An Improved Multi-Output Gaussian Process RNN with Real-Time Validation for Early Sepsis Detection

<u>Joe Futoma</u>
Duke University
Dept. of Statistical Science

Sanjay Hariharan, Katherine Heller (Duke Dept. of Statistical Science)
Mark Sendak, Nathan Brajer (Duke Institute for Health Innovation)
Meredith Clement, Armando Bedoya, Cara O'Brien (Duke Dept. of Medicine)

Sepsis

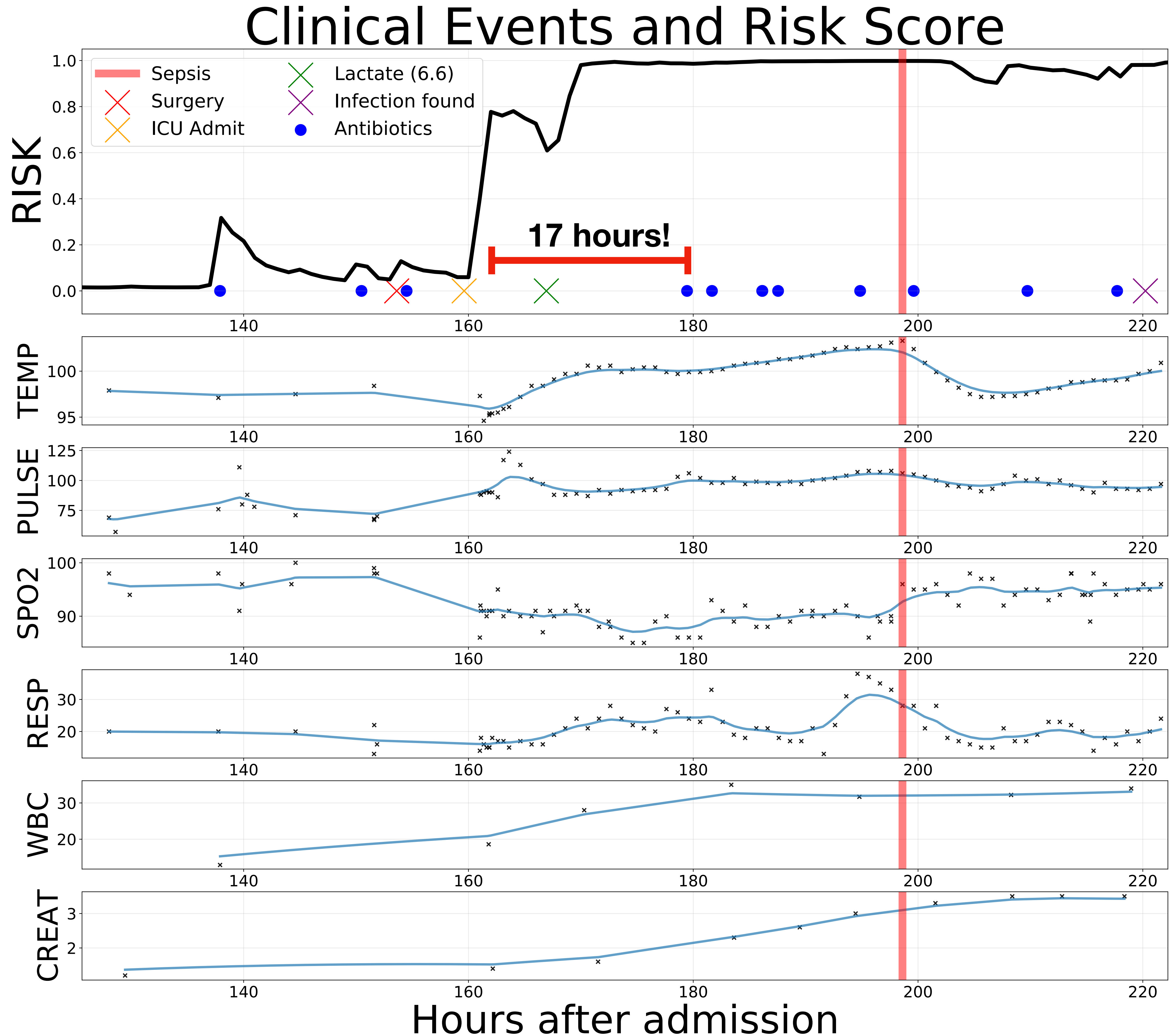
- Life-threatening complication from infection.
- 750,000+ new sepsis cases each year in US; high mortality (**30-50%**).
- Without intervention, progress to septic shock, organ failure, death.
- **Early identification is key:**
    - Earlier treatment associated with improved outcomes.
    - **We know what to do, if we know it's there!**
- **Early identification is hard:**
    - No clear time of onset, no reliable biomarker (yet).


