Machine Learning for Clinicians:

Advances for Multi-Modal Health Data

Michael C. Hughes

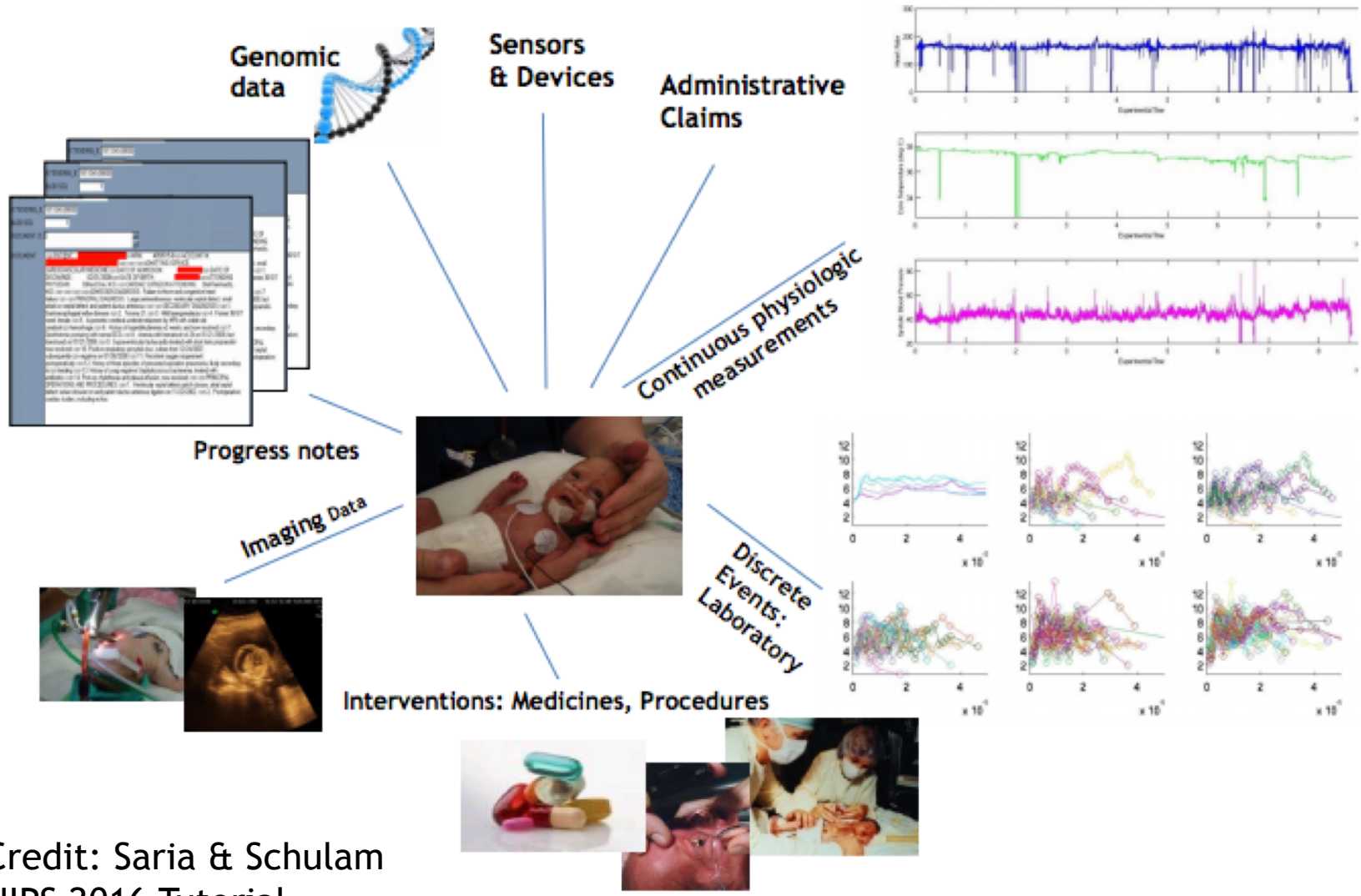
Postdoc (2016-2018)
HARVARD School of Engineering and Applied Sciences

