Life and Death with Congestive Heart Failure

Peniel Argaw, Irene Chen, Sebastian Gehrmann, Harlin Lee, Alisha Saxena

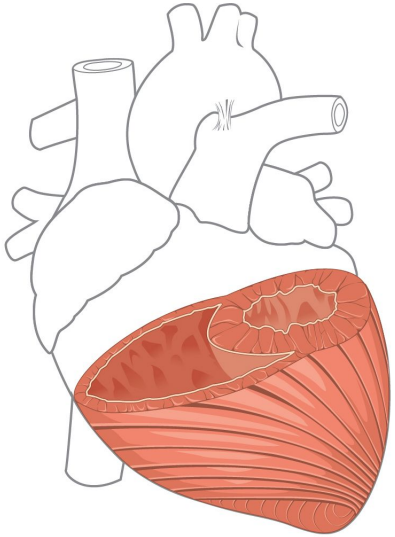
With help from Dr. Steven Horng, Dr. Sandeep Gangireddy

5.7 million

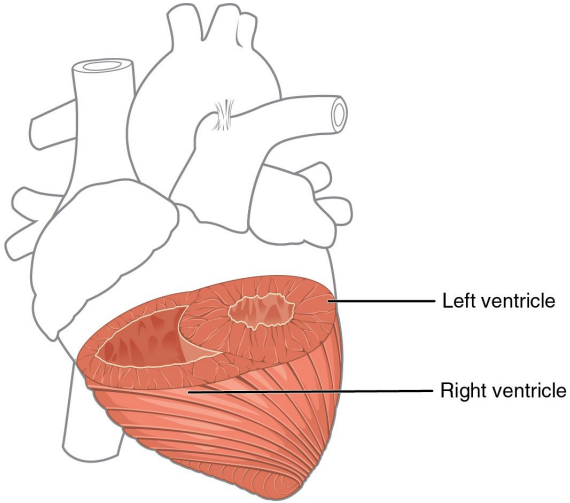
adults in US with congestive heart failure. **50%** die within 5 years.

[CDC 2016]

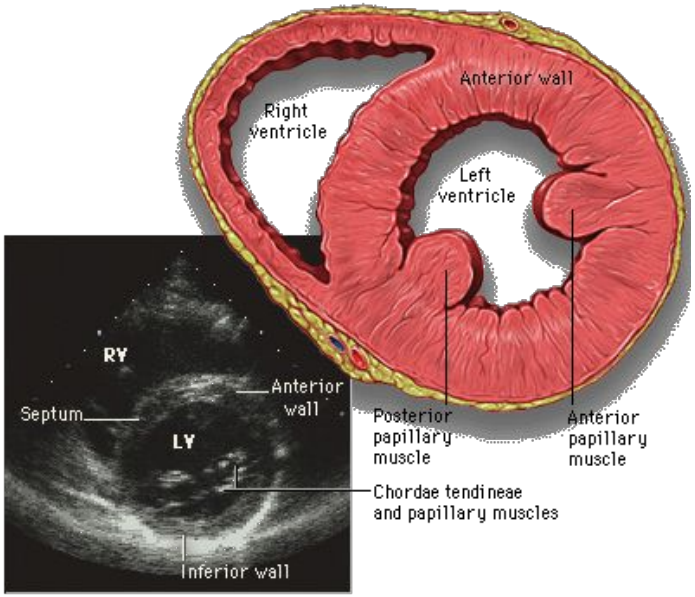
What is Congestive Heart Failure?



