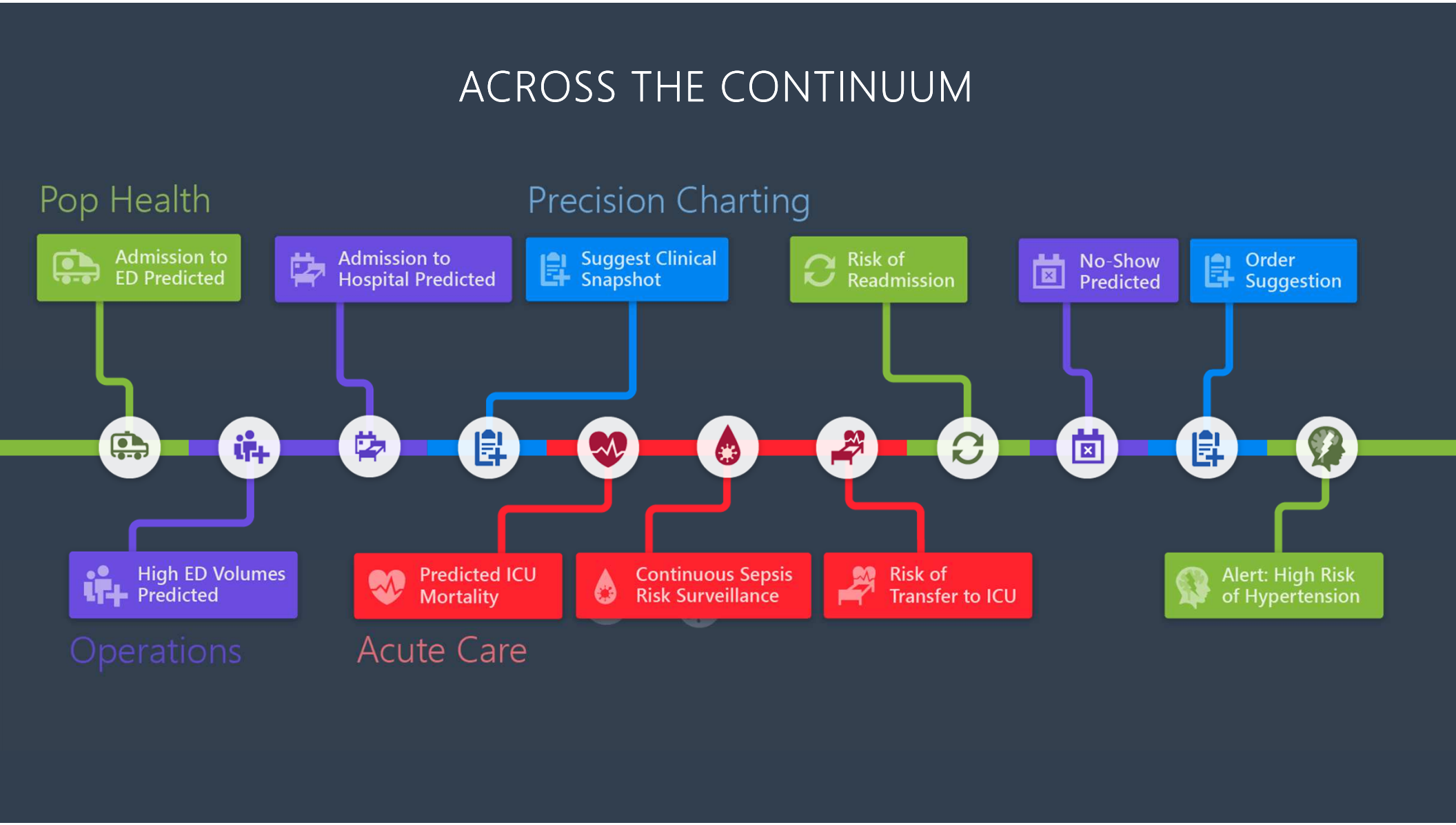
Background

Embed in Workflows

Measure & Monitor

Future Applications

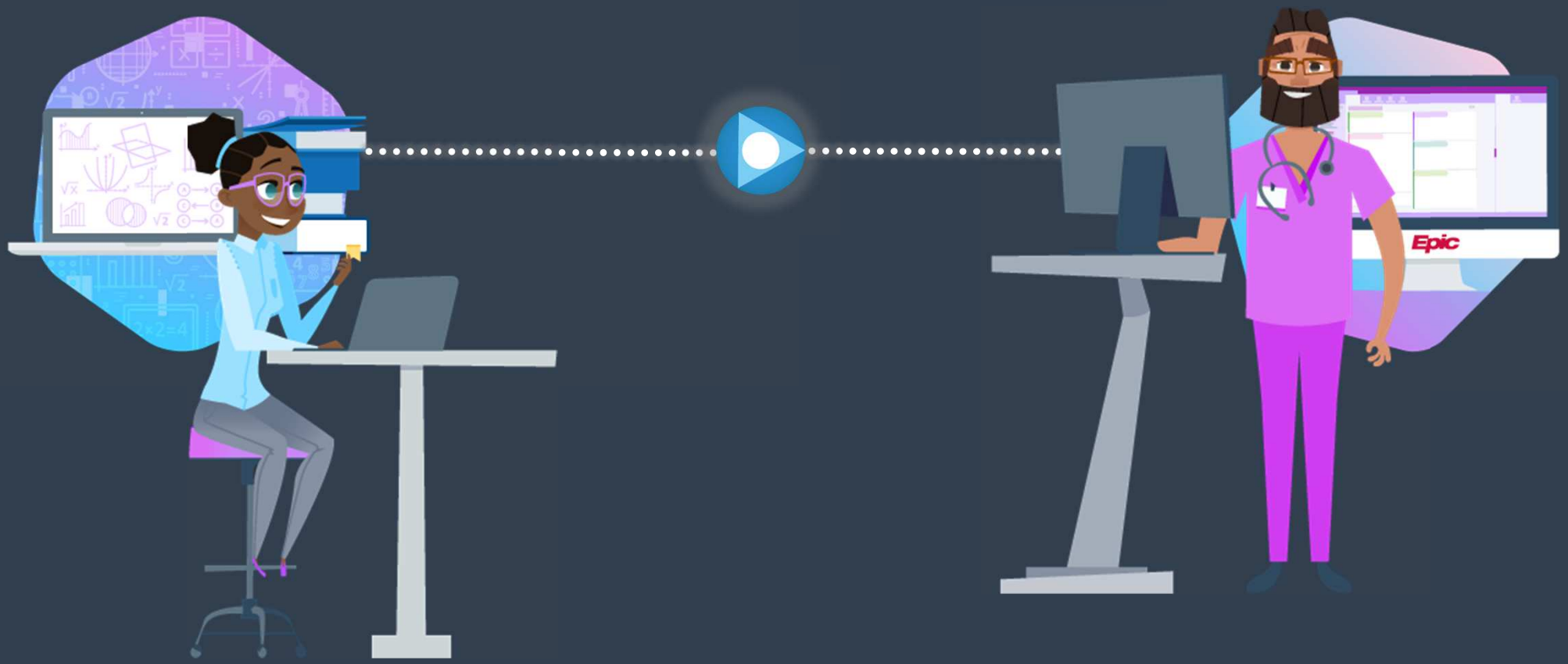
ACROSS THE CONTINUUM



Rule #1: Don't be afraid to launch a product without machine learning.

Google – Rules of Machine Learning

LAST MILE CHALLENGE



EPIC MACHINE LEARNING

160

organizations with models in production

17

machine learning **models**

Better Math:
Machine Learning Algorithms

More Data:
Weather, Local Events, Census

+

Adapt:
To local populations

Enhanced Decision-making.

