# Multiple Instance Learning for ECG Risk Stratification

Divya Shanmugam, Davis Blalock, John Guttag







## Feed in the whole signal?

• High sampling rate (11 million samples per patient)

#### Feed in the whole signal?

• High sampling rate (11 million samples per patient)

#### **Use statistical features?**

- Typical of existing work
- Approximates characteristics of risk signals

#### Feed in the whole signal?

• High sampling rate (11 million samples per patient)

#### **Use statistical features?**

- Typical of existing work
- Approximates characteristics of risky signals

## Learn what a risky set of heartbeats look like!

# We reframe the risk stratification task as a **multiple instance learning** problem.

- \* Labels of collections of instances are available, but labels of instances are not
- \* Assumption: patients at higher risk of cardiovascular death have more risky instances than those at lower risk



## Our method consists of three steps: instance extraction, classification, and aggregation



Classification

$$\mathsf{F}(\mathsf{M}) \to [0,1]$$

#### Aggregation



## Produces a state-of-the-art ECG-based risk score for cardiovascular death



#### **Takeaways**

- \* We propose a general-purpose method for incorporating very long time series into risk models
- \* Using this method, we produce a state-of-the-art ECGbased risk score!

Thanks!