UC San Diego
SCHOOL OF MEDICINE

Department of
BioMedical Informatics

Shamim Nemati, PhD
Director, Predictive Health Analytics
UC San Diego Health
shamim.nemati@alum.mit.edu
www.nematilab.info

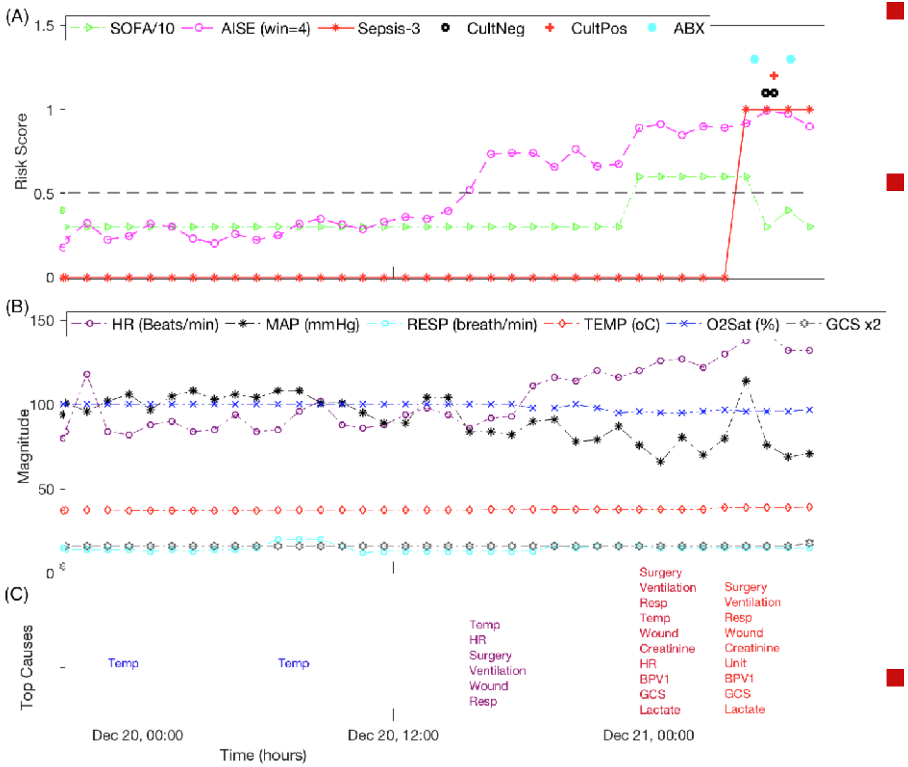
# Building and Evaluating Machine Learning Algorithms

Critical Care Medicine

Society of Critical Care Medicine

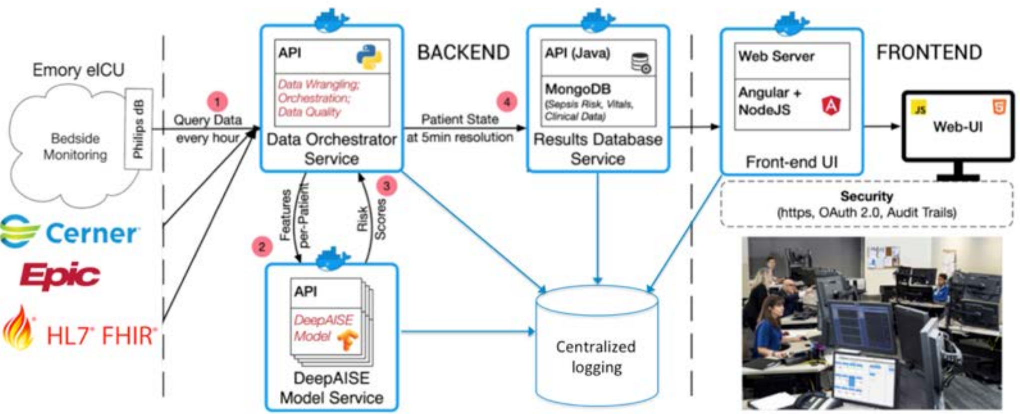
**An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU**

