

Learning to Detect Sepsis with a Multitask Gaussian Process RNN Classifier

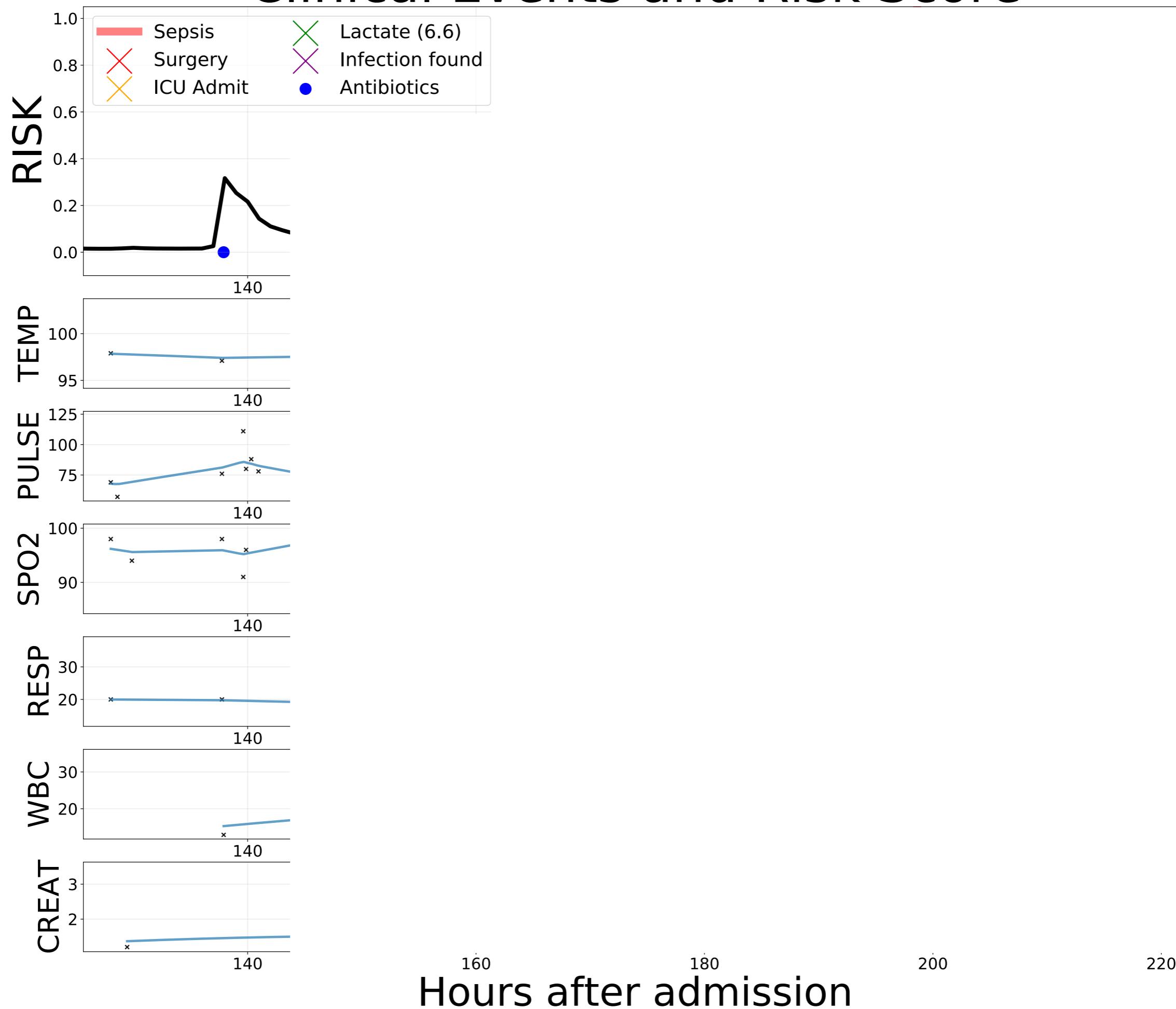
Joe Futoma
Duke University
Dept. of Statistical Science

Outline

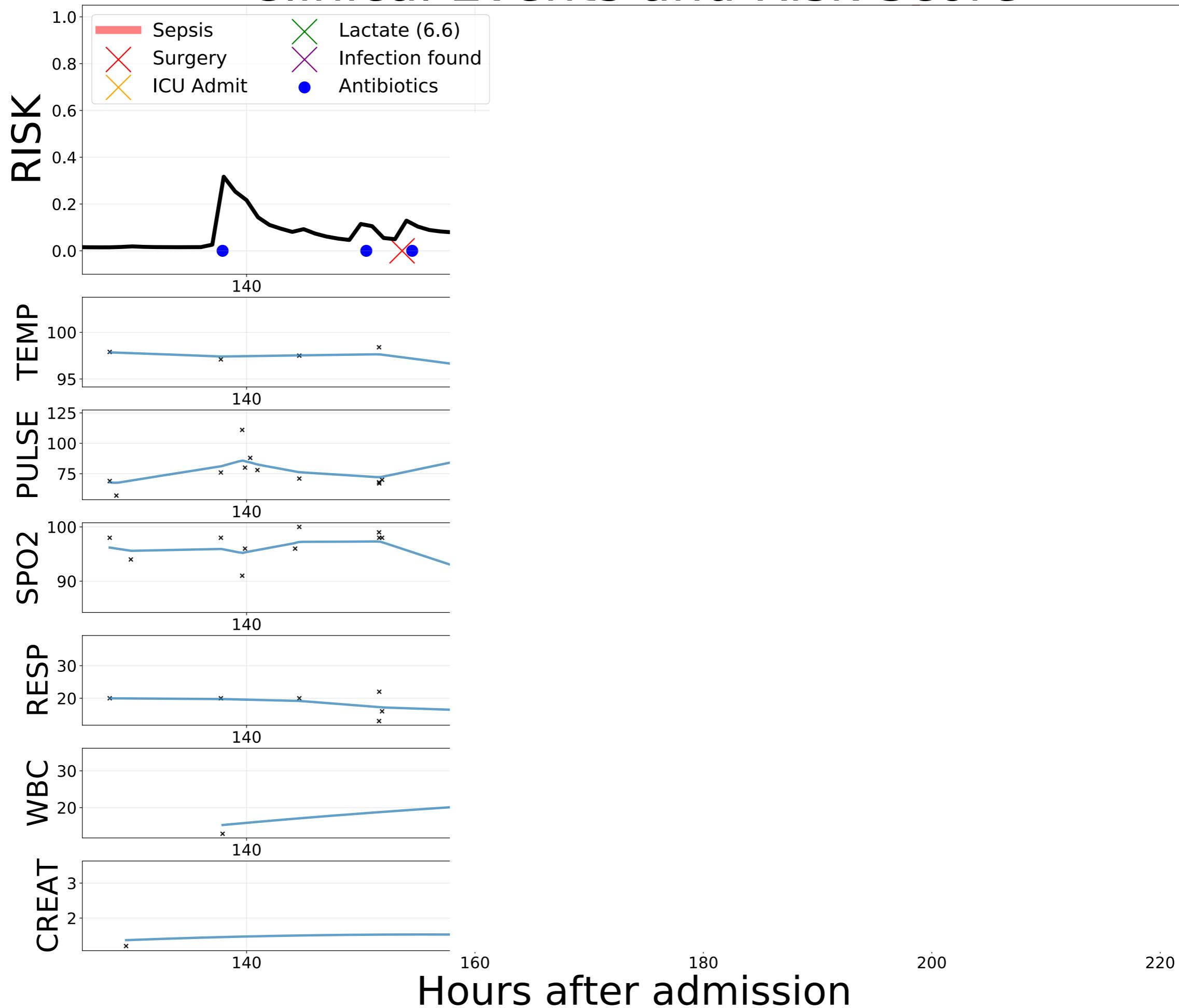
- Patient Story
- Background
- Proposed Model
- Experiments & Results
- In Clinical Practice