The NEW ENGLAND JOURNAL of MEDICINE

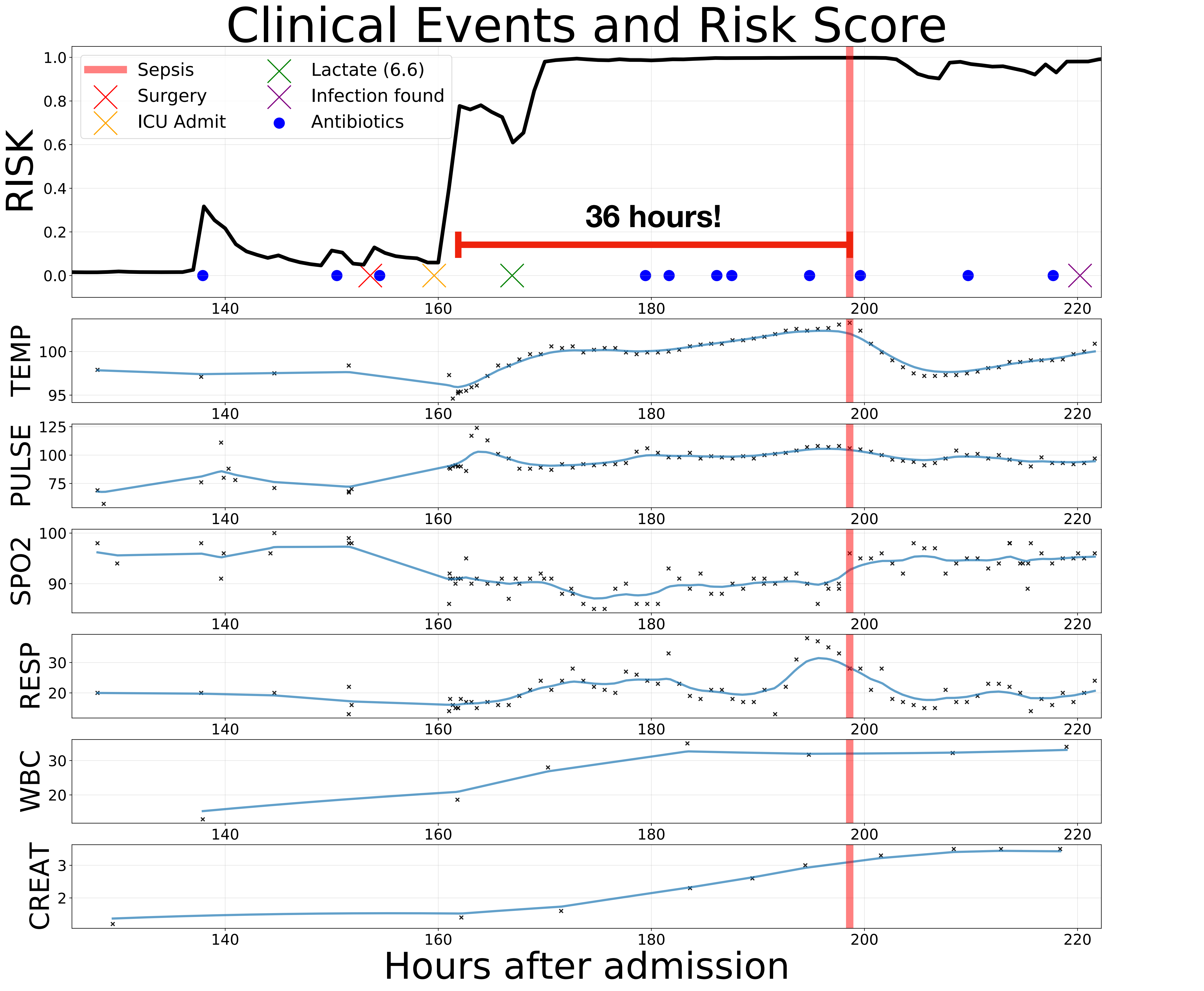
ORIGINAL ARTICLE

Time to Treatment and Mortality during Mandated Emergency Care for Sepsis

Christopher W. Seymour, M.D., Foster Gesten, M.D., Hallie C. Prescott, M.D., Marcus E. Friedrich, M.D., Theodore J. Iwashyna, M.D., Ph.D., Gary S. Phillips, M.A.S., Stanley Lemeshow, Ph.D., Tiffany Osborn, M.D., M.P.H., Kathleen M. Terry, Ph.D., and Mitchell M. Levy, M.D.



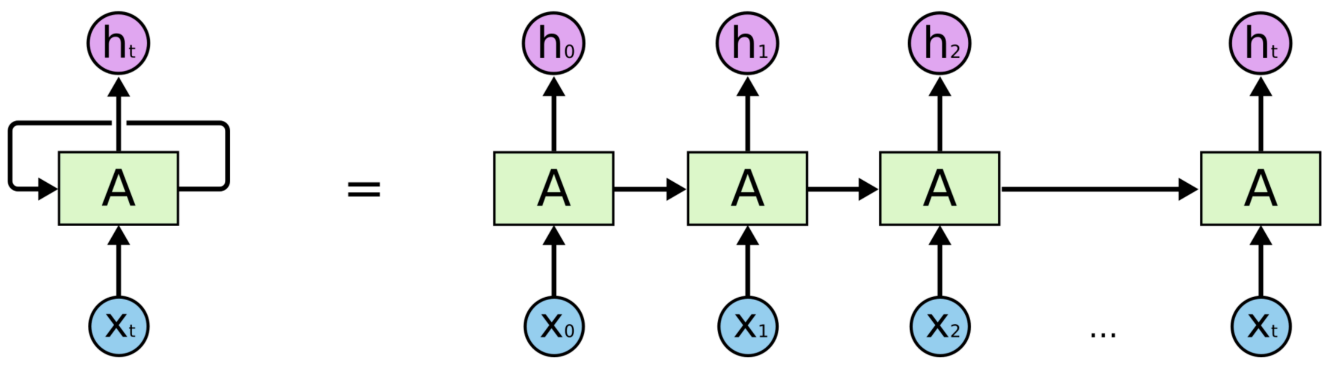
Clinical Events and Risk Score
RISK
Sepsis
Surgery
ICU Admit
Lactate (6.6)
Infection found
Antibiotics
17 hours!
TEMP
PULSE
SPO2
RESP
WBC
CREAT
Hours after admission



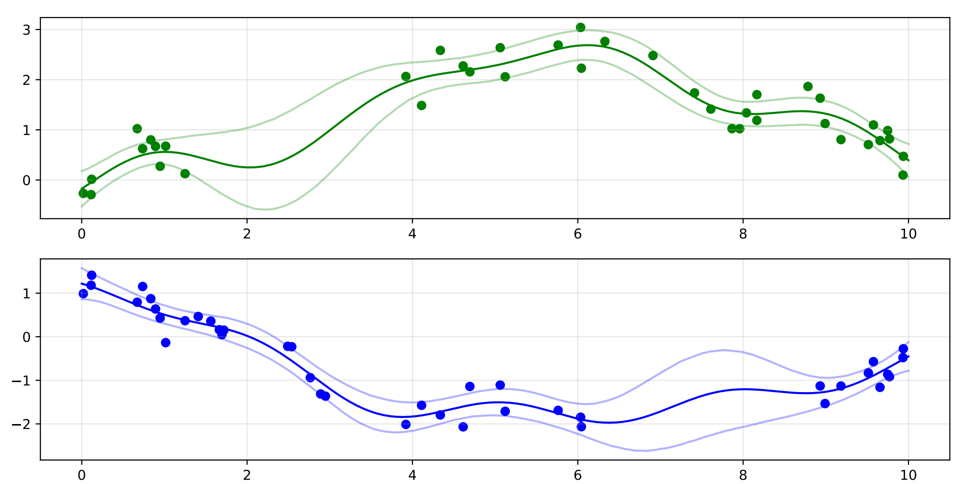
Clinical Events and Risk Score
RISK
Sepsis
Surgery
ICU Admit
Lactate (6.6)
Infection found
Antibiotics
36 hours!
TEMP
PULSE
SPO2
RESP
WBC
CREAT
Hours after admission

Model Main Idea

- Goal: detect onset of sepsis before it occurs.
- Data:
    1. Physiological time series (labs/vitals).
    2. Baseline admission info/comorbidities.
    3. Medication administration times.
- Multivariate time series classification: update a risk score (probability encounter is / will become septic).
- **Recurrent Neural Networks (RNNs)**: flexible functions, rich representational power for sequences of arbitrary length. But:
    - Requires a **complete** dataset with no missing values.
    - Requires **regularly** spaced inputs.
- **Multi-output Gaussian Processes**: model for multivariate time series.
    - Seamlessly handles **irregularly** spaced observation times.
    - Imputes missing values on a regular grid, along with an estimate of **uncertainty**.