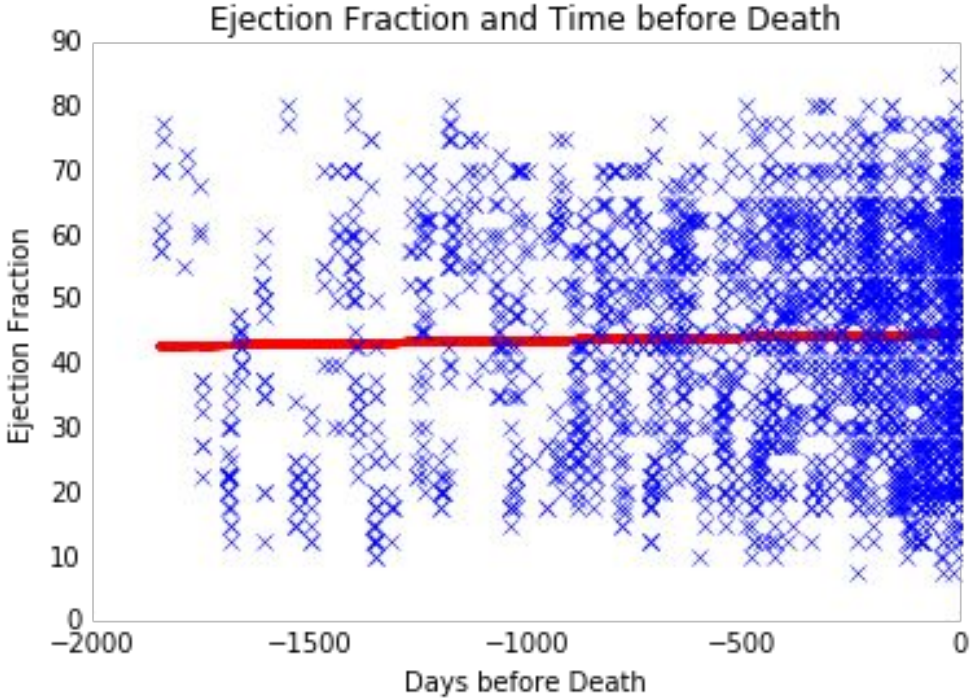
Relaxed



Contracted



Ejection Fraction != Time until Death



Class	Patient Symptoms
I	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnea (shortness of breath).
II	Slight limitation of physical activity. Comfortable at rest. Ordinary physical activity results in fatigue, palpitation, dyspnea (shortness of breath).
III	Marked limitation of physical activity. Comfortable at rest. Less than ordinary activity causes fatigue, palpitation, or dyspnea.
IV	Unable to carry on any physical activity without discomfort. Symptoms of heart failure at rest. If any physical activity is undertaken, discomfort increases.

Class	Objective Assessment
A	No objective evidence of cardiovascular disease. No symptoms and no limitation in ordinary physical activity.
B	Objective evidence of minimal cardiovascular disease. Mild symptoms and slight limitation during ordinary activity. Comfortable at rest.
C	Objective evidence of moderately severe cardiovascular disease. Marked limitation in activity due to symptoms, even during less-than-ordinary activity. Comfortable only at rest.
D	Objective evidence of severe cardiovascular disease. Severe limitations. Experiences symptoms even while at rest.

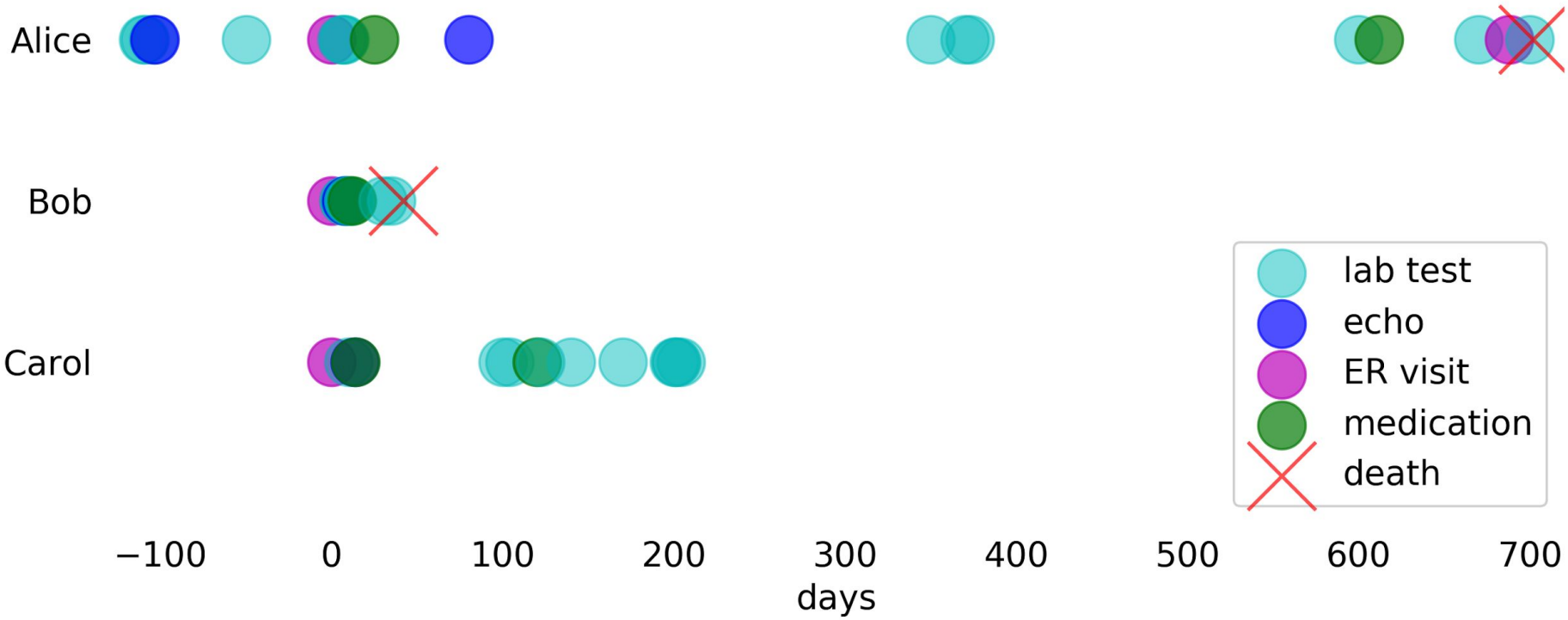
[NYHA]