Embed in Workflows

DESIGNING A WORKFLOW

Knowledge Driven

MDs & RNs

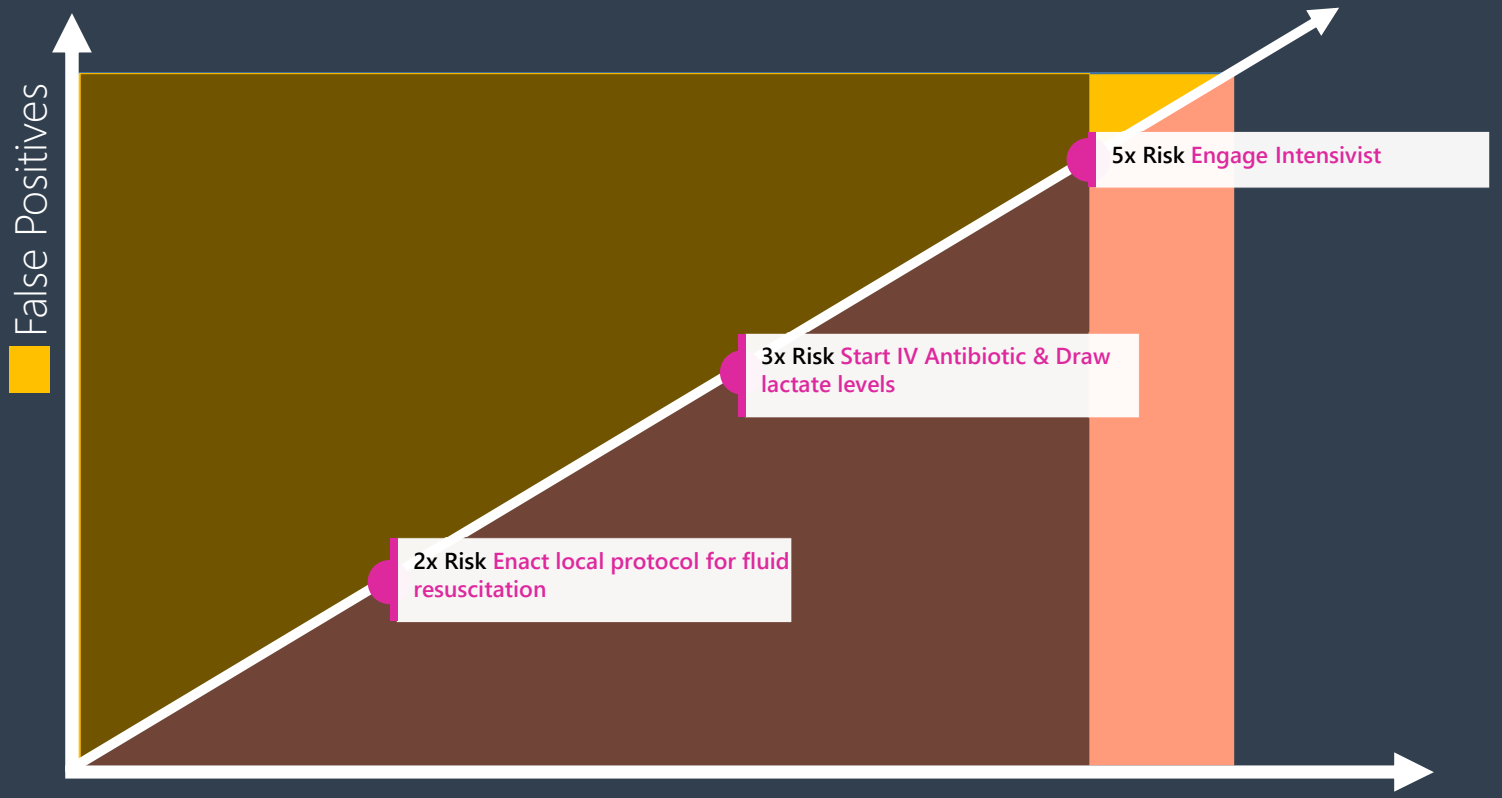
Domain knowledge
Workflow focus

Data Driven

Data Scientists

Statistical approach
Data reveals insight

INTERVENTION SLIDING SCALE



DEFINING INTERPRETABILITY

Explainable AI (XAI)