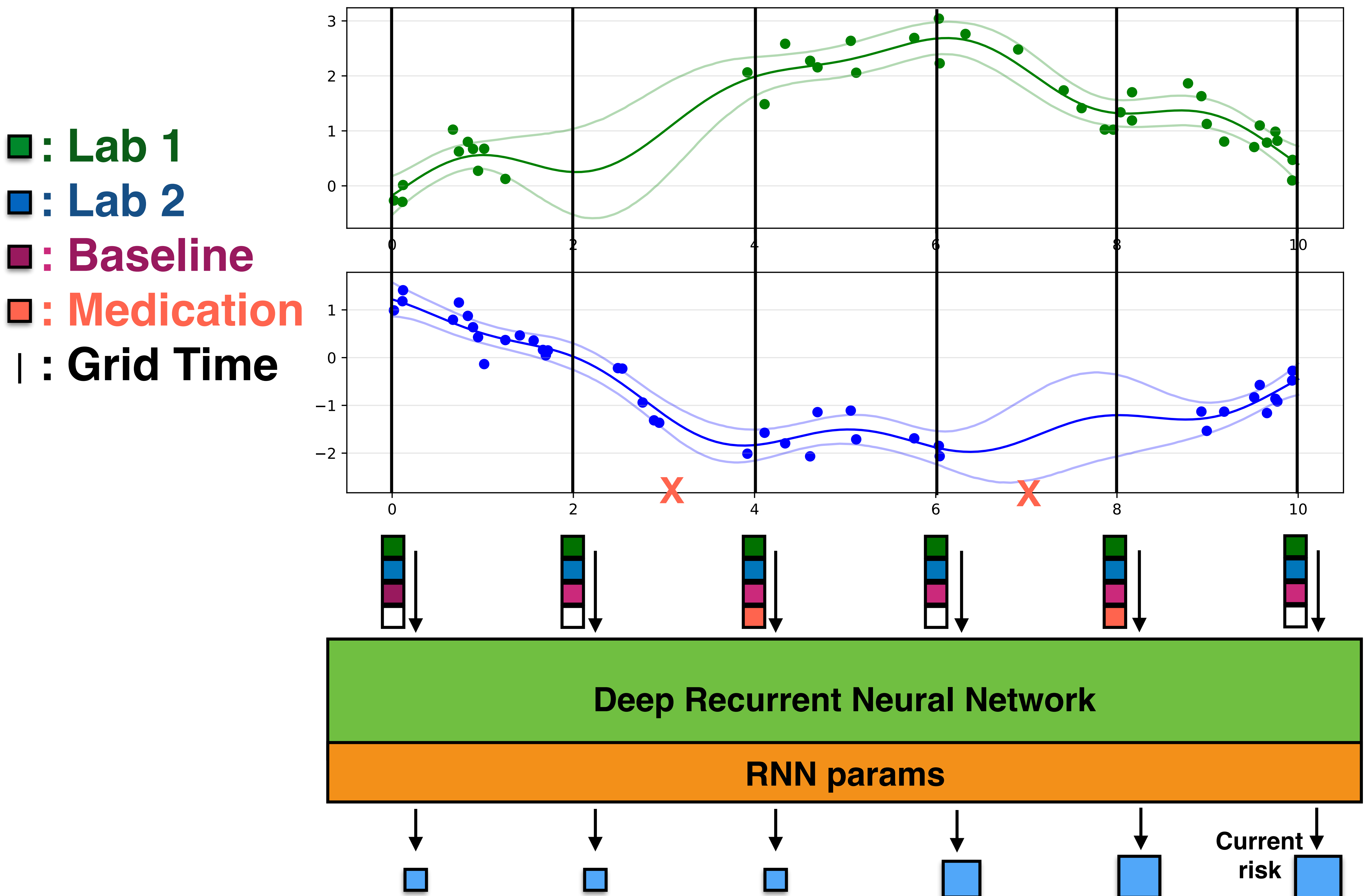


+



Futoma, Hariharan, Heller
ICML 2017

Model Schematic

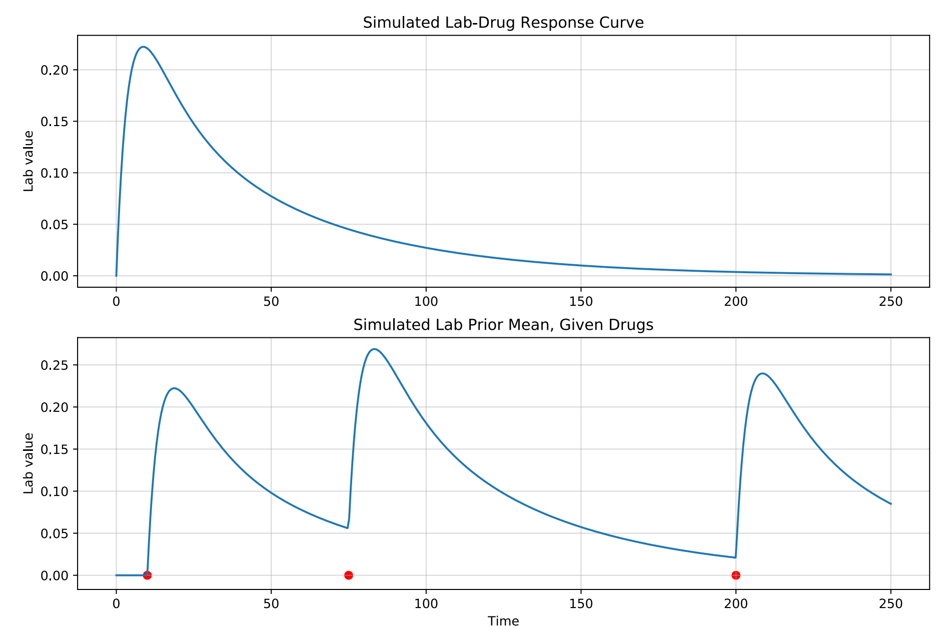


Improving the MGP

- <u>More flexible mean function</u>: extend zero-mean prior means to depend on administration of drugs.