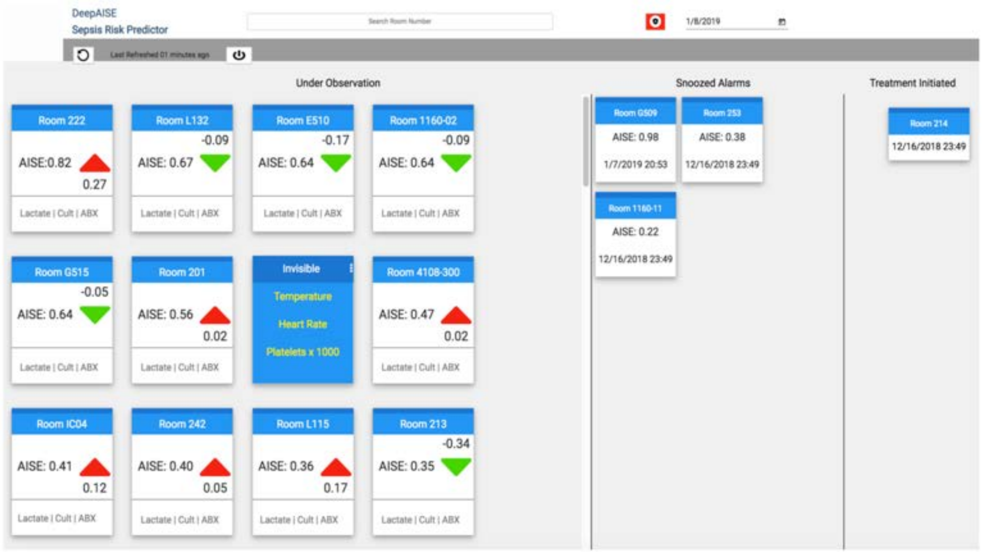
Sepsis-3 as a Clinical Construct:
ΔSOFA>2 + Clinical Suspicion