Understand results of the model

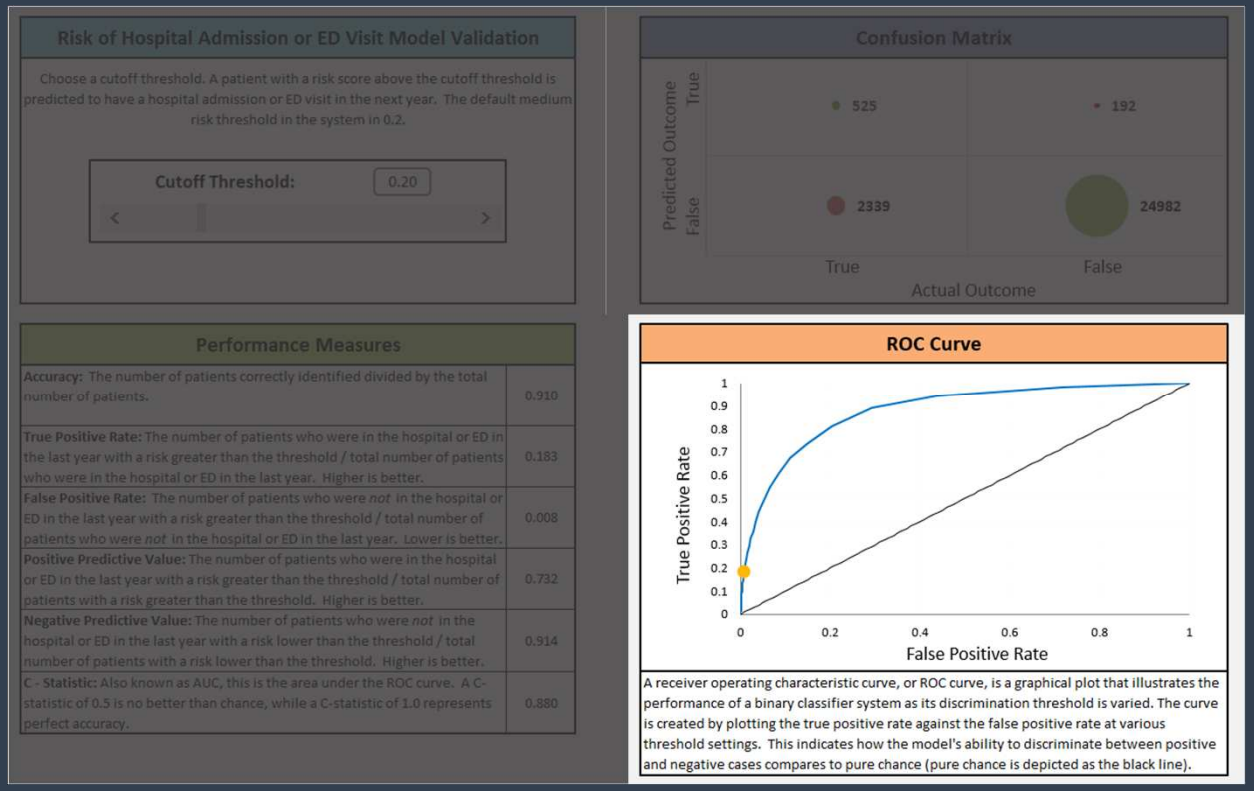
© 2018 Epic Systems Corporation. Confidential.

PRESENTING INSIGHTS

© 2018 Epic Systems Corporation. Confidential.