Asst. Prof. Comp. Sci. (2018-)
Tufts University

August 16, 2018
A Tutorial at MLHC 2018

Slides / Resources / Bibliography:
https://michaelchughes.com/mlhc2018_tutorial.html

The Promise of Healthcare Data



Credit: Saria & Schulam
NIPS 2016 Tutorial

Possible ML Use Cases in Health

- New Capabilities
  - Identify heart attack or falls from wearables?
  - Identify new drugs?
- Improve Patient Treatment Decisions
  - What drug to give this patient?
  - Will a ventilator be needed in 2 hours?
- Improve Information Management
  - Summarize this patient's chart?
  - Detect anomaly in dosage?
- Assist with Operations
  - How many beds will be free in the ICU tomorrow?

Exciting Application: Sepsis Risk

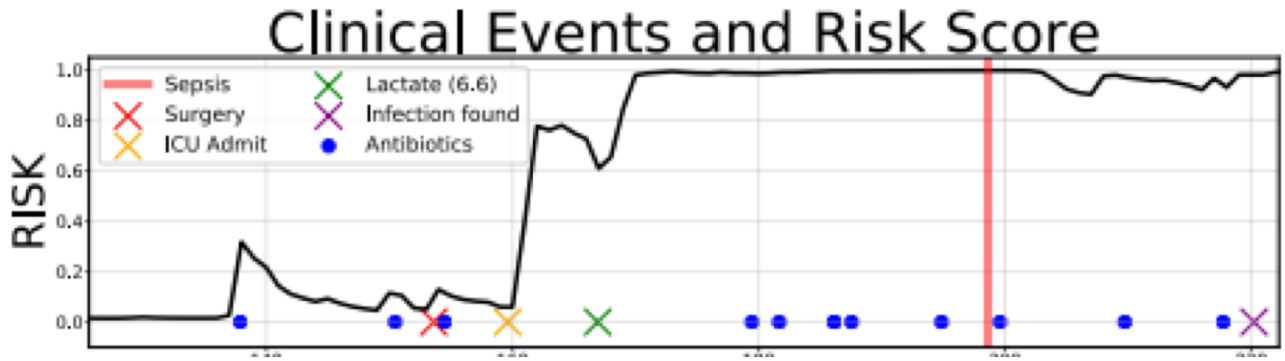
Proceedings of Machine Learning for Healthcare 2017 JMLR W&C Track Volume 68