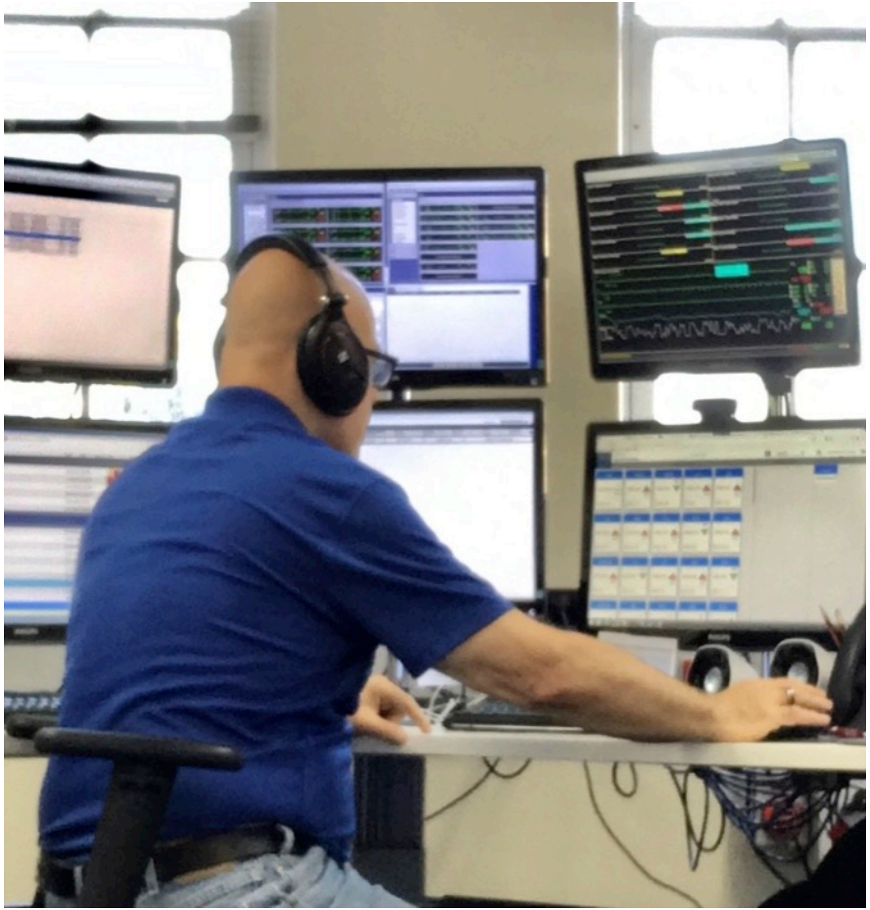



# Sepsis Predictive Analytics

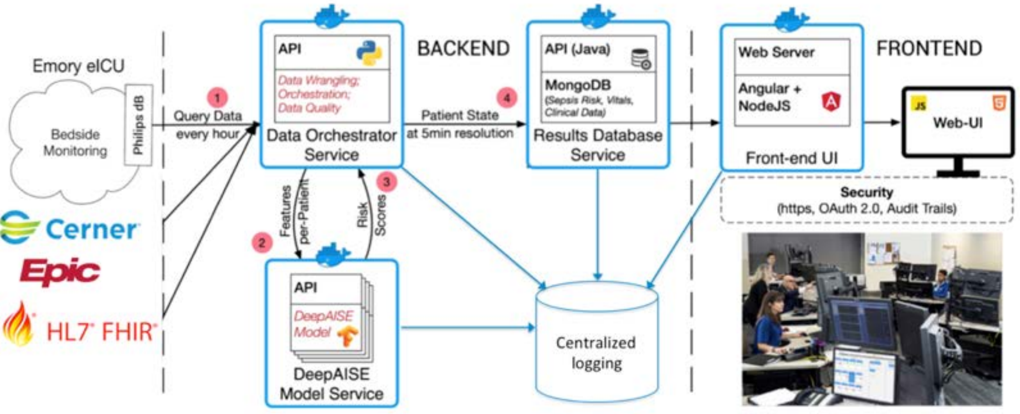
- Artificial Intelligence Sepsis Expert (AISE) (CCM, 2017-2018)
- Trained on 30,000 ICU patients from Emory Hospital and externally evaluated on 50,000 ICU patient from the MIMIC-III database
  - Input: commonly measured EHR data (labs, vitals, contextual info)
  - Output: 4-8 hours in advance risk of sepsis with an AUC of 0.85.
- Reveals top causes per each prediction (*interpretable/explainable*)