Measure & Monitor

DEFINING MODEL ACCURACY



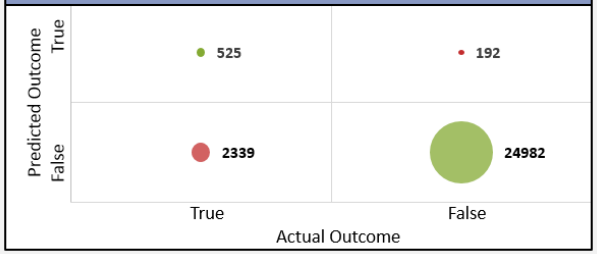
DEFINING MODEL ACCURACY

Risk of Hospital Admission or ED Visit Model Validation

Choose a cutoff threshold. A patient with a risk score above the cutoff threshold is predicted to have a hospital admission or ED visit in the next year. The default medium risk threshold in the system in 0.2.

Cutoff Threshold: 0.20

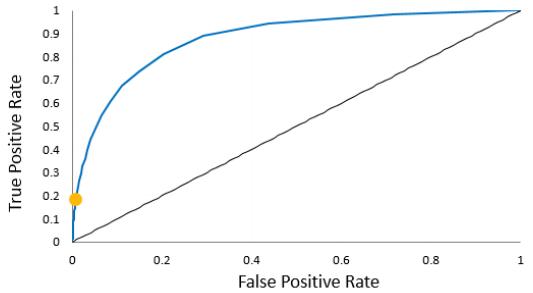
Confusion Matrix



Performance Measures

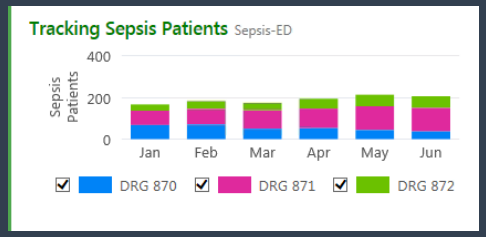
Accuracy: The number of patients correctly identified divided by the total number of patients.	0.910
True Positive Rate: The number of patients who were in the hospital or ED in the last year with a risk greater than the threshold / total number of patients who were in the hospital or ED in the last year. Higher is better.	0.183
False Positive Rate: The number of patients who were *not* in the hospital or ED in the last year with a risk greater than the threshold / total number of patients who were *not* in the hospital or ED in the last year. Lower is better.	0.008
Positive Predictive Value: The number of patients who were in the hospital or ED in the last year with a risk greater than the threshold / total number of patients with a risk greater than the threshold. Higher is better.	0.732
Negative Predictive Value: The number of patients who were *not* in the hospital or ED in the last year with a risk lower than the threshold / total number of patients with a risk lower than the threshold. Higher is better.	0.914
C - Statistic: Also known as AUC, this is the area under the ROC curve. A C-statistic of 0.5 is no better than chance, while a C-statistic of 1.0 represents perfect accuracy.	0.880

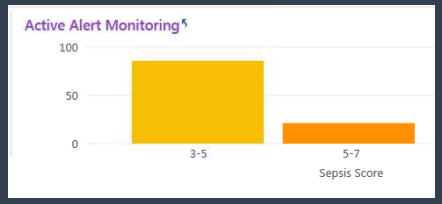
ROC Curve



A receiver operating characteristic curve, or ROC curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate against the false positive rate at various threshold settings. This indicates how the model's ability to discriminate between positive and negative cases compares to pure chance (pure chance is depicted as the black line).