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

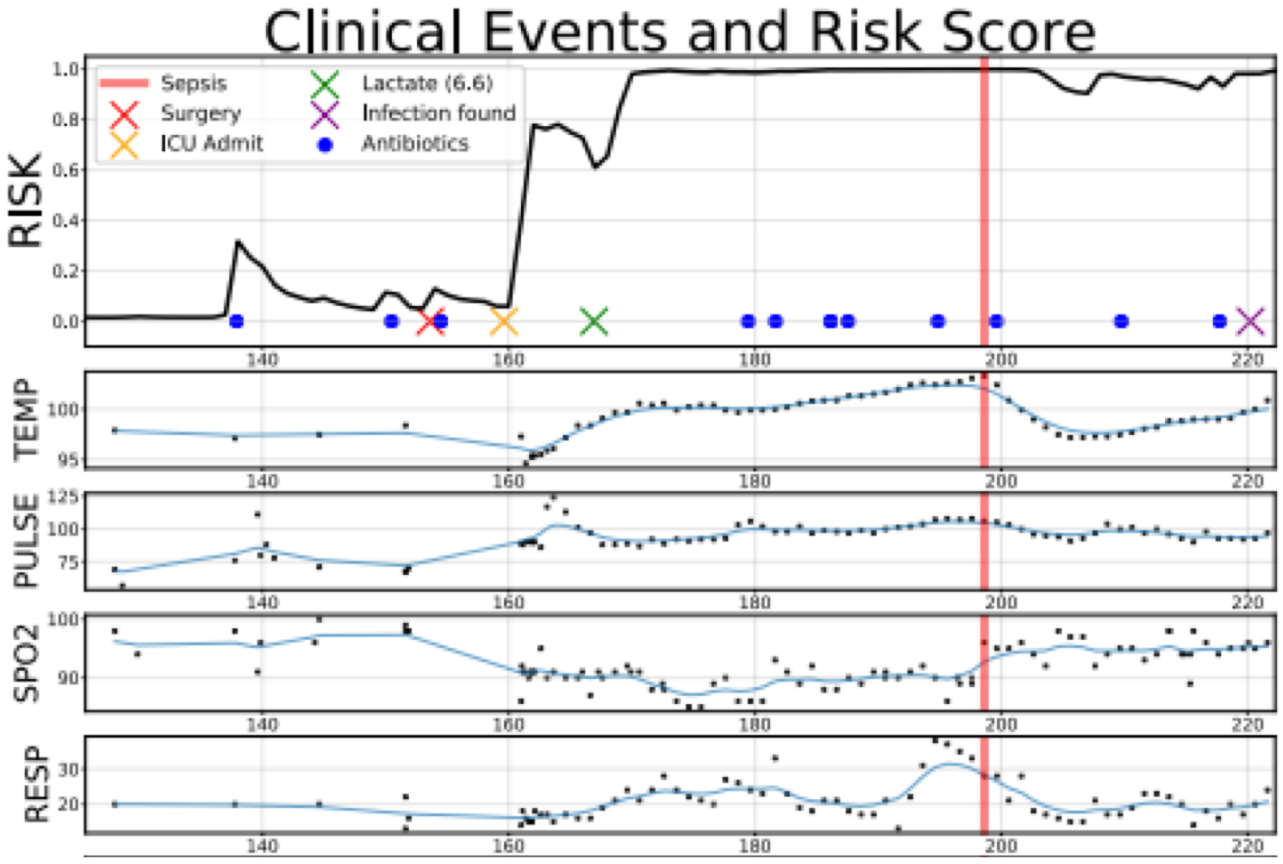
Joseph Futoma, Sanjay Hariharan, Katherine Heller JDF38,SH360,KH204@DUKE.EDU
Department of Statistical Science
Duke University, Durham, NC

Mark Sendak, Nathan Brajer MPD10,NJB23@DUKE.EDU
Institute for Health Innovation
Duke University, Durham, NC

Meredith Clement, Armando Bedoya, Cara O'Brien ME75,AB335,OBRIE028@DUKE.EDU
Department of M[illegible]
Duke University,