How can we better model the **disease progression**?

Data

- **1,961** Beth Israel Deaconess Medical Center (BIDMC) patients
  - **34% mortality rate**
  - Average age of 74
- Electronic health records
  - labs (3m)
  - emergency room admissions (3k)
  - echocardiograms (9k)
  - medications (3k)

Example Patient Timelines



Supervised Learning Approaches

Overview of Ideas

1. Predict **mortality** at a specific time, use regression scores to represent the CHF state. (baseline LR, SVM, RF models)

2. Predict **mortality** as a multi-task learning problem. (Convex Fused Sparse Group Lasso, CNN)

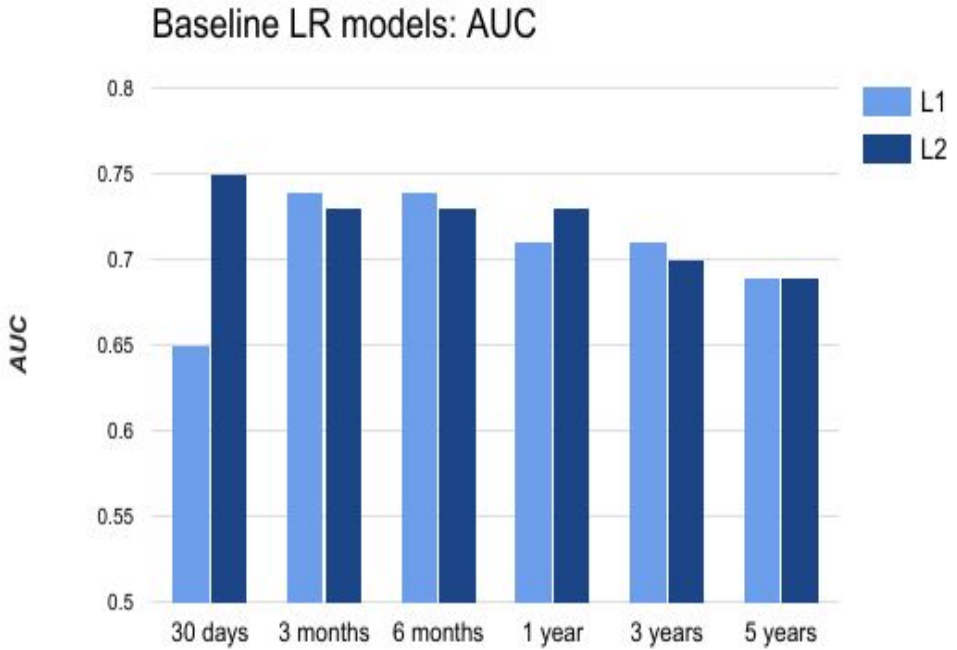
3. Predict **time until next visit**. (RNN)