Patient Story

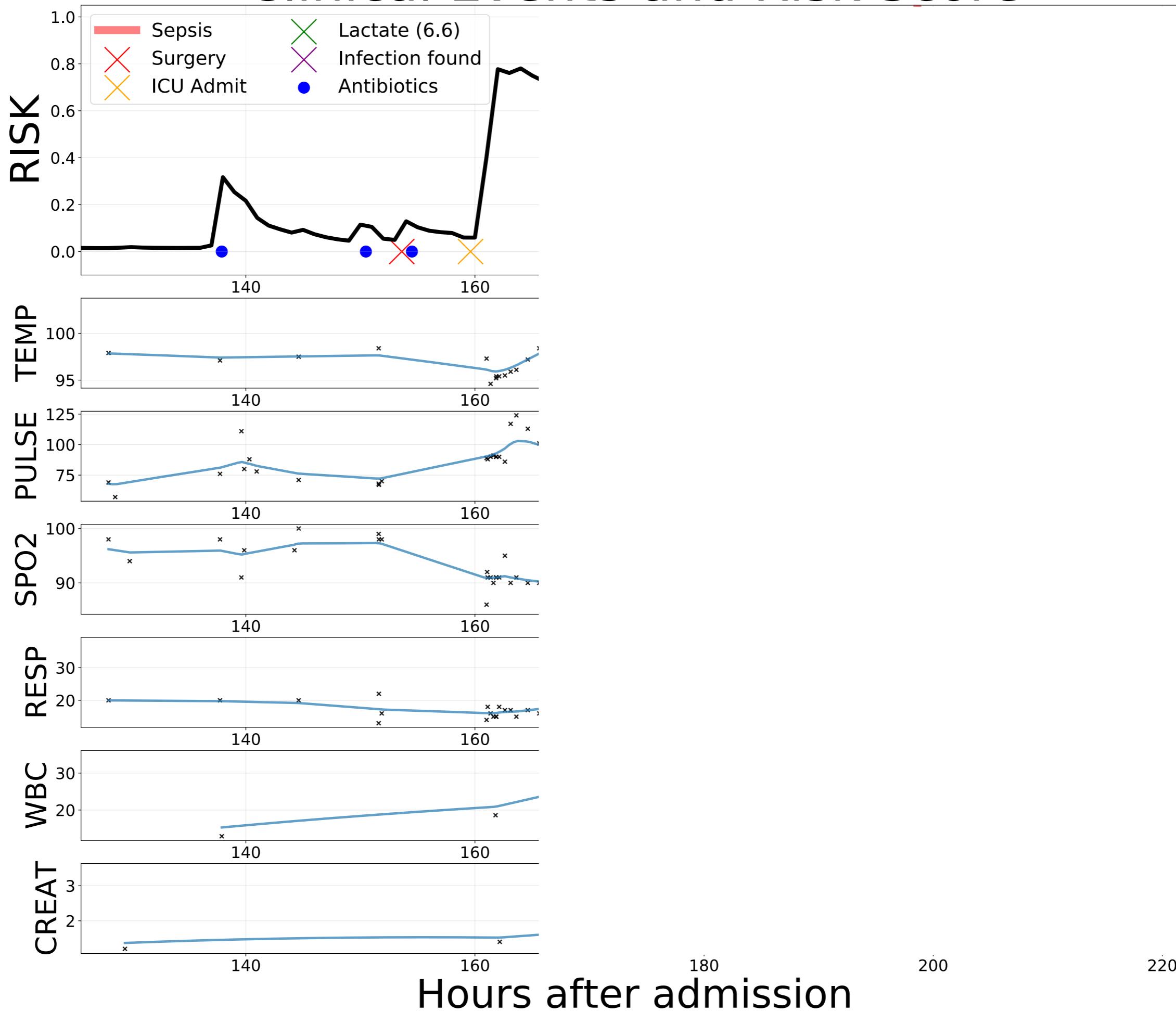
Clinical Events and Risk Score



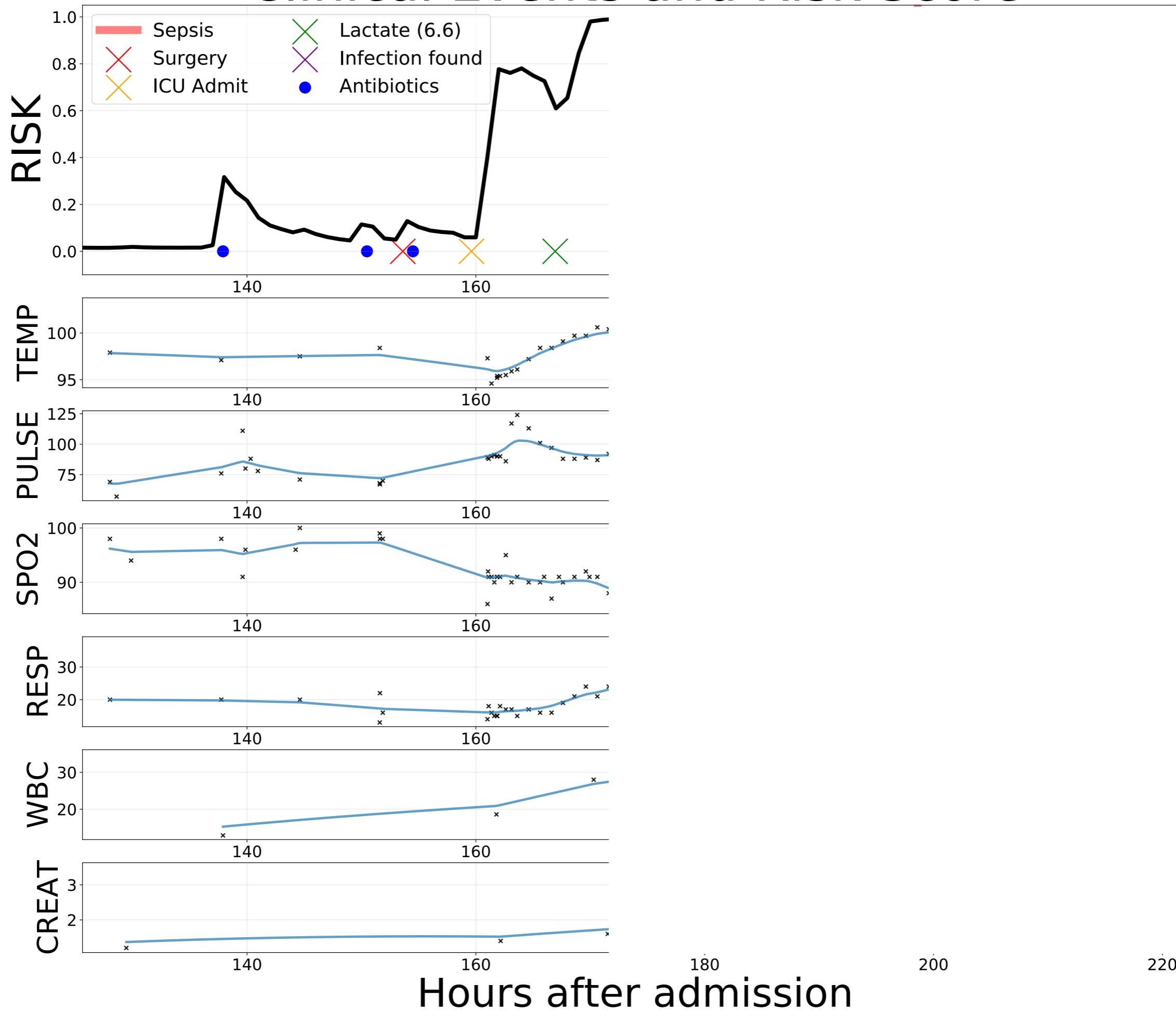
Clinical Events and Risk Score



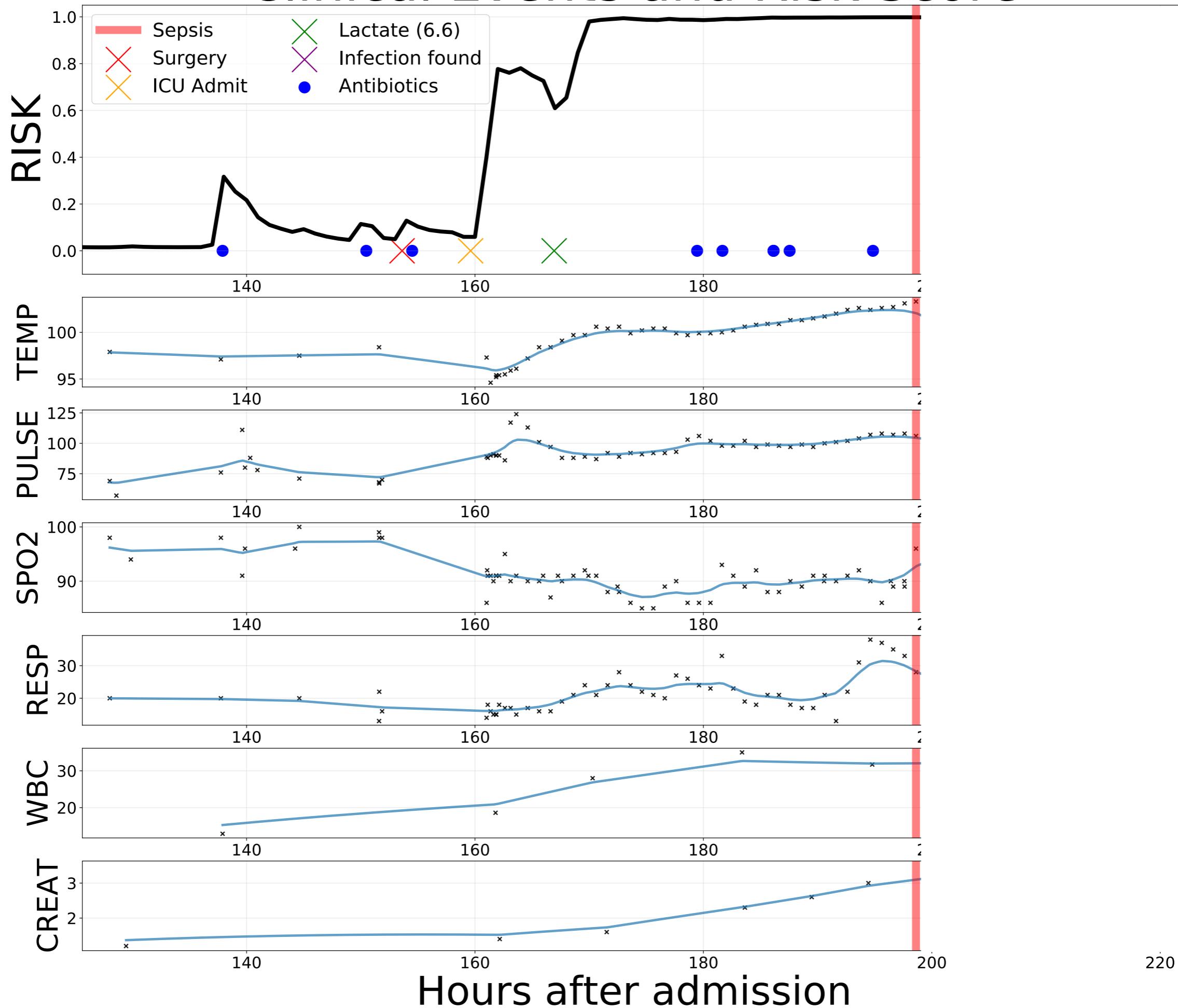
Clinical Events and Risk Score



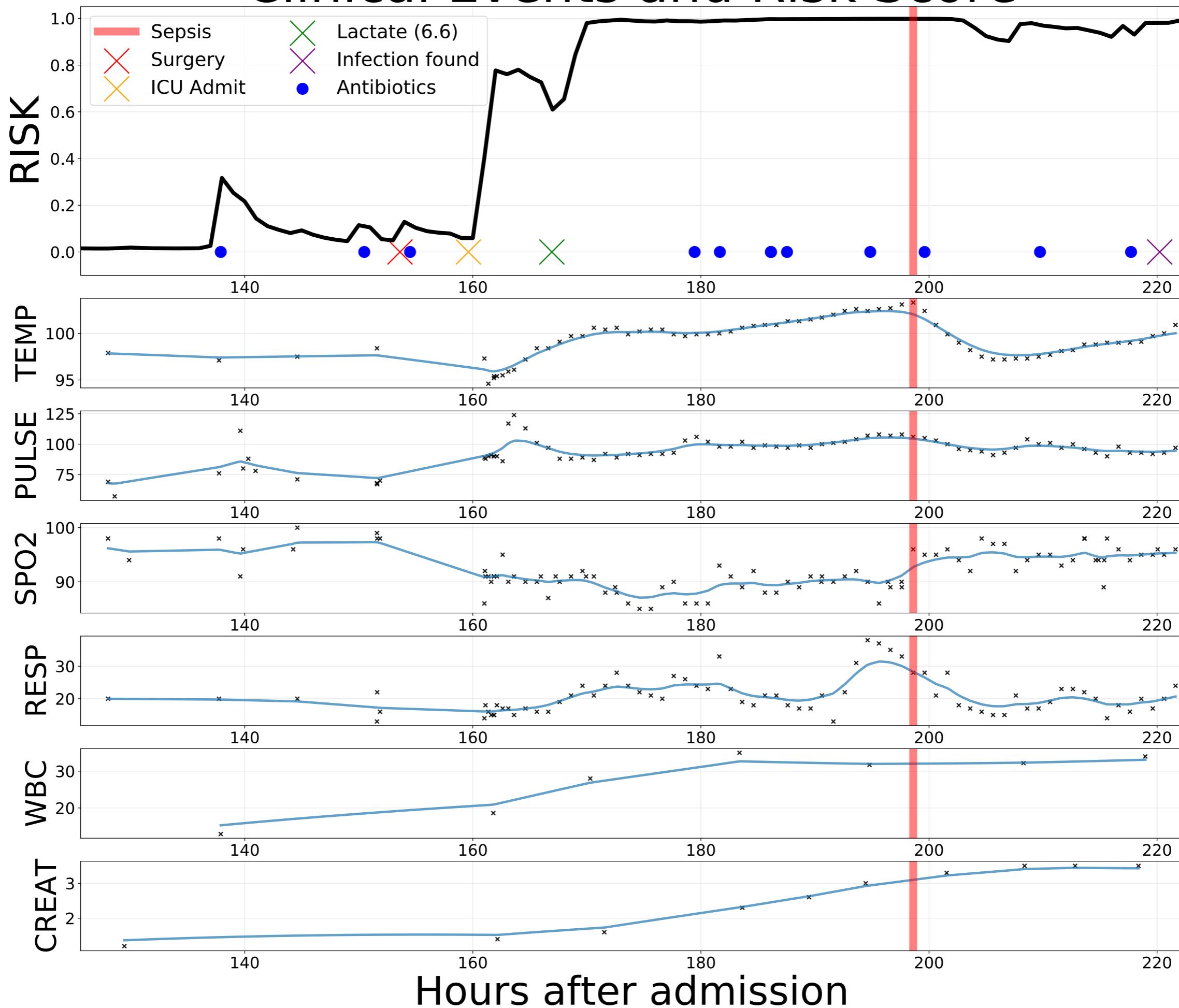
Clinical Events and Risk Score



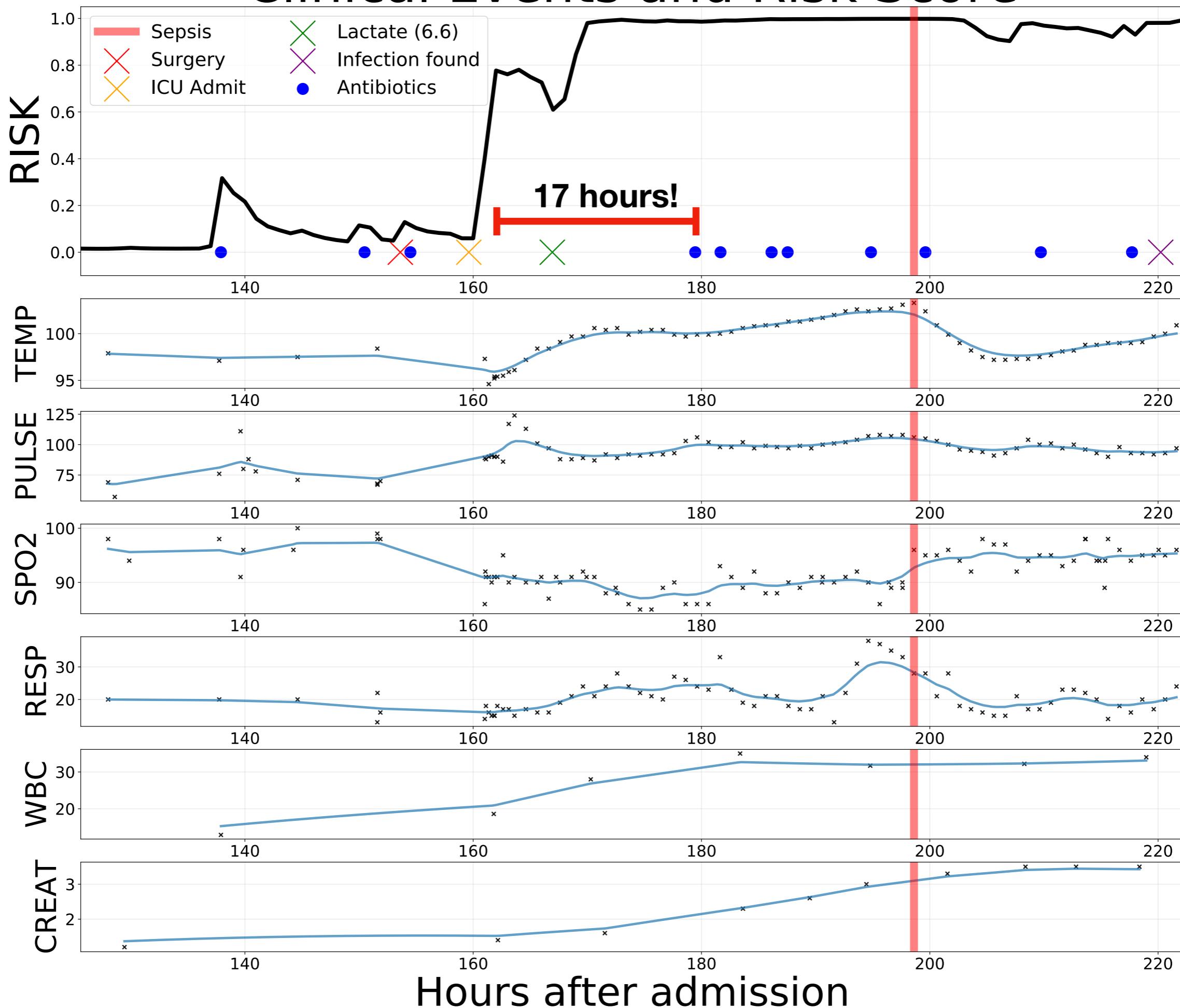
Clinical Events and Risk Score



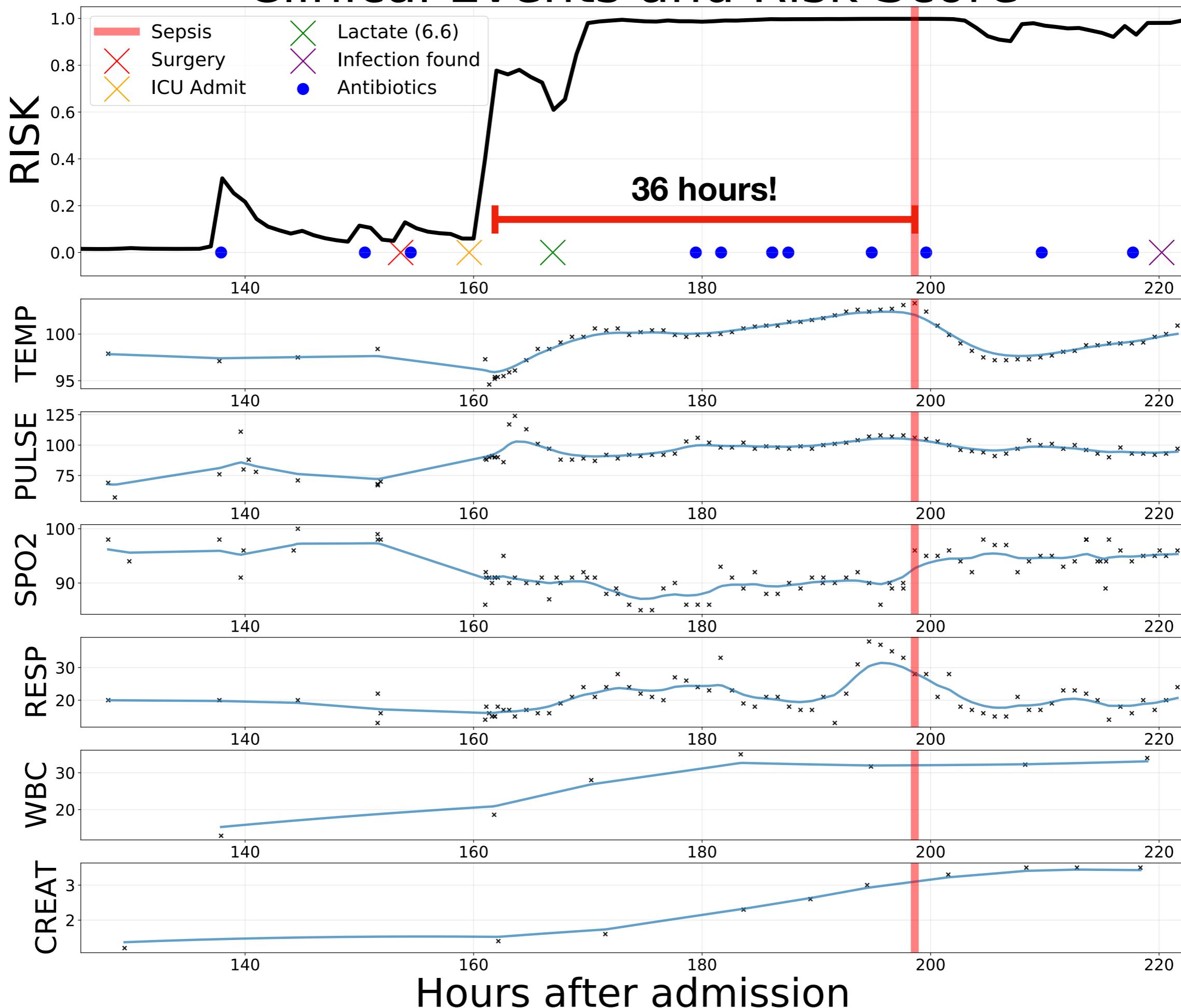
Clinical Events and Risk Score



Clinical Events and Risk Score



Clinical Events and Risk Score



Background

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.
- **Early identification is hard:**
 - No clear time of onset, no reliable biomarker (yet).

Surviving Sepsis Campaign



- **Sepsis Care Bundles:** selected elements of care from evidence-based practice guidelines.

The NEW ENGLAND JOURNAL of MEDICINE

- In first 3 hours:

1. Measure lactate.

2. Get blood cultures.

3. Give antibiotics.

ORIGINAL ARTICLE

Time to Treatment and Mortality during Mandated Emergency Care for Sepsis

Christopher W. Seymour, M.D., Foster Gosten, 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.

- Other actions at 6 hours if no improvement.