$$\mu_m(t) = \sum_{p=1}^{P} \sum_{t_p<t} f_{pm}(t - t_p)$$

$$f_{pm}(t) = \sum_{l=1}^{L} \alpha_{lpm} e^{-\beta_{lpm} t}$$

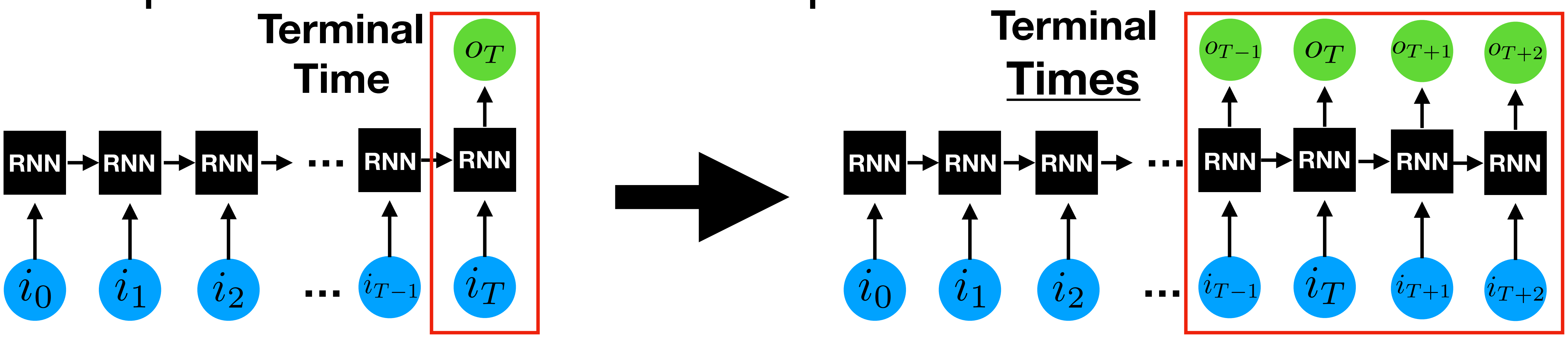


- <u>More flexible kernel</u>: extend assumption of separable covariance function in Multitask GP (Q=1).

$$\mathrm{cov}(f_{im}(t), f_{im'}(t')) = \sum_{q=1}^{Q} K_q^M(m, m') k_q^t(t, t')$$

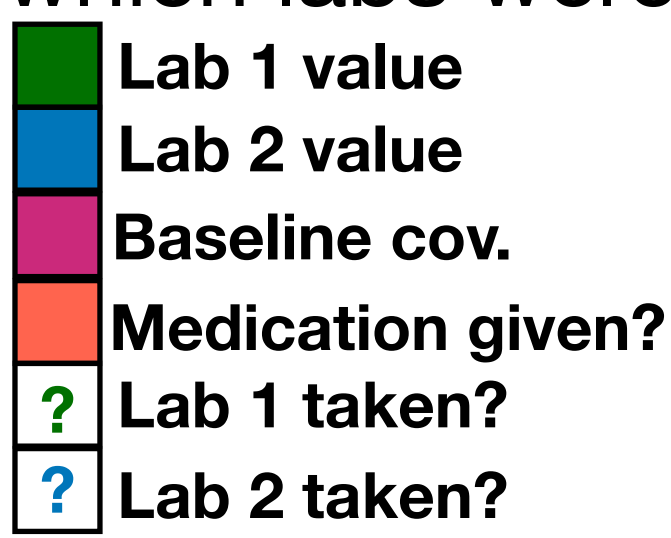
Improving the RNN

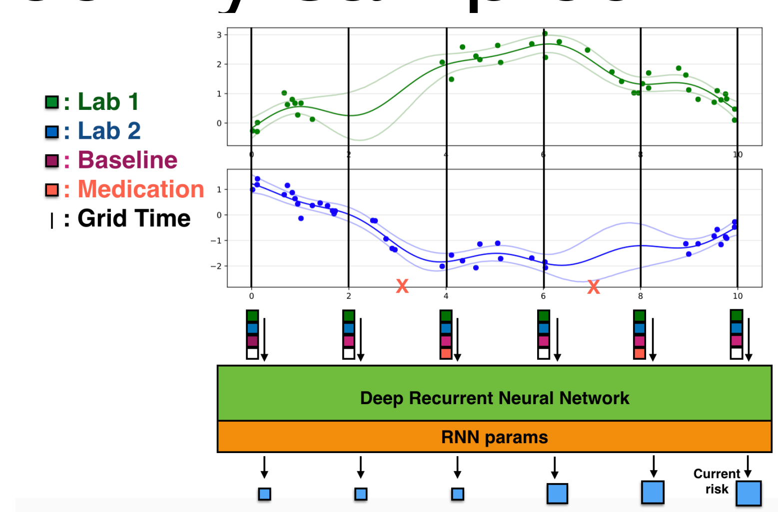
- Target replication: classification loss depends on multiple outputs instead of last time point.



- Missingness indicators: directly model pattern of informative missingness in the RNN by passing indicator vector with which labs were recently sampled.

Input Vector to RNN:

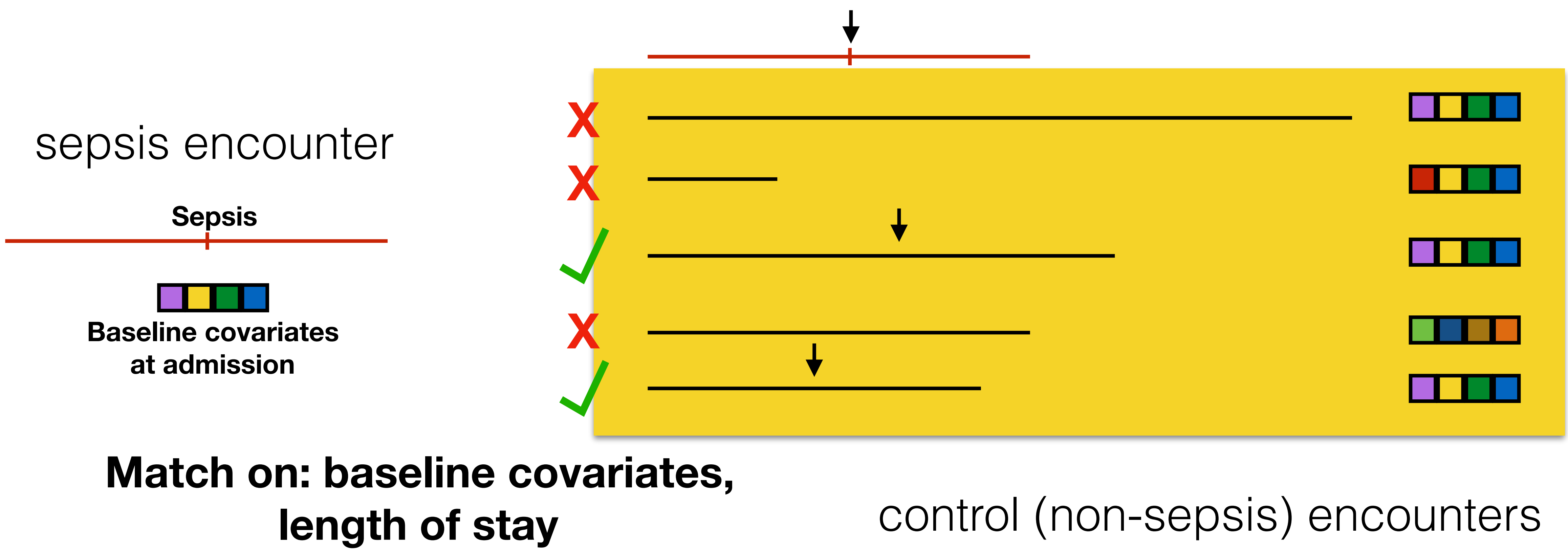




Dataset

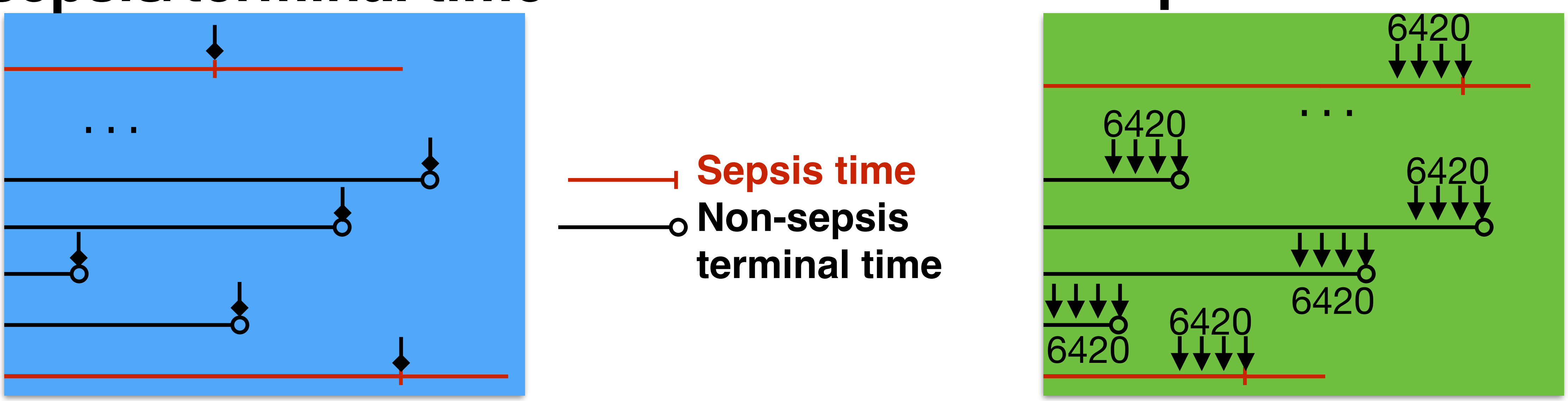
- **51,697** inpatient encounters at Duke Hospital over 18 months, 21.4% with a sepsis event; no specific inclusion/exclusion criteria.
- **34** physiological variables (5 vitals, 29 labs).
    - At least one value for each vital in 99% of encounters.
    - Some labs rarely measured (2-4%), most measured 20-80% of the time.
- **35** baseline covariates (e.g. age, transfer status, comorbidities).
- **8** medication classes (e.g. antibiotics, opioids, heparins).
- Mean length of stay 121.7 hours (sd: 108.1); highly variable.