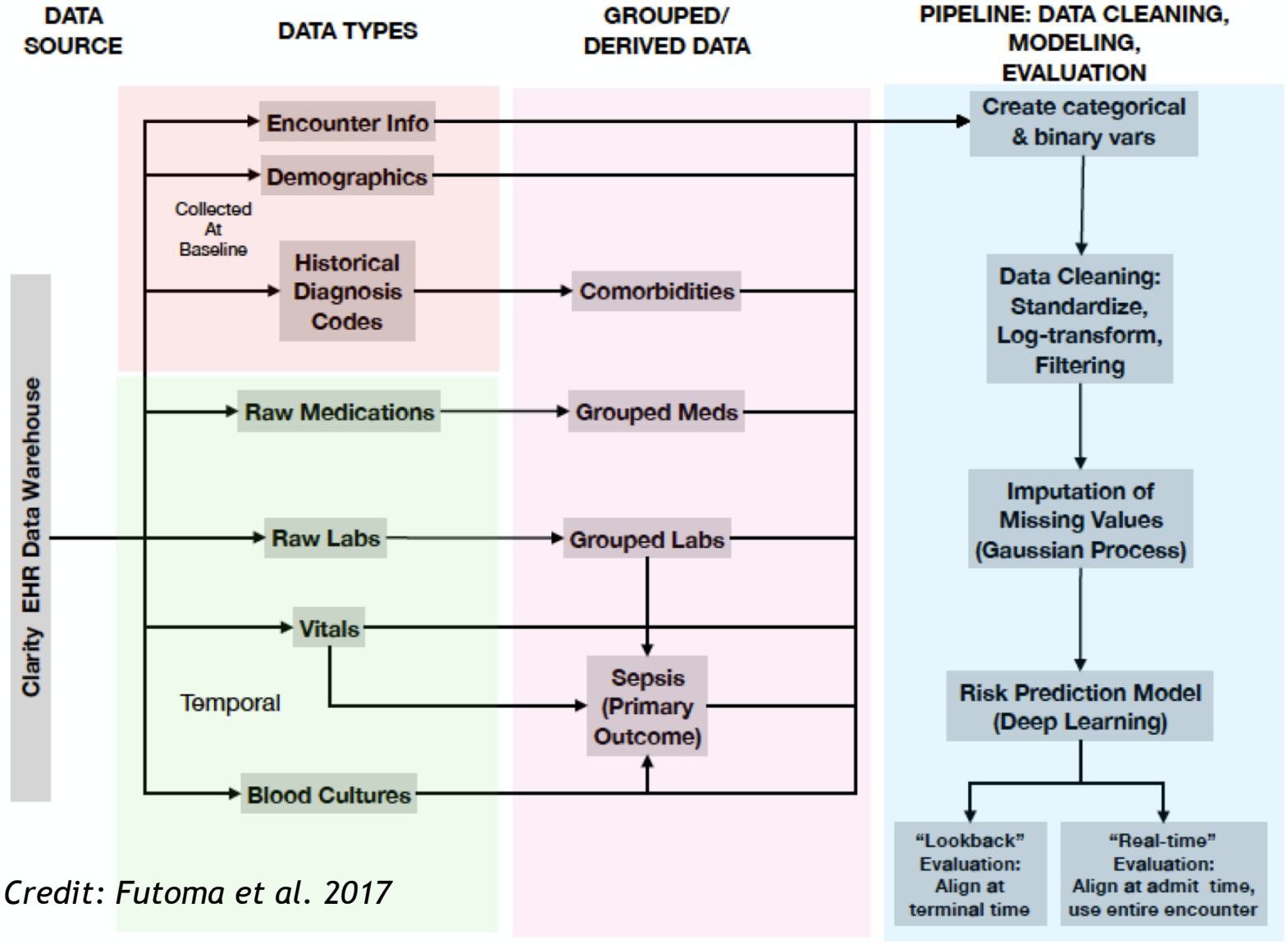


Risk Prediction Illustration



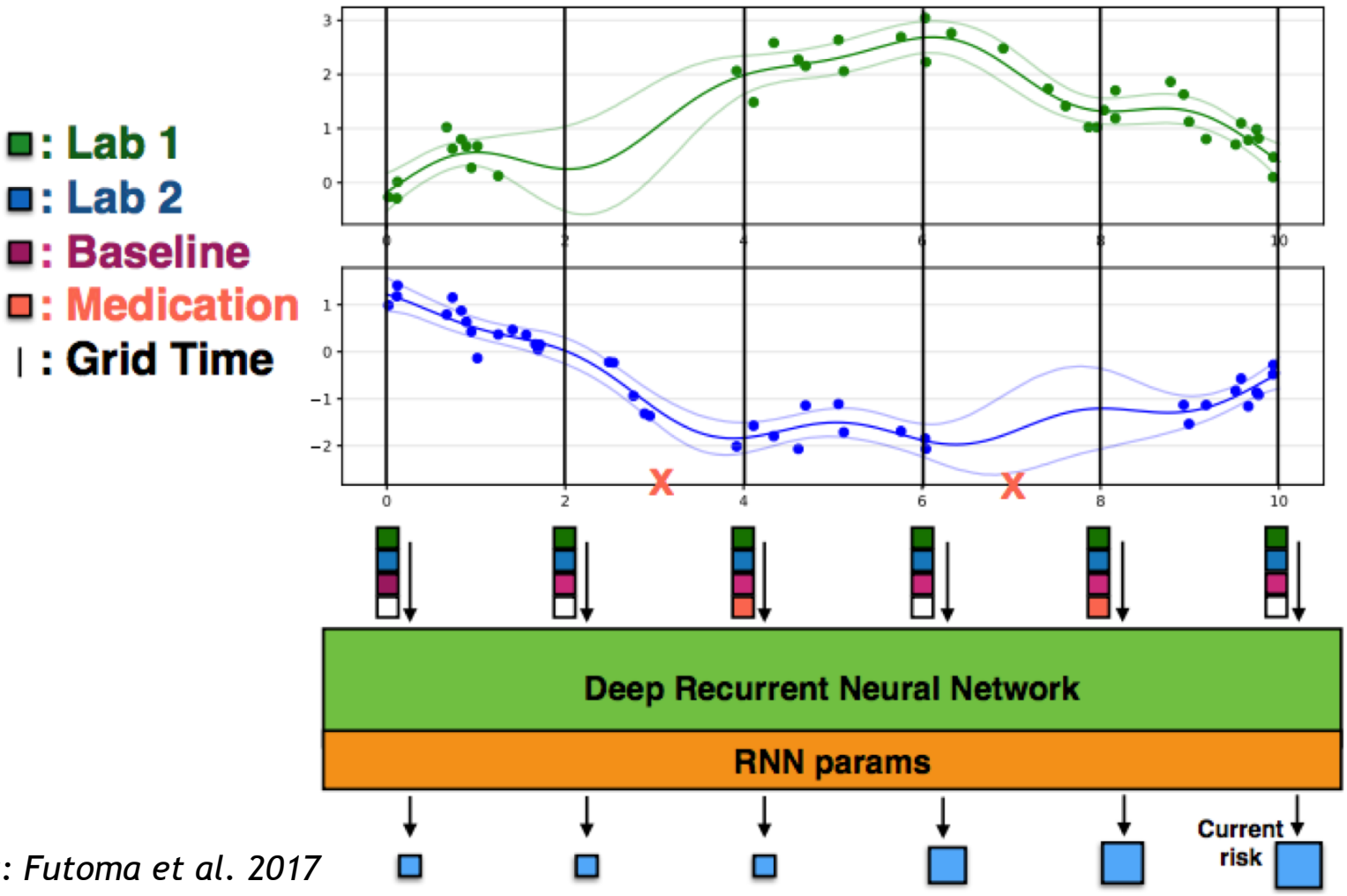
Credit: Futoma et al. 2017

Data



Credit: Futoma et al. 2017

Model



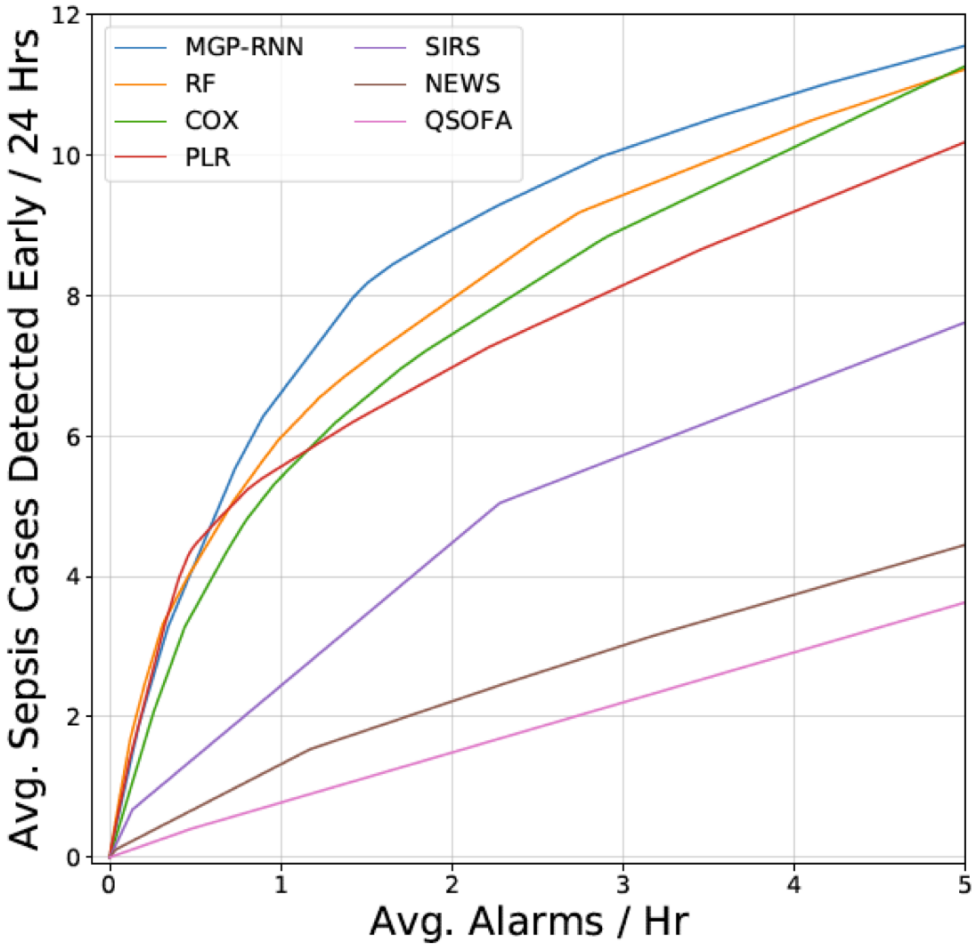
Credit: Futoma et al. 2017

Evaluation

AUC for sepsis classifier

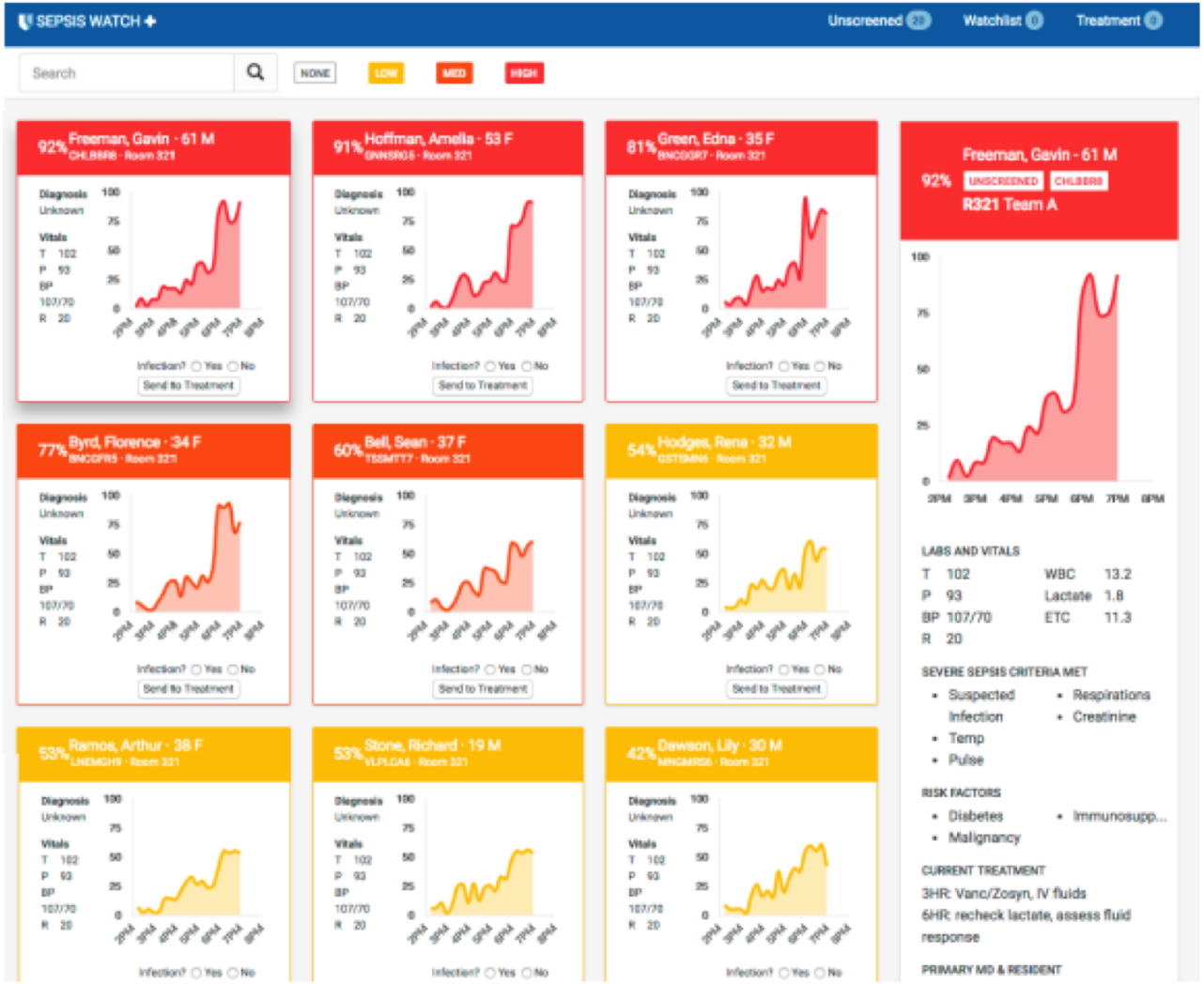
4 hrs beforehand

- 0.84 MGP-RNN
- 0.73 RNN
- 0.71 NEWS



Credit: Futoma et al. 2017

Deployment into Clinical Workflow

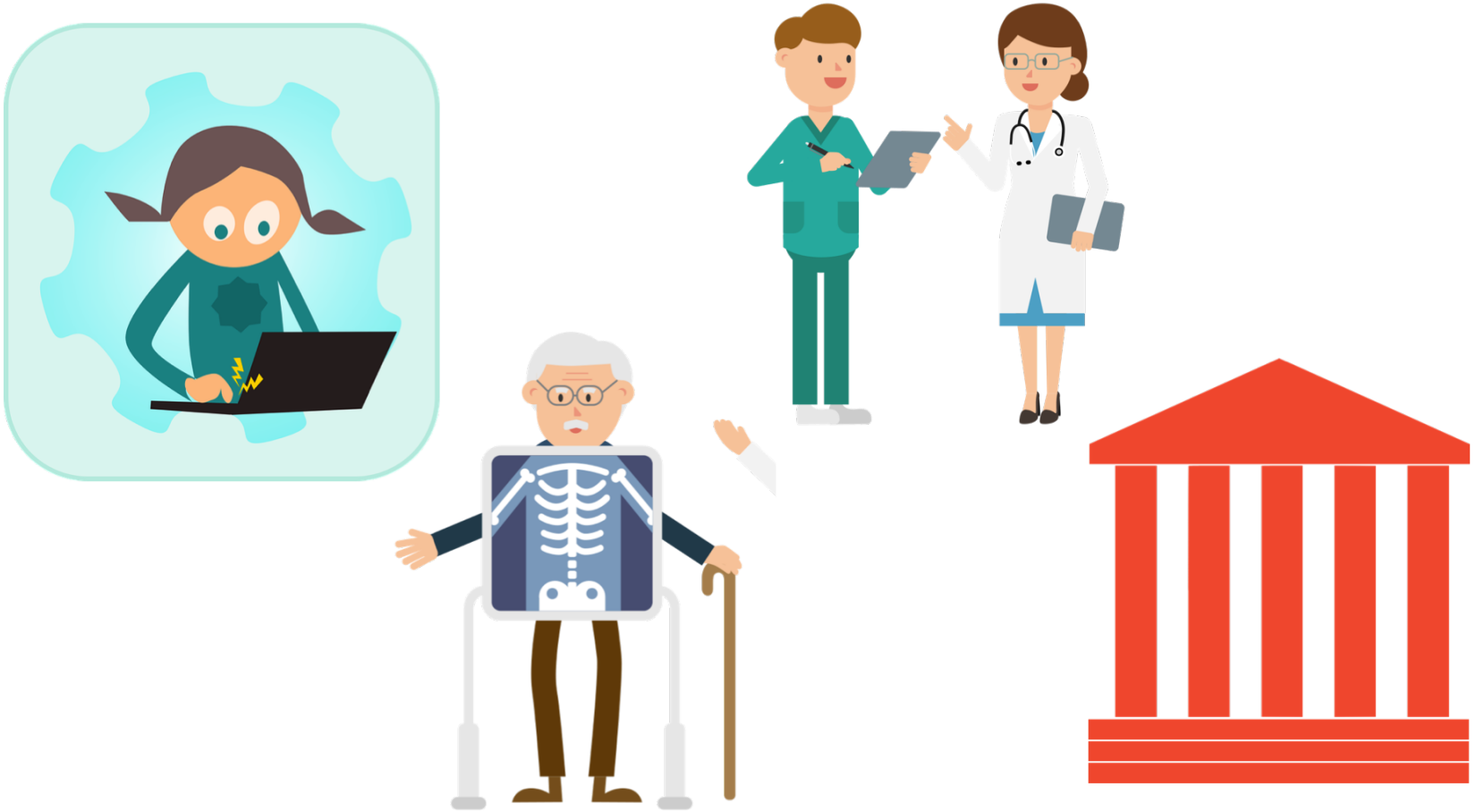