- **We know what to do, if we know it's there!**

Proposed Model

Related Works

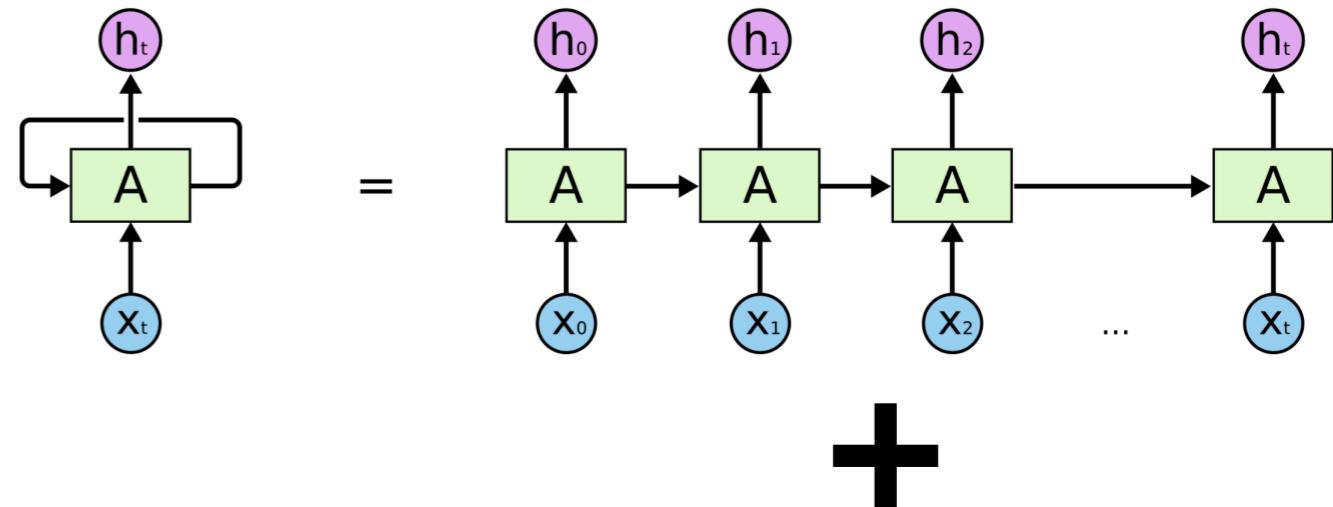
- Clinical Early Warning Scores, e.g. NEWS, SIRS, MEWS, Apache II.
 - NEWS at Duke: 63.4% of triggered alerts cancelled by nurse.
 - Typically broad, not targeted for particular conditions.
 - Low precision, leading to high alarm fatigue.
- (Henry et al, Science Translational Medicine 2015): TREWS score: Cox regression to predict time to septic shock, using 54 potential features [MIMIC data].
- (Ghassemi et al, AAAI 2015): Use MGPs for modeling multivariate physiological time series data from the ICU [MIMIC data].
- (Yoon et al, ICML 2016), (Hoiles & van der Schaar, NIPS 2016): related problem of predicting time to ICU admission, using streams of clinical data [UCLA in-house data].
- (Cheng-Xian & Marlin, NIPS 2016): “GP-adapter” for classifying univariate irregularly spaced time series, of the same fixed length.

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.



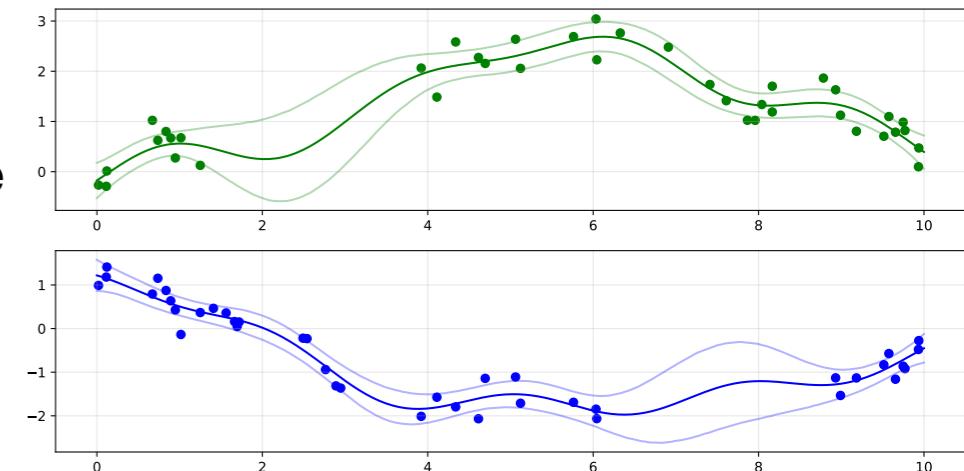
- Multivariate time series classification: update a risk score (probability encounter is / will become septic).

- **Recurrent Neural Networks (RNNs)**: deep learning, learn flexible and powerful functions from sequences of arbitrary length. But:

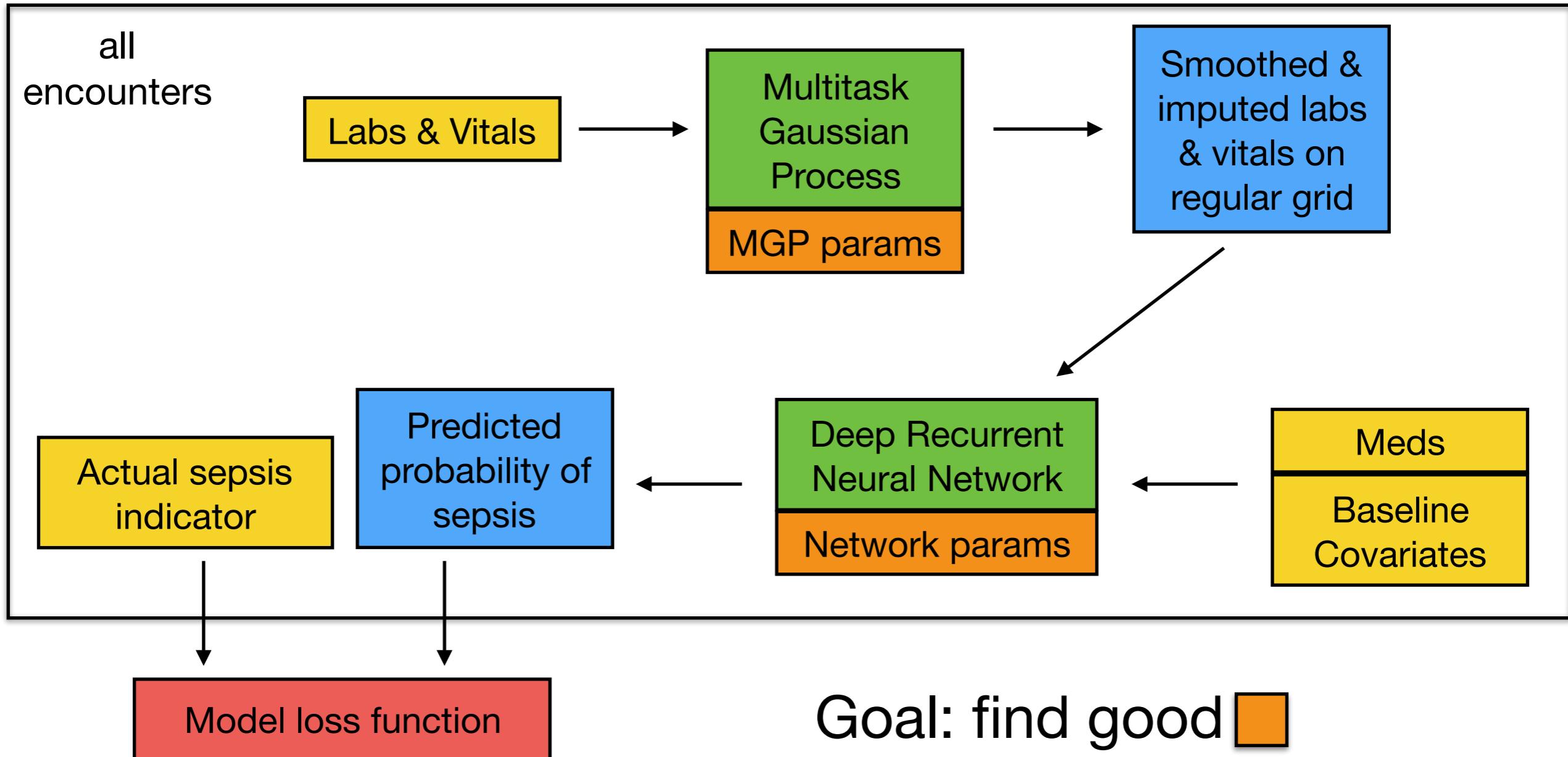
- Requires a **complete** dataset with no missing values.
- Requires **regularly** spaced inputs.

- **Multitask 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**.