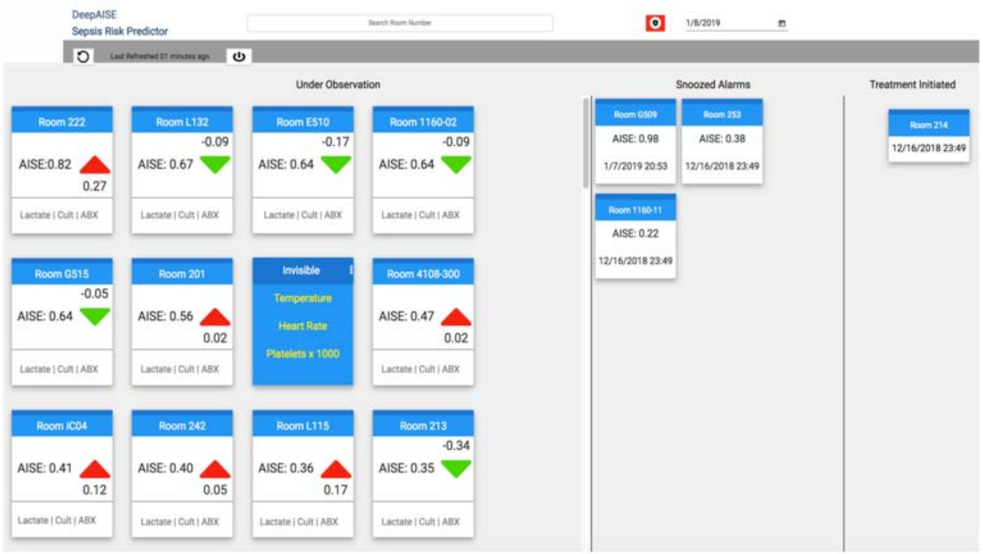


## Real-time analytics, workflow integration

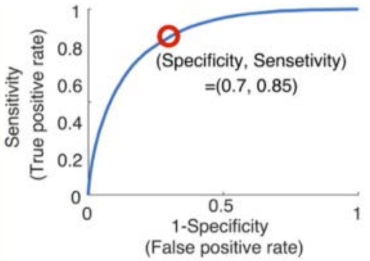



# The Five Rights of CDS:

1. The **right information** (reliable predictions, top causes / context)
2. To the **right person** (tele-ICU nurses)
3. In the **right CDS intervention** format (e.g., alert)
4. Through the **right channel** (e.g., EMR)
5. At the **right time**
   - Actionable window (4-12 hours in advance)
   - Prior to clinical recognition
   - Prior to initiation of treatments

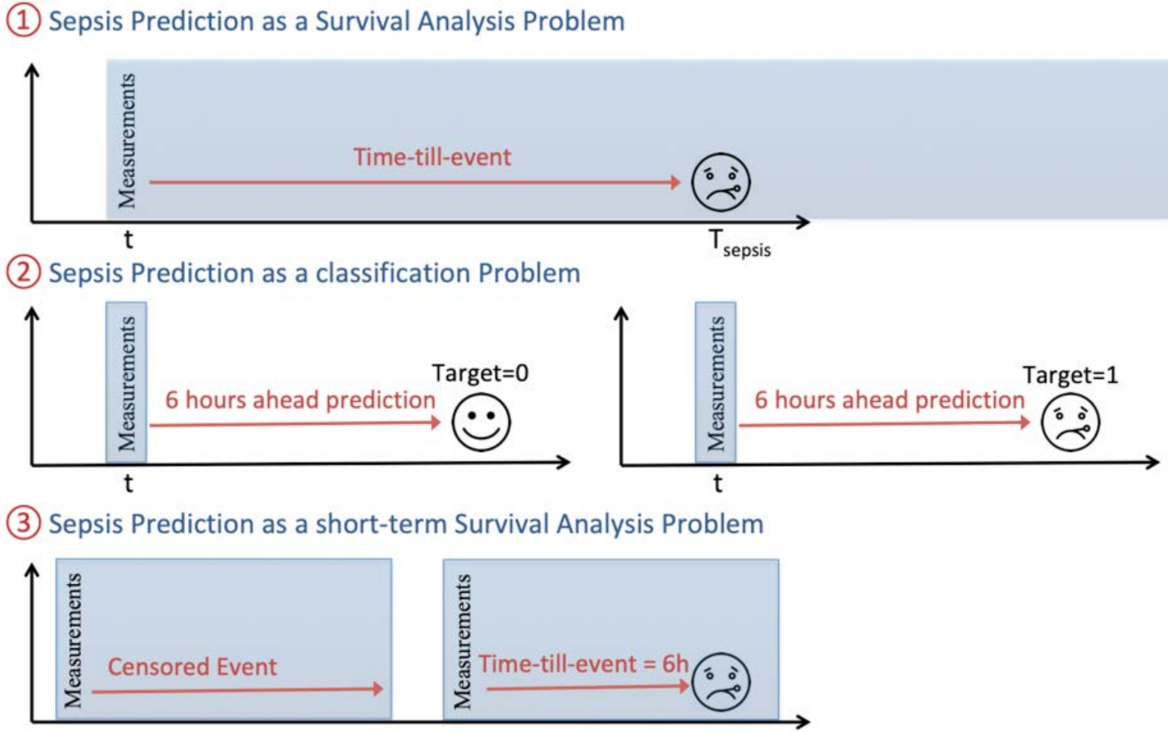
Real-time analytics, workflow integration