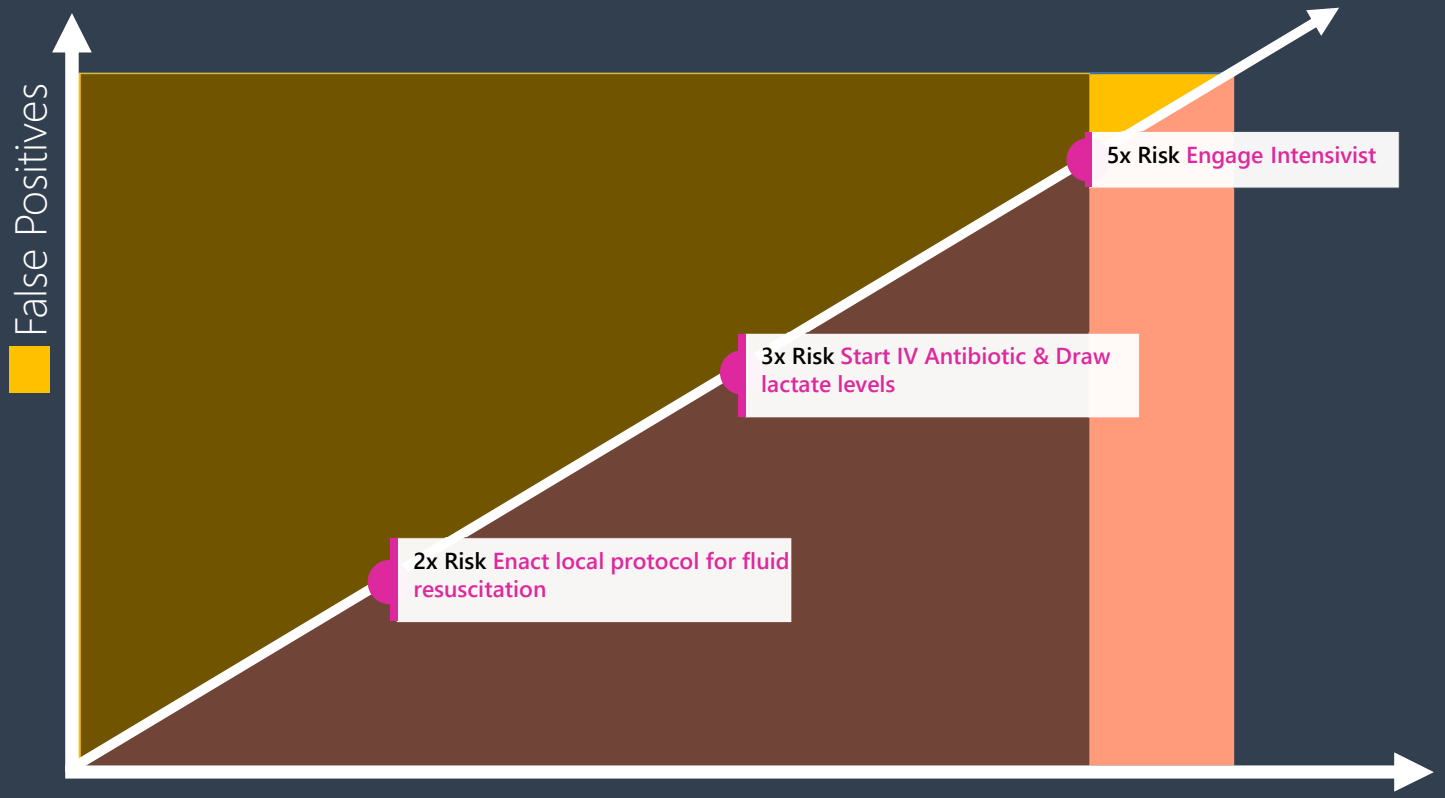
MONITORING OUTCOMES





© 2018 Epic Systems Corporation. Confidential.

OPTIMIZING THRESHOLDS



Future Applications

EXPANDING MACHINE LEARNING

Resource Maximization

Streamline operations

Helping Physicians

real-time guidance and recommendations

HELPING PHYSICIANS

Shopper One
Shopper Two

HELPING PHYSICIANS

ONGOING MONITORING

CONCLUSIONS

Design a workflow to **maximize benefit to end users**

Establish a strategy for **optimizing thresholds & monitoring impact**

Be thoughtful in ongoing **model management capabilities**