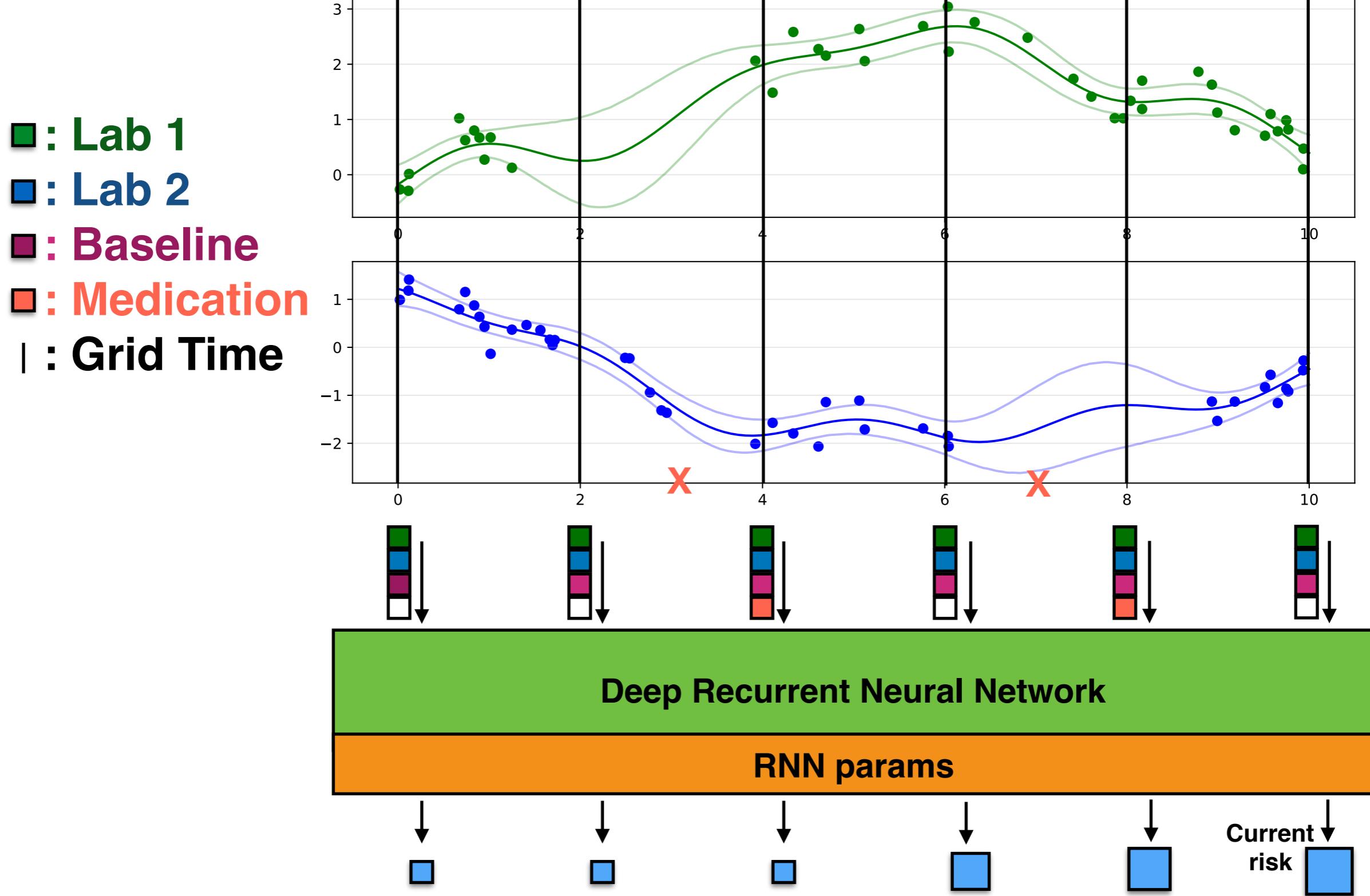


Model Architecture



End-to-end learning!

Model Example



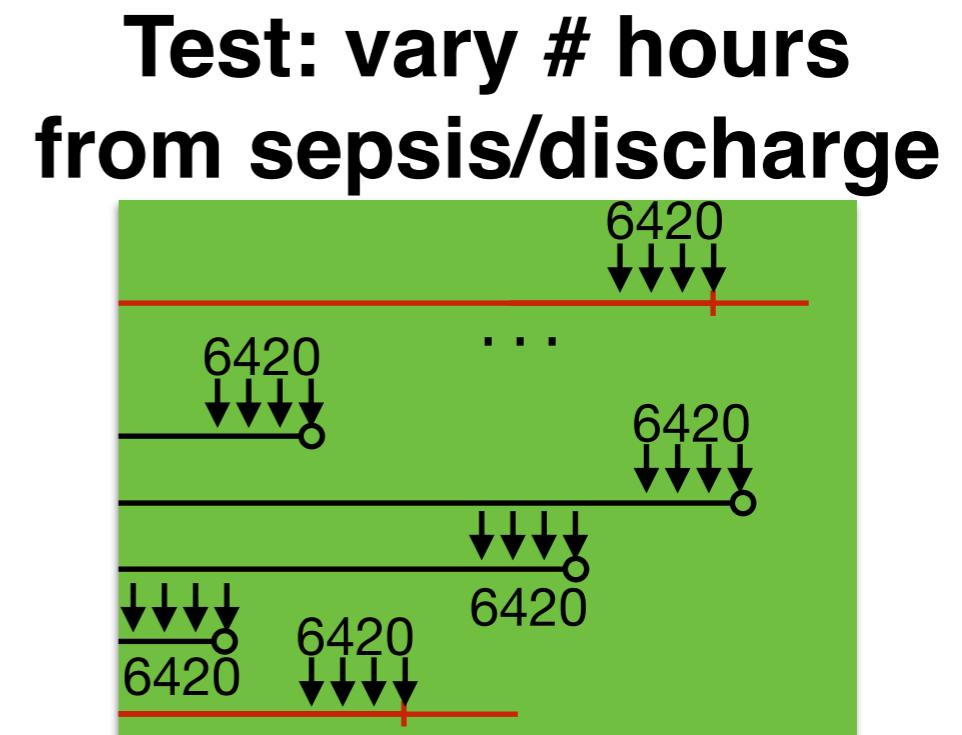
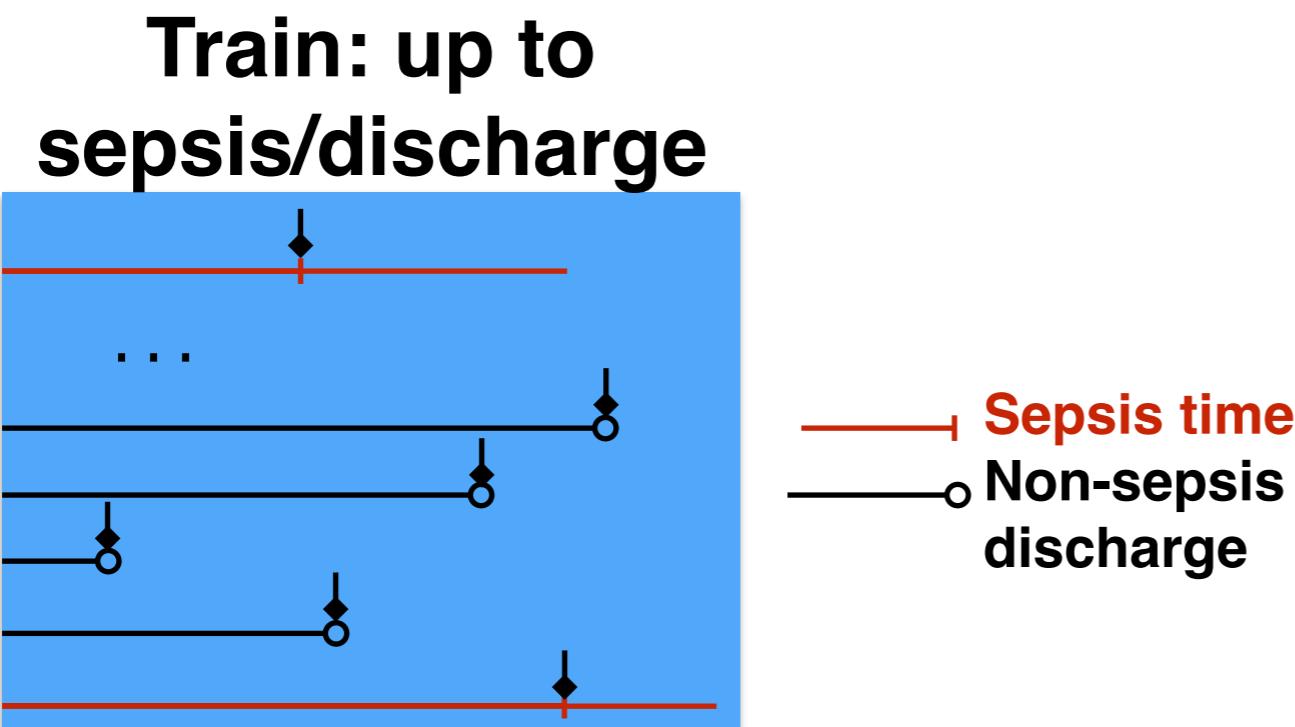
Experiments & Results

Dataset

- **49,312** 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.