# How many evaluation protocols are there?

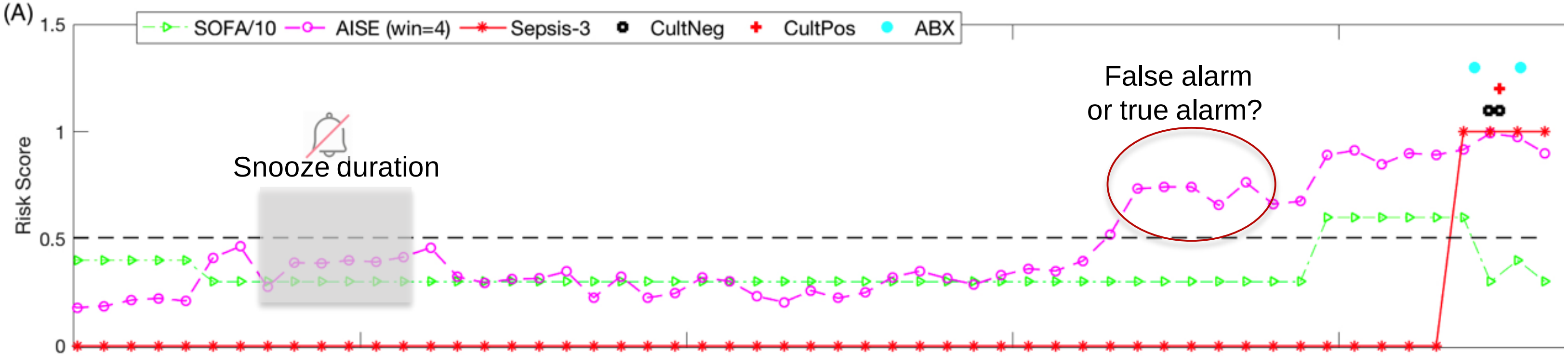
1. Infinite horizon
2. Look-back
3. Sequential
4. Case-control
5. Max value
6. One-shot (at time of admission or time of clinical suspicion
7. ...

Can we compare AUCs, PPVs, across all these protocols? *No!*



# Designing Clinically Meaningful Evaluation Metrics

- **Utility** of a model is intimately connected to the clinical protocols under which the model is evaluated
  - Snooze
  - Prediction horizon and actionable window (e.g., 6hr)
  - Setting (e.g., ED versus ICU)