Case Control Matching & Experimental Setup



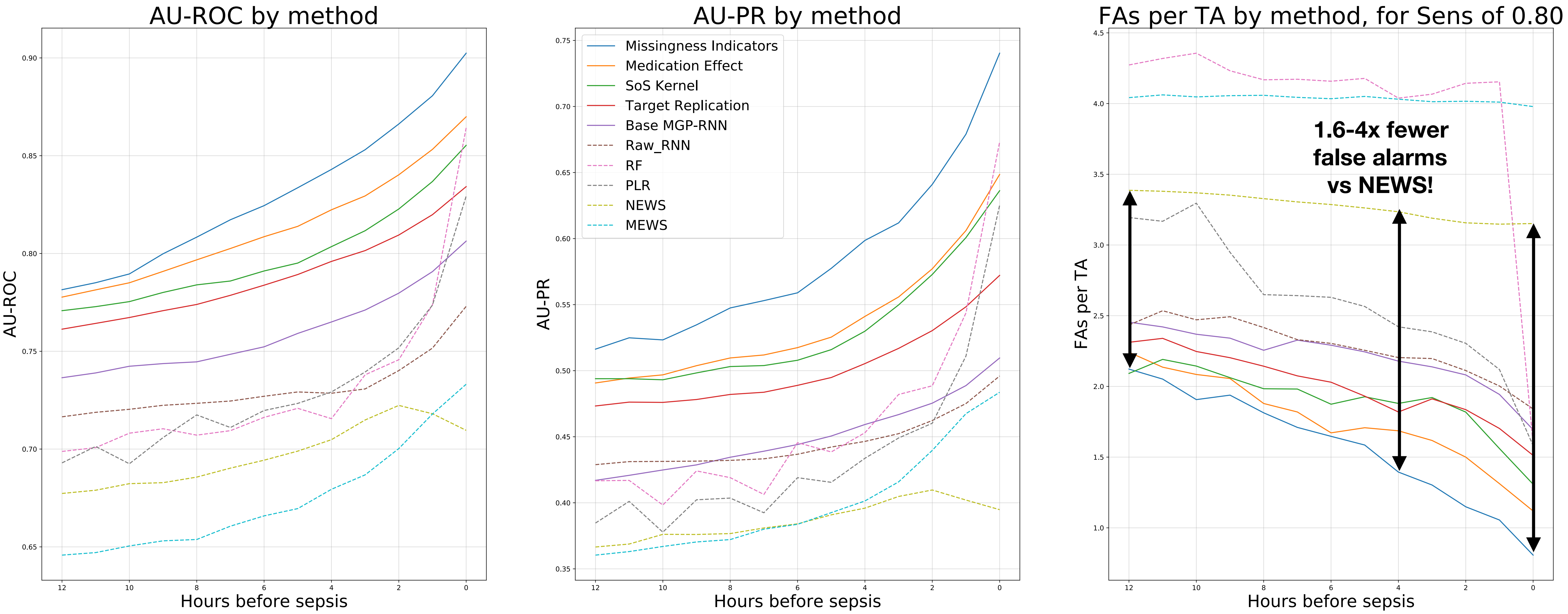
Train: up to sepsis/terminal time

Test: vary # hours from sepsis/terminal time



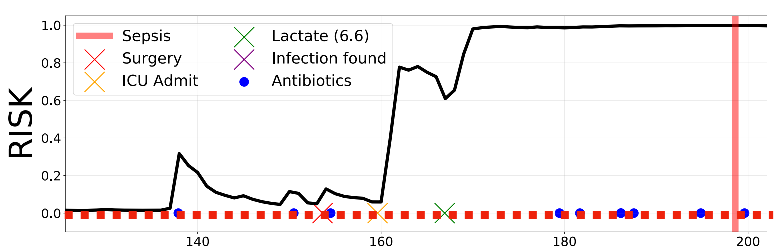
“Lookback” Results

- **Solid lines:** 5 MGP-RNN versions, adding in each extension.
  - Base; Target replication; SoS kernel; Med effects; Missing Indicators
- **Dashed lines:** RNN, Random Forest, Penalized Logistic Regression (impute with last observation carried forward); NEWS, MEWS clinical scores

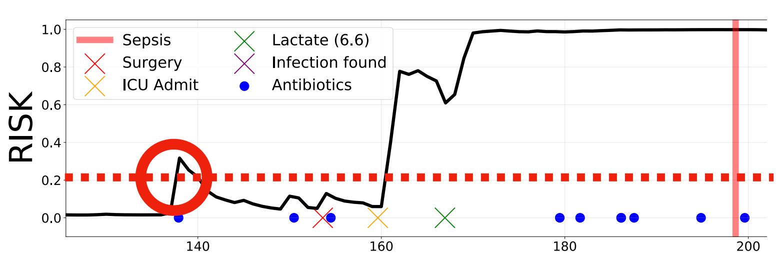


Real-Time Validation

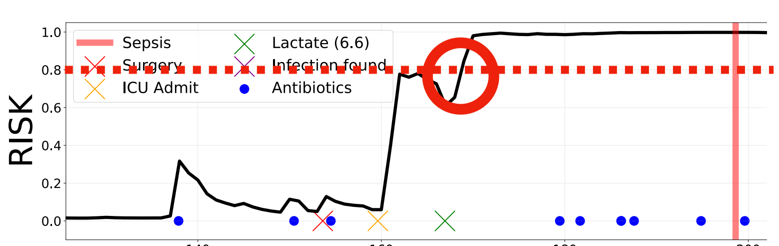
- Idea: Previous “Lookback” Results require alignment of time series at some terminal time.
- In practice, terminal time / sepsis time not known in advance.
- Want evaluation that more closely mirrors actual use case.



False Negative (Never detected)



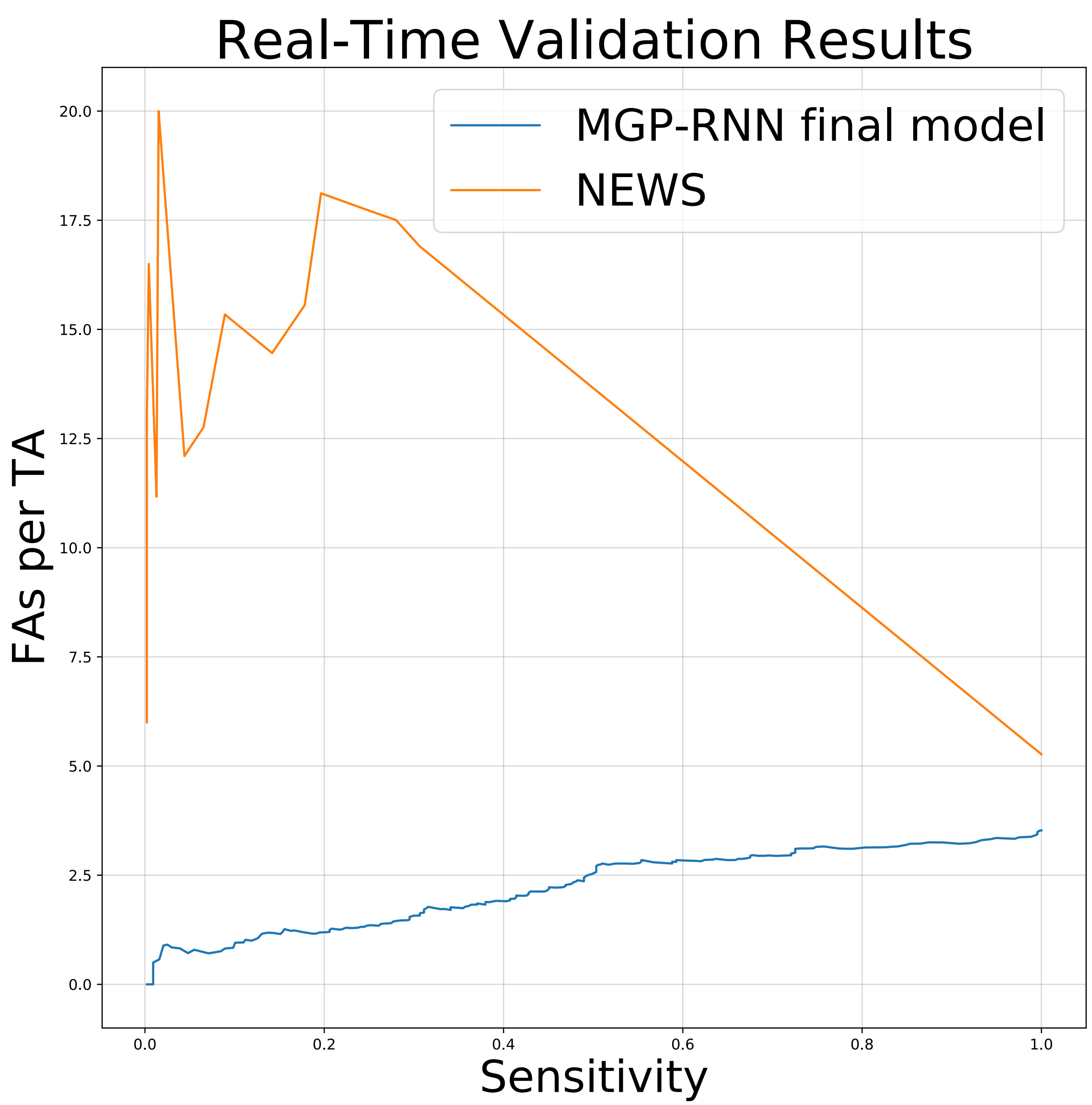
False Postive (Detected too early) → Get confusion matrices and aggregate metrics