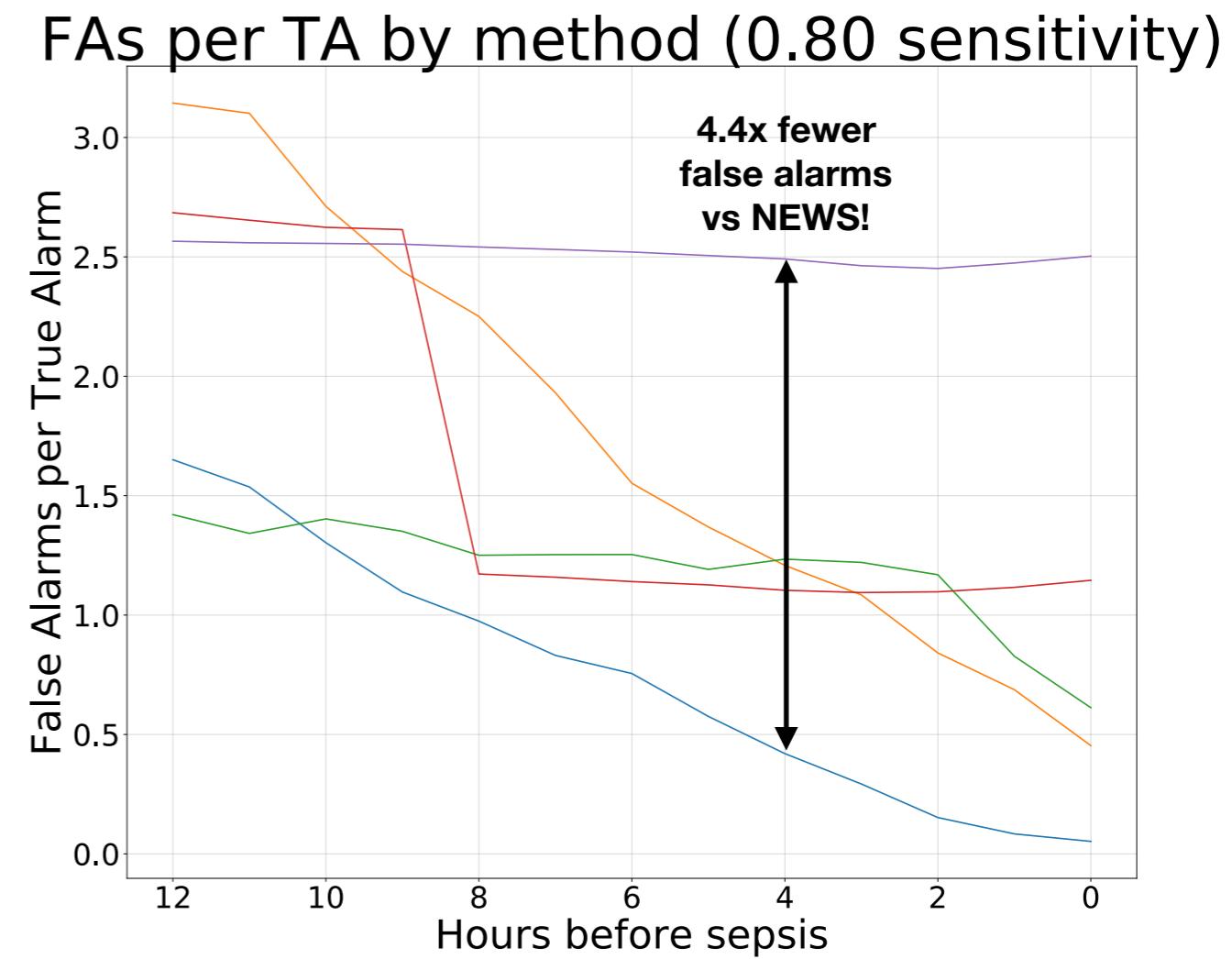
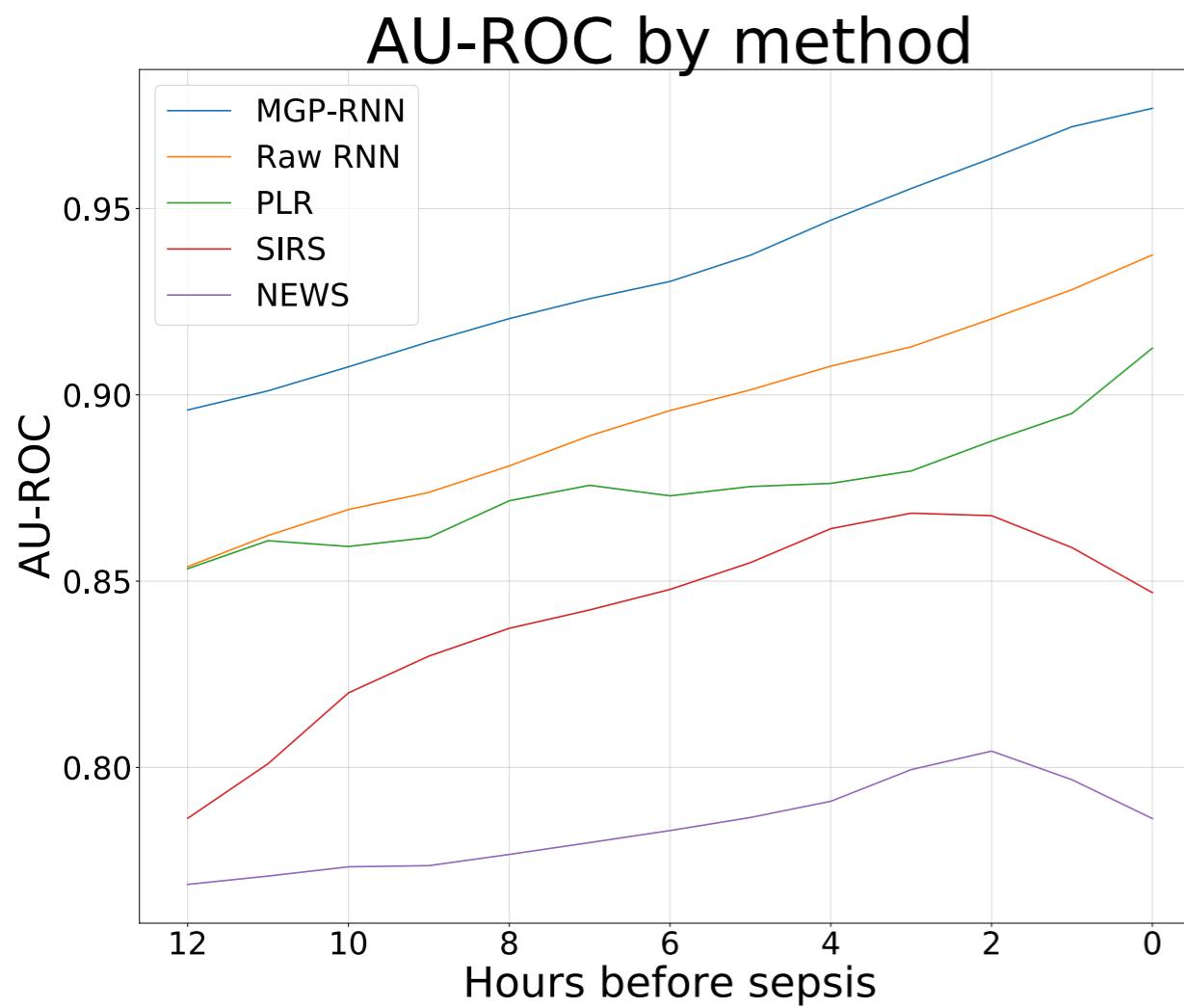
Experimental Setup

- Compare metrics hours in advance of sepsis/discharge:
 - **AU-ROC:** Area under ROC curve / C-statistic.
 - **False Alarms per True Alarm:** For fixed sensitivity, number of false alarms for each true alarm (related to precision).



Results

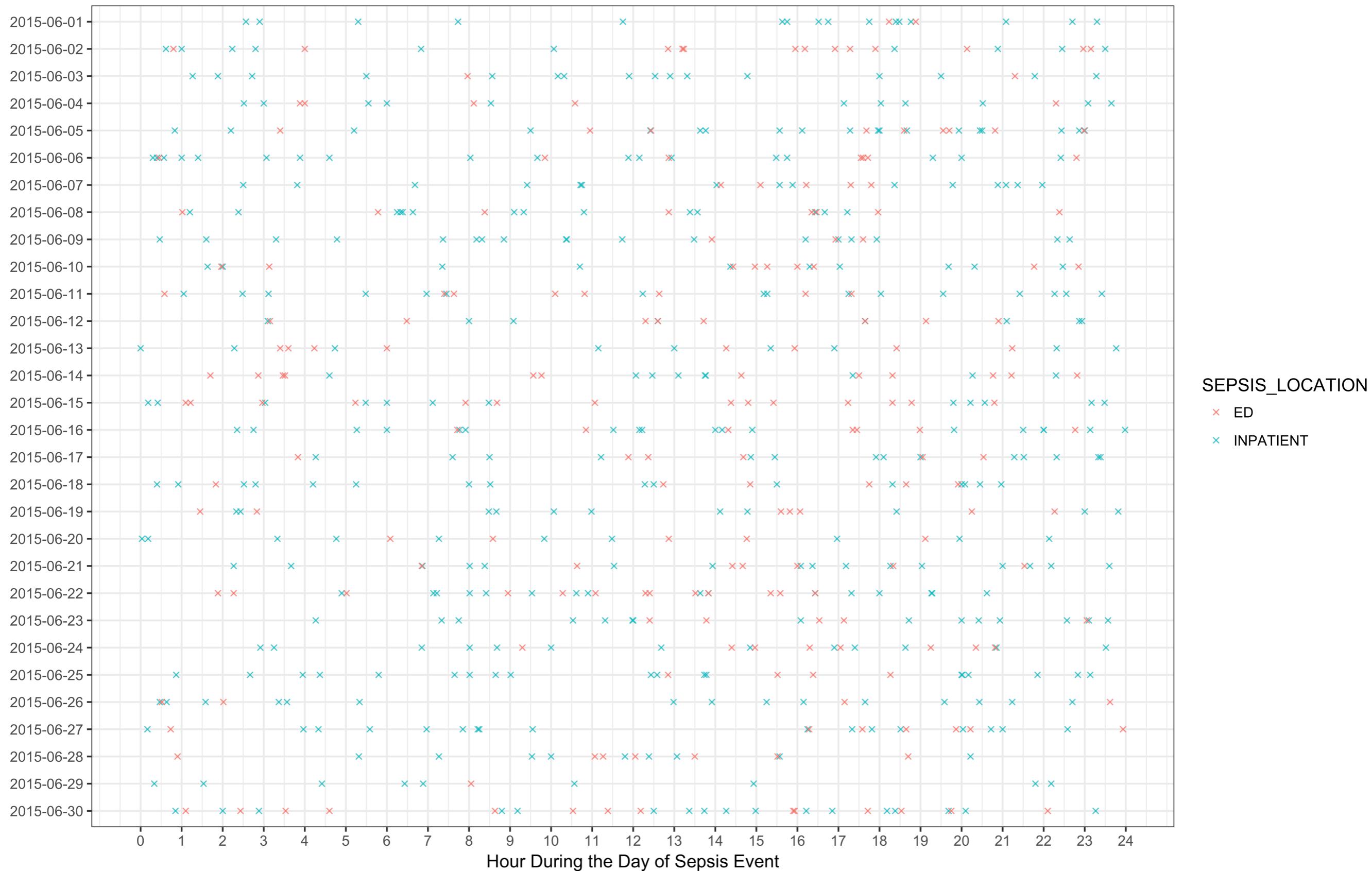
- **MGP-RNN:** our approach
- **Raw RNN:** RNN trained on raw data (missing: carry forward last observed value)
- **PLR:** Penalized logistic regression, same imputation as Raw RNN
- **SIRS, NEWS:** clinical scores



In Clinical Practice

Sepsis Events in June 2015

Split By Setting



SepsisWatch

SEPSIS WATCH +

Unscreened 20

Watchlist 0

Treatment 0

Search



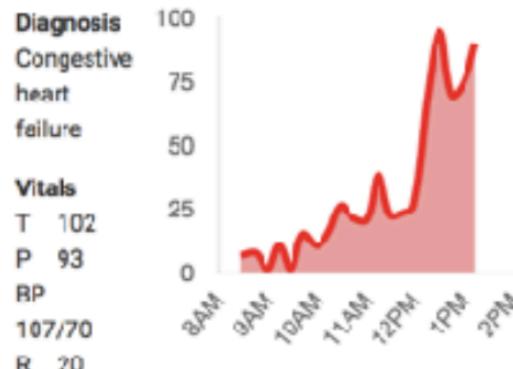
NONE

LOW

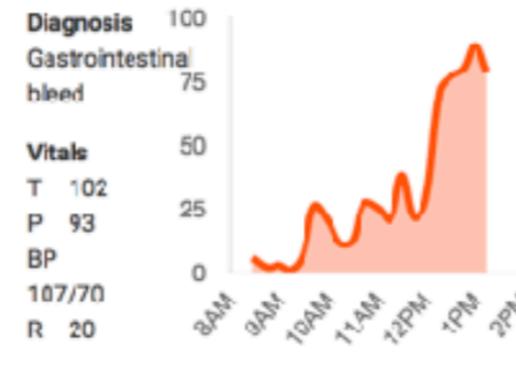
MED

HIGH

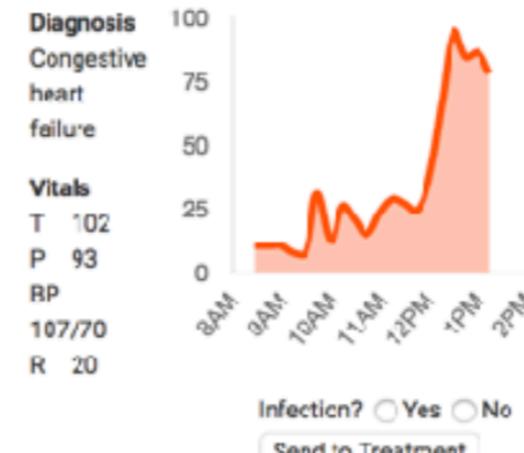
90% Frazier, Larry - 38 M
CLLGRG6 · Room 321