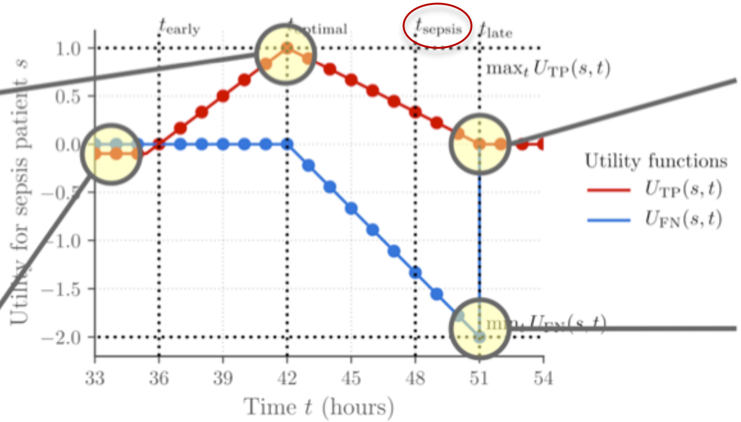


# Early Prediction of Sepsis from Clinical Data: the PhysioNet/Computing in Cardiology Challenge 2019

**Utility function:** rewards early predictions and penalizes late predictions as well as false alarms.

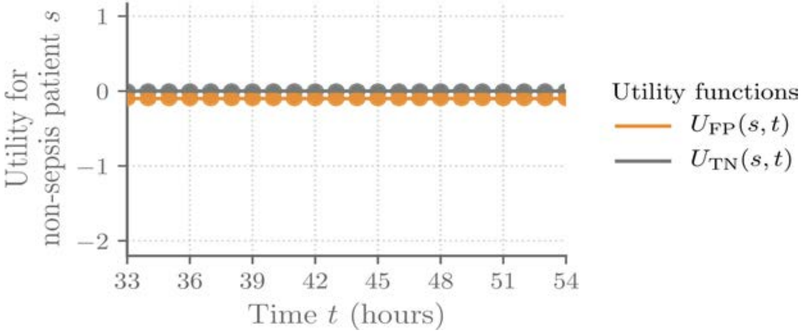
Early prediction: most beneficial

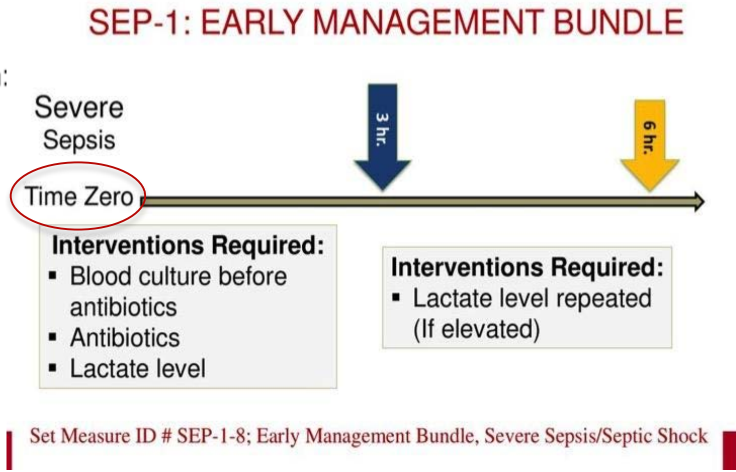
Too early or false alarm: poor use of hospital resources and overtreatment



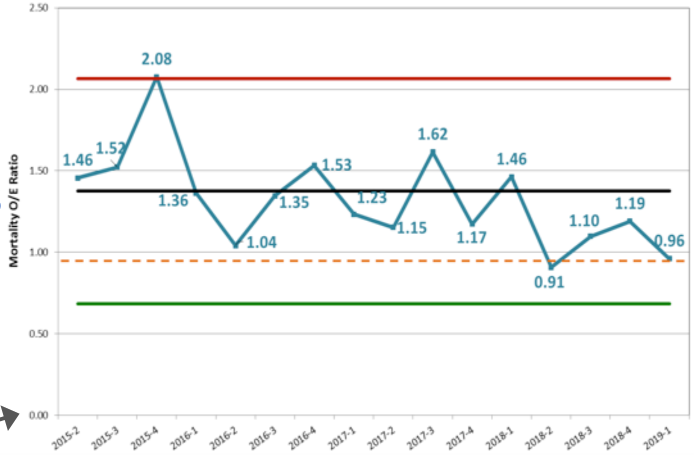

Late prediction: not beneficial

Missed prediction: very harmful



# Designing Prospective Evaluation Methods




- Multiple Reader Multiple Case (MRMC)
- Quality Improvement (QI) study
  - Improving SEP-1 compliance rate
  - Lowering Mortality O/E ratio
  - Hospital ranking
- Single Center (pilot) and Multicenter clinical trials
  - Mortality, LOS, days on MV, Pressors, etc.

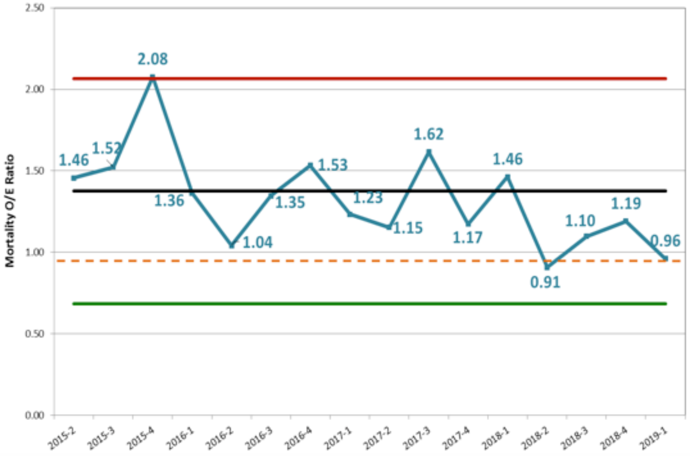
# Designing ~~Clinically Meaningful~~ Evaluation Protocols for National Ranking