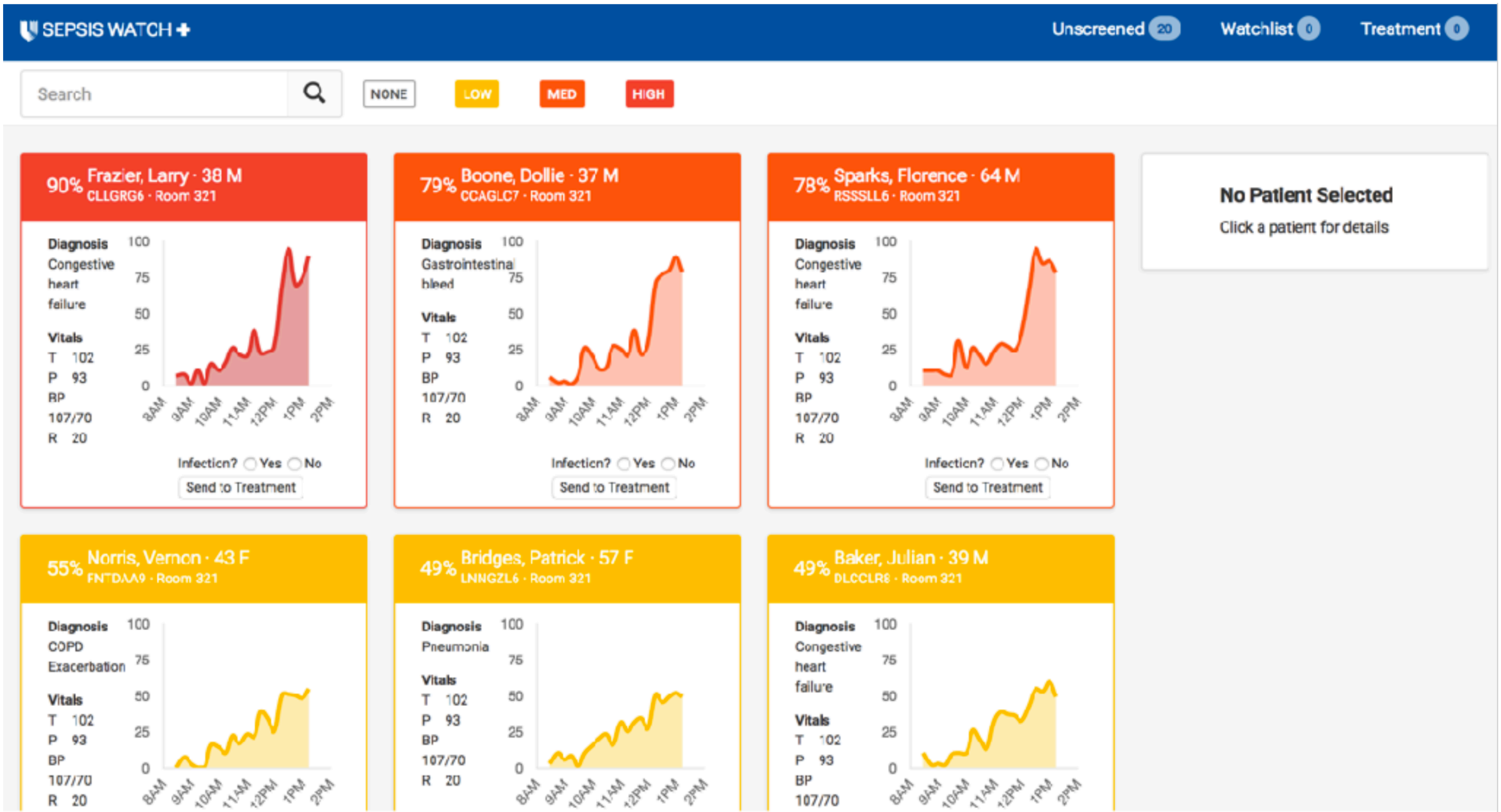
True Positive!

“Real-Time” Results

- **MGP-RNN:** our approach.
- **NEWS:** clinical baseline, previously used at Duke.