79% Boone, Dolly - 37 M
CCAGLC7 · Room 321



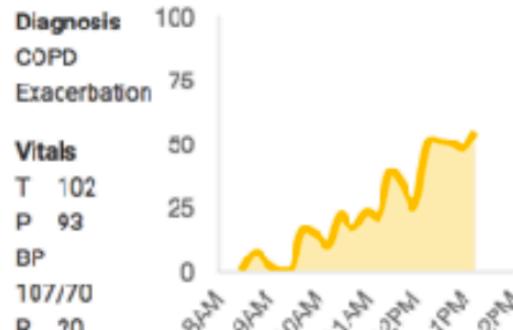
78% Sparks, Florence - 64 M
RSSSL6 · Room 321



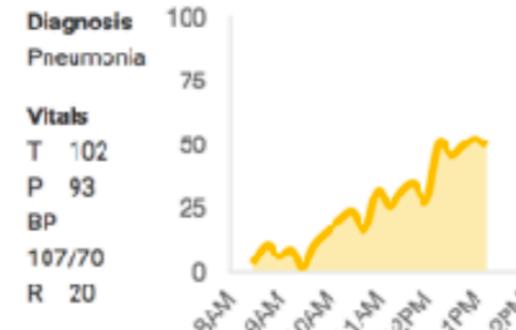
No Patient Selected

Click a patient for details

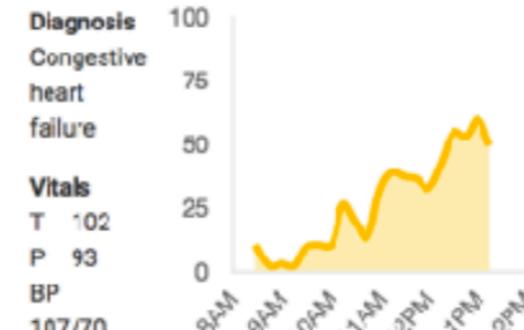
55% Norris, Vernon - 43 F
FNTDAA9 · Room 321



49% Bridges, Patrick - 57 F
LNNGZL6 · Room 321



49% Baker, Julian - 39 M
DLCCLR8 · Room 321



SepsisWatch

Watchlist

SEPSIS WATCH +

Search  NONE LOW MED HIGH

78% Sparks, Florence · 64 M
RSSLL6 · Room 321

Diagnosis: Congestive heart failure (100)
Vitals: T 102, P 93, BP 107/70, R 20
Infection: Genitourinary Urine Cx +



Send to Treatment Chart Review Exam
 Called MD Called Nurse

Treatment

SEPSIS WATCH +

5:30AM Boone, Dollie · 37 SEPTIC SHOCK
3/2/17 CCAGLC7 · Room 321 CMS SEP-1

3 Hour Bundle
1:58 remaining

- Fluids
- Lactate
- Blood
- Cultures
- Antibiotics

6 Hour Bundle
4:58 remaining

- Repeat Lactate
- Vasopressors
- Volume assessment

...but does it work? Let's find out!

- **Sepsis Rapid Response Team (RRT):** Fast-responding team to help coordinate care for patients with suspected sepsis.
 - Cardiac care unit nurses, pharmacists, hospitalists, respiratory therapists, administrators (logistics).
- Patients at high risk, or that meet the sepsis definition, will be reviewed by a care nurse.
- In planning stages of **randomized clinical trial!**
 - Goal: evaluate effect of using the app on clinical outcomes (in-hospital mortality).
 - Secondary outcomes: compliance to completing 3, 6 hour bundles on time.
 - On target to launch this fall!

Conclusion

- Novel model for early detection of sepsis, leveraging **deep learning** and **Gaussian processes**.
- Significantly improved performance over NEWS used at Duke.
- To be used in actual practice in a **clinical trial!**
- Many exciting new directions:
 - Model other clinical events (e.g. code blue). Methods are **generalizable!**
 - RNN Attention mechanism for interpretability.
 - Learn treatment-response curves from medications data.
 - 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:

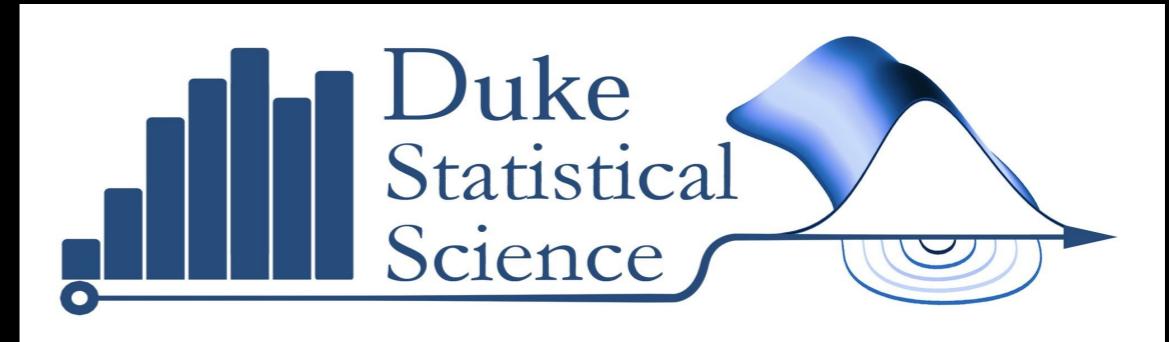
Cara O'Brien, MD

Armando Bedoya, MD

Meredith Clement, MD

Software Developer:

Faraz Yashar



jdf38@duke.edu

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

<https://arxiv.org/abs/1706.04152>

Defining Sepsis



accp/sccm consensus conference

Intensive Care Med (2003) 29:530–538
DOI 10.1007/s00134-003-1662-x

EXPERT PANEL

Clinical Review & Education

Definitions for Sepsis and Organ Failure and Guidelines for the Use of Innovative Therapies in Sepsis

THE ACCP/SCCM CONSENSUS CONFERENCE COMMITTEE:

Roger C. Bone, M.D., F.C.C.P., Chairman

Robert A. Balk, M.D., F.C.C.P.

Frank B. Cerra, M.D.

R. Phillip Dellinger, M.D., F.C.C.P.

Alan M. Fein, M.D., F.C.C.P.

William A. Knaus, M.D.

Roland M. H. Schein, M.D.

William J. Sibbald, M.D., F.C.C.P.

Mitchell M. Levy
Mitchell P. Fink
John C. Marshall
Edward Abraham
Derek Angus
Deborah Cook
Jonathan Cohen
Steven M. Opal
Jean-Louis Vincent
Graham Ramsay
for the International Sepsis Definitions Conference

2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference

Special Communication | CARING FOR THE CRITICALLY ILL PATIENT

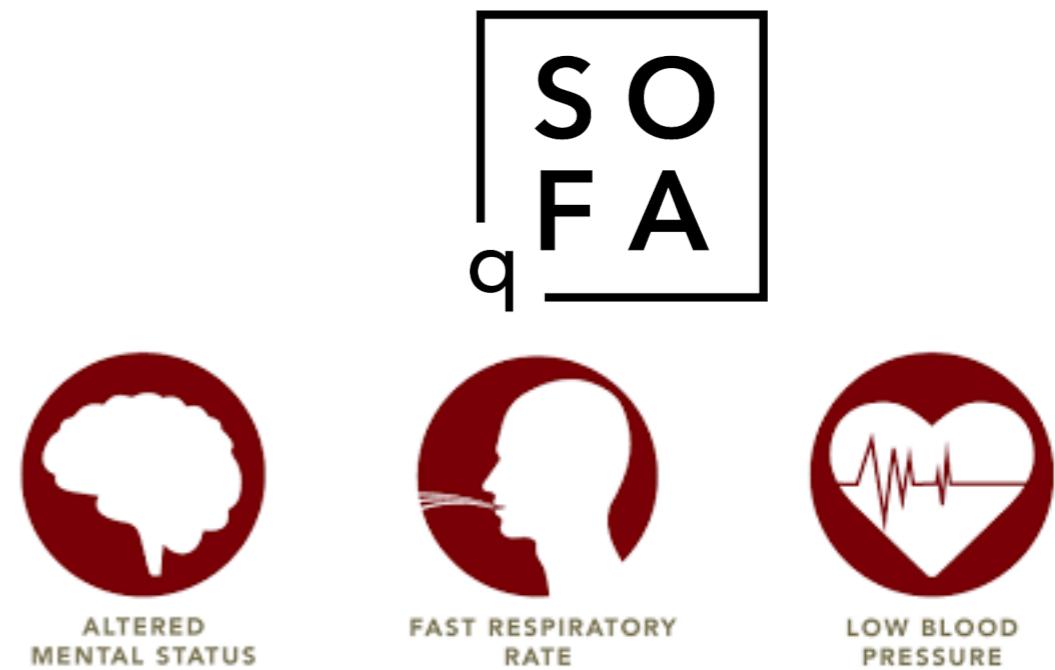
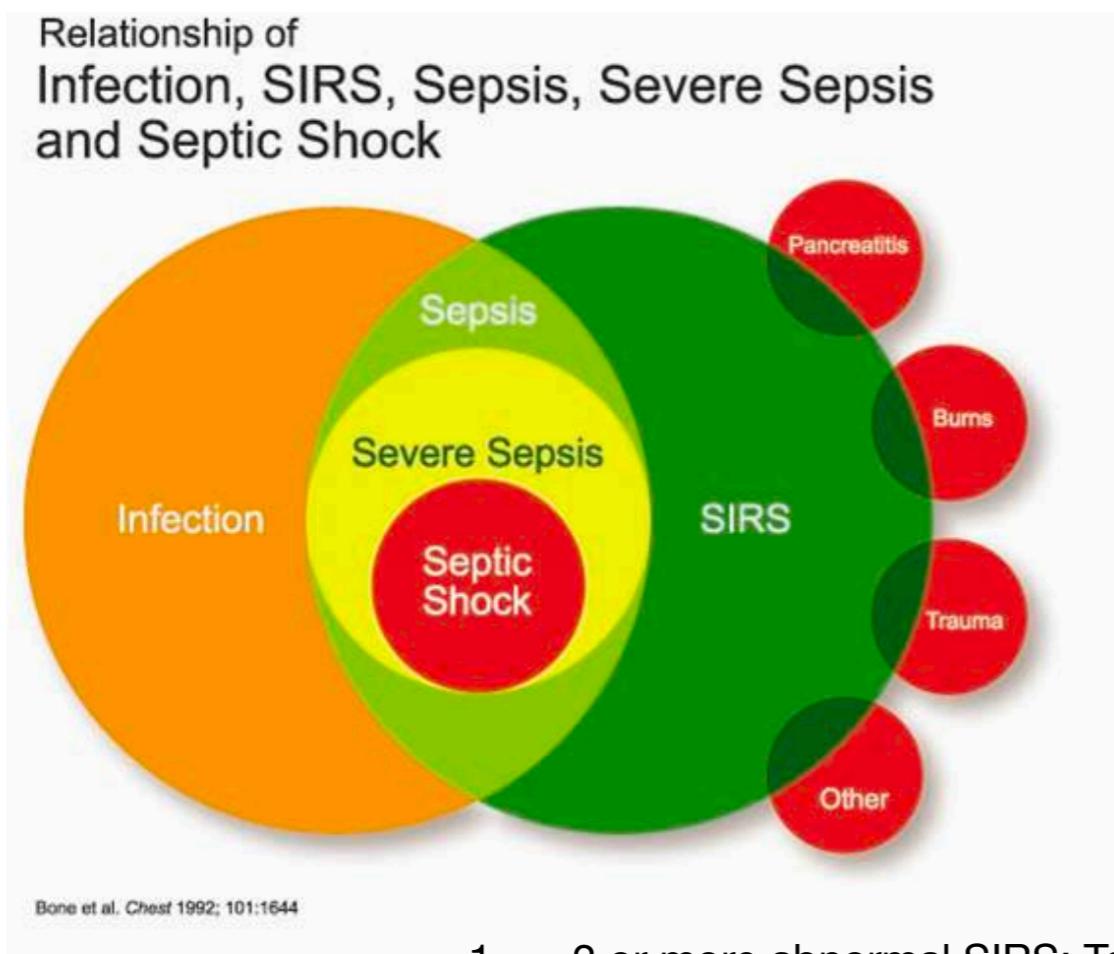
The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)