- Task: Prediction of shock
- End-point: O/E (Observed over ~~Expected Mortality~~)

**Model Results (Significant Predictors)**

| Explanatory Variable | Beta | OR |
|---|---|---|
| Intercept | -4.644 | |
| Vent on Admission Day | 1.445 | 4.242 |
| Female, Age >= 85 | 1.437 | 4.208 |
| ECMO on Admission Day | 1.434 | 4.195 |
| DIC | 1.380 | 3.975 |
| Necrotizing Fasciitis | 1.318 | 3.737 |
| Tumor Lysis Syndrome | 1.316 | 3.730 |
| Shock | 1.269 | 3.557 |

:
.

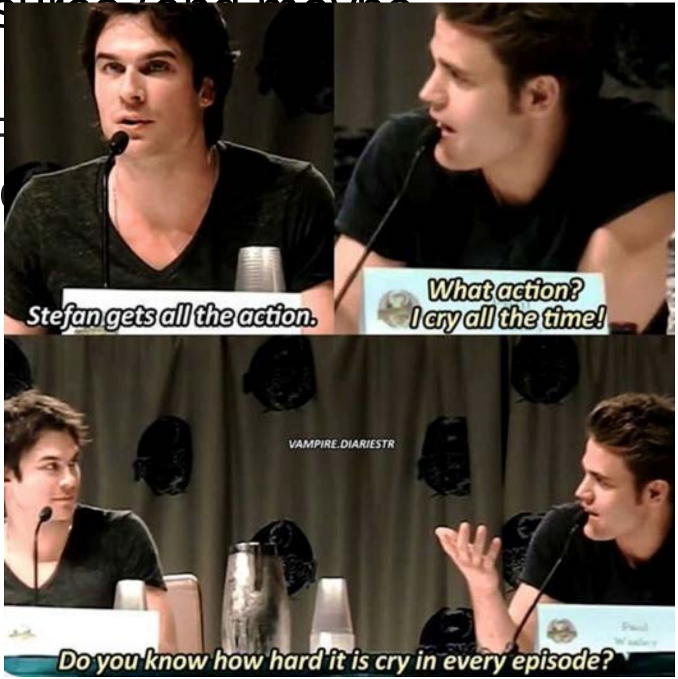



Snooze for
3 hours

Snooze for
10 hours

# Conclusion

- Performance metrics that take into account care protocols, case mix complexity (CPT modifiers), cost, quality meas
hospital rankings)
  - Can this AI tool make me more compliant with SEP
  - Can this AI tool reduce cost within a given MS-DR




- All the buzz is about AI, but QI gets all the
  - To encourage adoption of AI/ML we need to rethir
metrics and encourage QI-AI team building

Do you know how hard it is to change a menu item in Epic?

# Thank You!

shamim.nemati@alum.mit.edu
http://www.nematilab.info
@ShamimNemati

## Funding Acknowledgements

- Gordon and Betty Moore Foundation, “SEP1+: A Composite Measure to Accurately Assess Early Sepsis Management” (2020-2021)
- National Institutes of Health (NIH), “Deep Learning and Streaming Analytics for Prediction of Adverse Events in the ICU” (2015-2021)
- Biomedical Advanced Research and Development Authority (BARDA), “Multicenter Deployment and Evaluation of DeepAISE-on-FHIR for Early Prediction of Sepsis” (2019-2020)
- National Science Foundation (NSF), “Leveraging Heterogeneous Data Across International Borders in a Privacy Preserving Manner for Clinical Deep Learning” (2018-2021)
- National Institutes of Health (NIH), “DeepFHIR: A Fast Healthcare Interoperability Resources (FHIR)-enabled Deep Learning Platform for Prediction of Decompensation in Critically Ill Patients” (2017-2019)