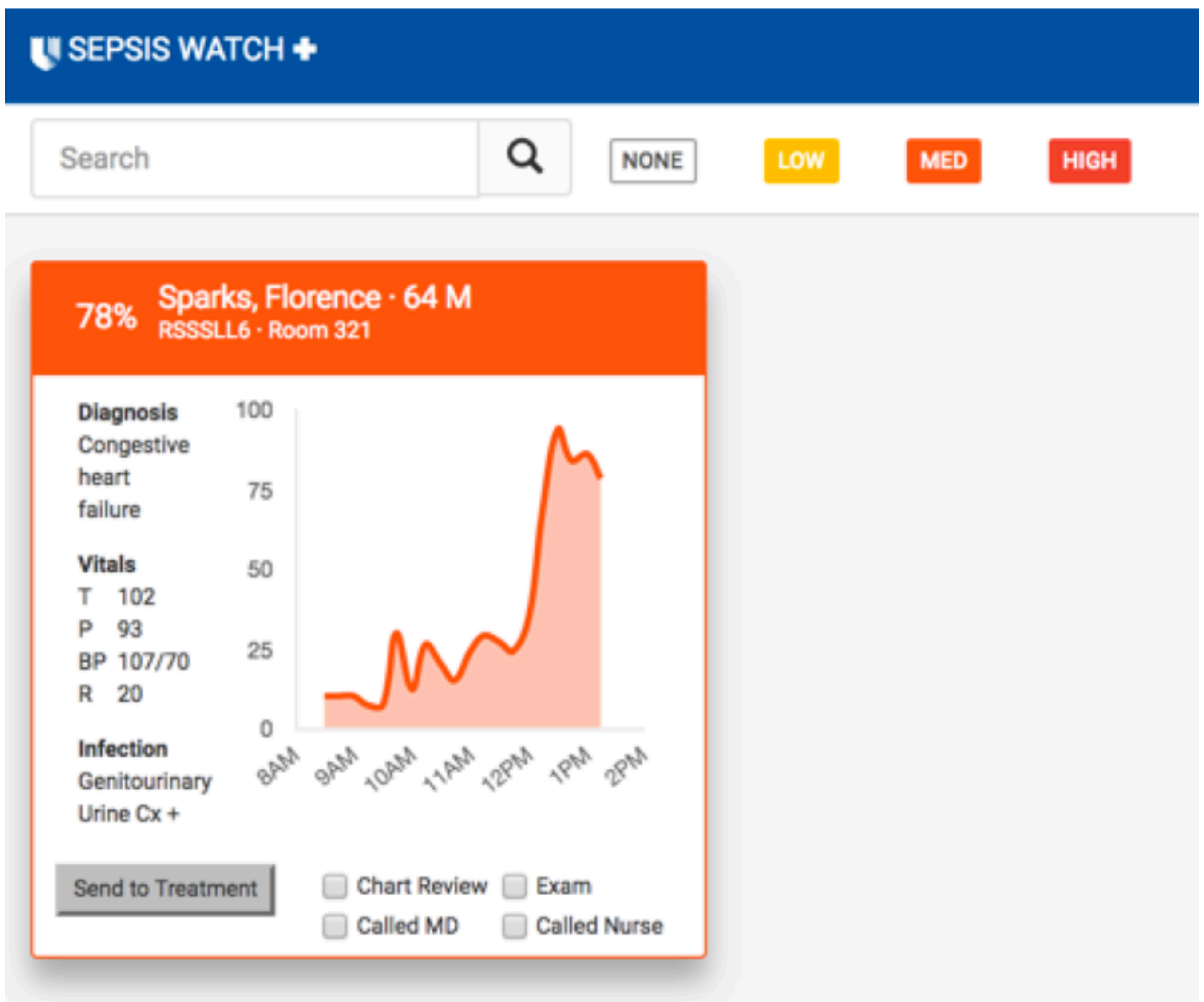


SepsisWatch

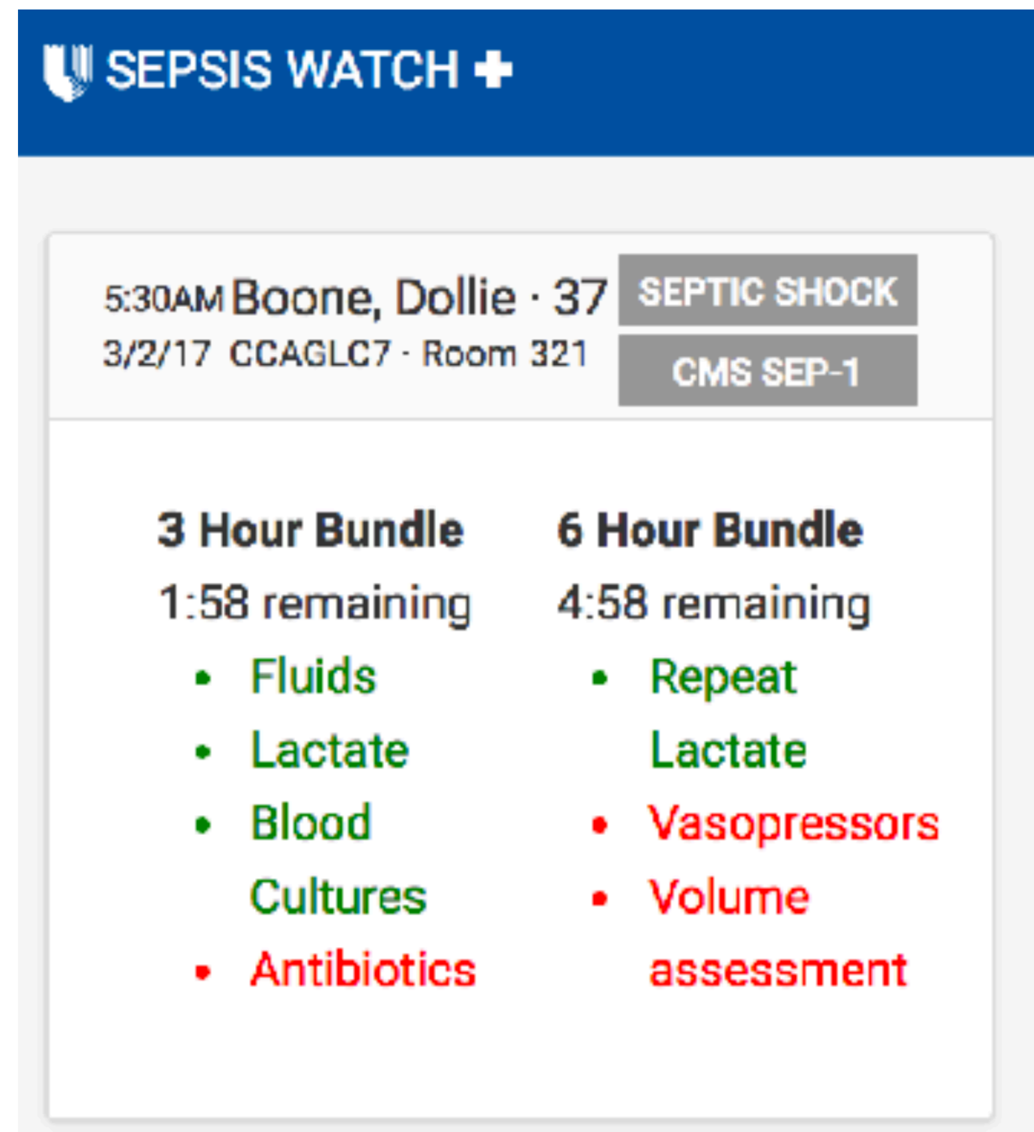


SepsisWatch

Watchlist



Treatment



Conclusion

- Improved an existing classifier that uniquely combines **Gaussian Processes** and **deep learning**.
- Developed a more realistic validation scheme to simulate model performance in real-time.
- To be used in actual practice in a **clinical trial!**
- Many exciting new directions:
  - Model other clinical events (e.g. code blue).
  - RNN Attention mechanism for interpretability.
  - Reinforcement Learning to recommend optimal treatments.

Acknowledgements

DIHI Team:
Mark Sendak, MD, MPP
Nathan Brajer, MD Candidate
Michael Gao
Suresh Balu, MBA
Machine Learning:
Sanjay Hariharan, MS
Katherine Heller, PhD
Sepsis Clinicians:
Cara O'Brien, MD
Armando Bedoya, MD
Meredith Clement, MD
Software Developer:
Faraz Yashar

Duke Statistical Science

NDSEG
NATIONAL DEFENSE SCIENCE & ENGINEERING GRADUATE FELLOWSHIP

Duke Institute for Health Innovation

jdf38@duke.edu

https://github.com/jfutoma/MGP-RNN