Credit: Futoma et al. 2017

AI is seen less of a threat and is welcomed by health professionals, research reveals

By **CATHERINE STURMAN** • Aug 06, 2018, 6:53AM

How to do ML + health collaboration?



Goal: Team Science at every stage of the process

Tutorial Goals

After this tutorial, you should (I hope) be able to:

- Collaborate with ML researchers more effectively
  - Frame relevant sub-problems in computational terms
  - Understand what is possible and what isn't possible

- Critically read an ML for health paper
  - Understand modeling approach (maybe not all details)
  - Critique performance measures and experimental design
    - Were appropriate baselines considered?
    - Are the evaluation metrics used reasonable?

Tutorial Outline

- 1:30-2:20 Part 1: Making and Evaluating Predictions
- 2:20-2:30 10 min Q&A break + **SURVEY**

- 2:30-3:20 Part 2: Learning Representations
- 3:20-3:30 10 min Q&A break

- 3:30-4:30 Part 3: Addressing Challenges in Health Data

Part 1 outline: Classic predictors

Evaluation Best Practices

- Metrics
- Calibration
- Decision-theory and utility

Quick tour of common predictors

- Linear models
- Decision trees and forests
- Neural nets (MLPs)

Part 2 outline: Modern methods

Learned Representations for

- Images (CNNs)
- Time series (RNNs)
- Text (RNNs and embeddings)

Tricks of the trade

Models that generate data

Part 3 outline: Challenges

M : “missing data”

I : ”incomplete labels” (semisupervised learning)

M : “multimodal data” (text + images + EHR codes)

I : “interpretability”

C : “causality”

S : “sequential decision making” (reinforcement learning)

Caveats

- Can't fit everything into this tutorial
- Please send feedback if I missed something!
  - mike@michaelchughes.com
- Known blindspots
  - Bias towards research at US institutions
  - Bias towards my conferences (NIPS/ICML/AISTATS)
  - Bias towards my areas (e.g. critical care)