Mervyn Singer, MD, FRCP; Clifford S. Deutschman, MD, MS; Christopher Warren Seymour, MD, MSc; Manu Shankar-Hari, MSc, MD, FRCR; Djillali Annane, MD, PhD; Michael Bauer, MD; Rinaldo Bellomo, MD; Gordon R. Bernard, MD; Jean-Daniel Chiche, MD, PhD; Craig M. Coopersmith, MD; Richard S. Hotchkiss, MD; Mitchell M. Levy, MD; John C. Marshall, MD; Greg S. Martin, MD, MSc; Steven M. Opal, MD; Gordon D. Rubenfeld, MD, MS; Tom van der Poll, MD, PhD; Jean-Louis Vincent, MD, PhD; Derek C. Angus, MD, MPH

1992

2001

2016



1. 2 or more abnormal SIRS: Temperature, Heart Rate, Respiration Rate, WBC Count.
2. Blood culture (suspected infection).
3. End organ damage lab.

Our definition
("Severe Sepsis")

Early Warning Scores

National Early Warning Score (NEWS)*

PHYSIOLOGICAL PARAMETERS	3	2	1	0	1	2	3
Respiration Rate	≤8		9 - 11	12 - 20		21 - 24	≥25
Oxygen Saturations	≤91	92 - 93	94 - 95	≥96			
Any Supplemental Oxygen		Yes		No			
Temperature	≤35.0		35.1 - 36.0	36.1 - 38.0	38.1 - 39.0	≥39.1	
Systolic BP	≤90	91 - 100	101 - 110	111 - 219			≥220
Heart Rate	≤40		41 - 50	51 - 90	91 - 110	111 - 130	≥131
Level of Consciousness				A			V, P, or U

Can we do better?

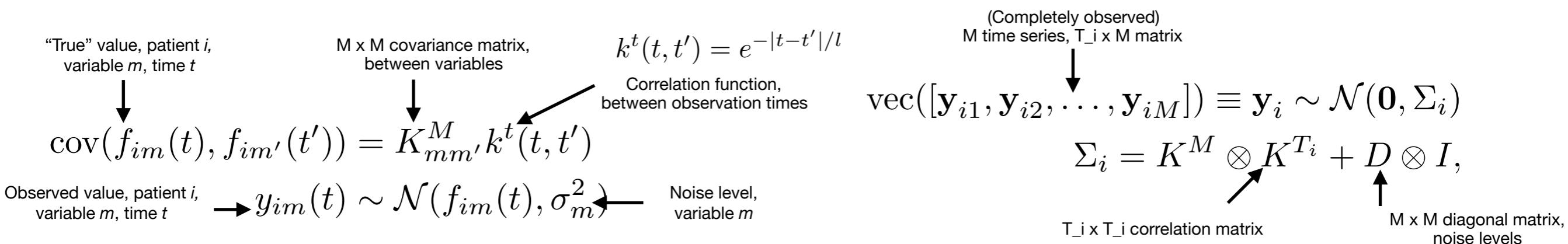
Technical Details

- **Gaussian process:** prior distribution over functions:

$$f_i(t) \sim \mathcal{GP}(\mu(t), K(t, t'))$$

Observation times $\longrightarrow \mathbf{t}_i = (t_{i1}, \dots, t_{iT_i})^T$ covariance matrix
 $f_i(\mathbf{t}_i) \sim \mathcal{N}(\mu(\mathbf{t}_i), K(\mathbf{t}_i, \mathbf{t}_i))$

- **Multitask GP:** extension to multivariate time series.



- Define some regularly spaced (e.g. every hour) reference times, shared across all encounters.

$$\begin{aligned} & \xrightarrow{x_i \text{ shared reference times}} \mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iX_i}) \\ & \xrightarrow{X_i \times M \text{ matrix, latent values at } \mathbf{x}_i} \text{vec}(\mathbf{Z}_i) \equiv \mathbf{z}_i \xleftarrow{\text{Given } \mathbf{y}_i, \text{ conditional normal posterior:}} \\ & \quad \Sigma_{z_i} = (K^M \otimes K^{X_i T_i}) \Sigma_i^{-1} (K^M \otimes K^{T_i X_i}) \\ & \quad \mu_{z_i} = (K^M \otimes K^{X_i T_i}) \Sigma_i^{-1} \mathbf{y}_i \\ & \quad \Sigma_{z_i} = (K^M \otimes K^{X_i}) - (K^M \otimes K^{X_i T_i}) \Sigma_i^{-1} (K^M \otimes K^{T_i X_i}) \\ & \quad \boxed{\text{MGP parameters to learn, shared across all encounters}} \\ & \quad \theta = (K^M, \{\sigma_m^2\}_{m=1}^M, l) \end{aligned}$$

- MGP posterior for \mathbf{Z}_i : the M labs at X_i times; maintaining uncertainty.

Technical Details

- **Gaussian process:** prior distribution over functions:

$$f_i(t) \sim \mathcal{GP}(\mu(t), K(t, t'))$$

$$\mathbf{t}_i = (t_{i1}, \dots, t_{iT_i})$$

$$f_i(\mathbf{t}_i) \sim \mathcal{N}(\mu(\mathbf{t}_i), K(\mathbf{t}_i, \mathbf{t}_i))$$

- **Multitask GP:** extension to multivariate time series.

$$\text{cov}(f_{im}(t), f_{im'}(t')) = K_{mm'}^M k^t(t, t')$$

$$\text{vec}([\mathbf{y}_{i1}, \mathbf{y}_{i2}, \dots, \mathbf{y}_{iM}]) \equiv \mathbf{y}_i \sim \mathcal{N}(\mathbf{0}, \Sigma_i)$$

$$y_{im}(t) \sim \mathcal{N}(f_{im}(t), \sigma_m^2)$$

$$\Sigma_i = K^M \otimes K^{T_i} + D \otimes I,$$

- Define some regularly spaced (e.g. every hour) reference times, shared across all encounters.

$$\theta = (K^M, \{\sigma_m^2\}_{m=1}^M, l)$$

$$\mathbf{z}_i = \text{vec}(\mathbf{Z}_i) \equiv \mathbf{z}_i$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iX_i})$$

$$\mu_{z_i} = (K^M \otimes K^{X_i T_i}) \Sigma_i^{-1} \mathbf{y}_i$$

$$\Sigma_{z_i} = (K^M \otimes K^{X_i}) - (K^M \otimes K^{X_i T_i}) \Sigma_i^{-1} (K^M \otimes K^{T_i X_i})$$

- MGP posterior for \mathbf{z}_i : the M labs at X_i times; maintaining uncertainty.

- RNN input: latent values \mathbf{z}_i , baseline covariates, medication indicators.

- To learn RNN parameters: optimize loss comparing model predictions to true label.

- **Problem:** \mathbf{z}_{ij} are not observed!

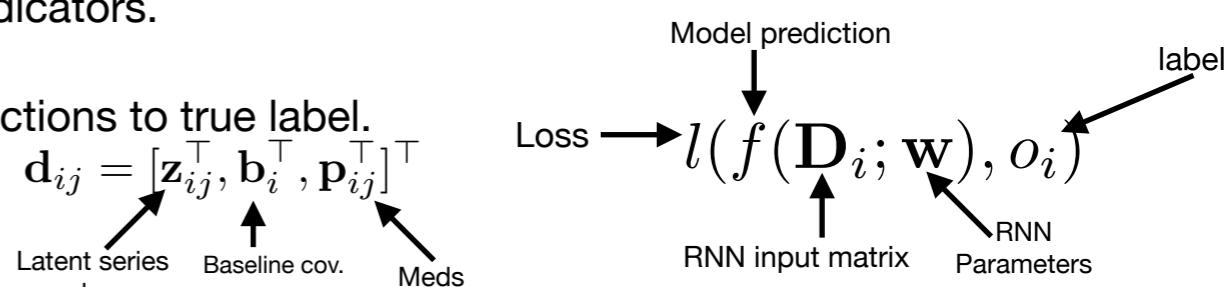
- Instead, optimize expected loss with respect to MGP posterior on \mathbf{z}_i . Overall learning problem:

- Jointly learn MGP, RNN parameters (“end-to-end learning”).

- Optimize with stochastic gradient descent (ADAM).

- Reparameterization trick to get gradients of expectation (MC approx.)

- Conjugate gradient, Lanczos method to speed computation.



$$\mathbf{w}^*, \theta^* = \underset{\mathbf{w}, \theta}{\operatorname{argmin}} \sum_{i=1}^N \mathbb{E}_{z_i \sim N(\mu_{z_i}, \Sigma_{z_i}; \theta)} [l(f(\mathbf{D}_i; \mathbf{w}), o_i)]$$

Risk score for new patient i'

$$\mathbb{E}_{z_{i'} \sim N(\mu_{z_{i'}}, \Sigma_{z_{i'}}; \theta^*)} [f(\mathbf{D}_{i'}; \mathbf{w}^*)]$$