Input Features

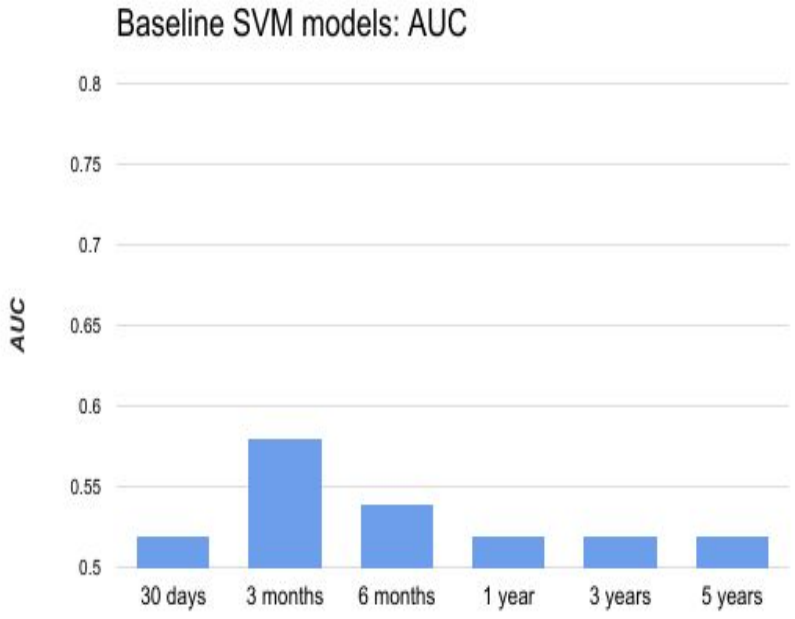
Date Source	**Variables**
First ER visit (12 total)	Age, Sex, Diabetes, Chronic kidney failure, Triage acuity, Triage vitals
Labs done within 72 hours (31 total)	ANION GAP, BASOS, CALCIUM, CHLORIDE, **CREAT**, EOS, GLUCOSE, HCT, HGB, INR(PT), LYMPHS, MAGNESIUM, MCH, MCHC, MCV, MONOS, NEUTS, PHOSPHATE, PLT COUNT, POTASSIUM, PT, PTT, RBC, RDW, SODIUM, TOTAL CO2, **UREA N**, WBC, cTropnT, estGFR, proBNP
Echo done within 72 hours (21 total)	Aorta - Ascending, Aorta - Valve Level, Aortic Valve - LVOT Diam, Aortic Valve - Peak Velocity, BP, Left Atrium - Four Chamber Length, Left Atrium - Long Axis Dimension, Left Ventricle - Diastolic Dimension, **Left Ventricle - Ejection Fraction**, Left Ventricle - Fractional Shortening, Left Ventricle - Inferolateral Thickness, Left Ventricle - Septal Wall Thickness, Left Ventricle - Systolic Dimension, Mitral Valve - A Wave, Mitral Valve - E Wave, Mitral Valve - E Wave Deceleration Time, Mitral Valve - E/A Ratio, Right Atrium - Four Chamber Length, Height, Weight

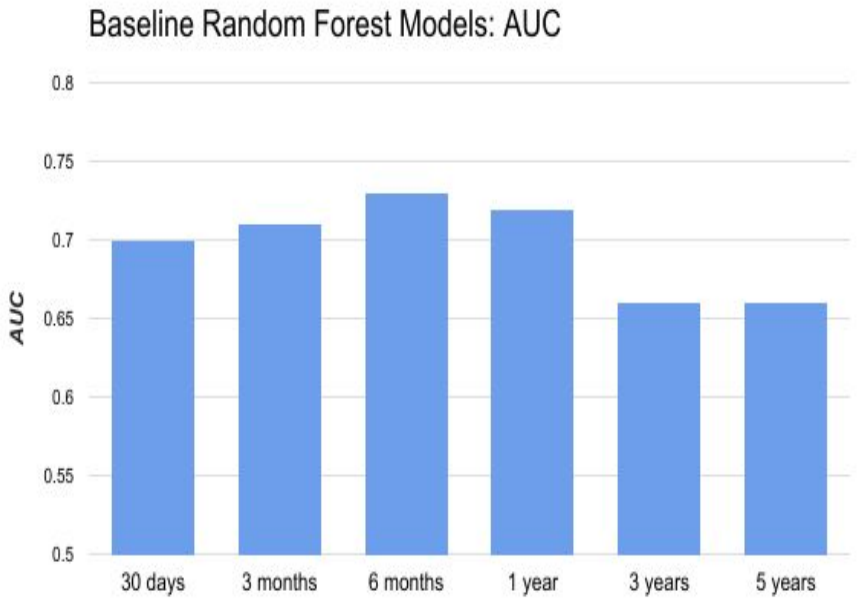
Output: Death at different times

Time point	Number of deaths
30 days	107
3 months	181
6 months	259
1 year	359
3 years	560
5 years	606



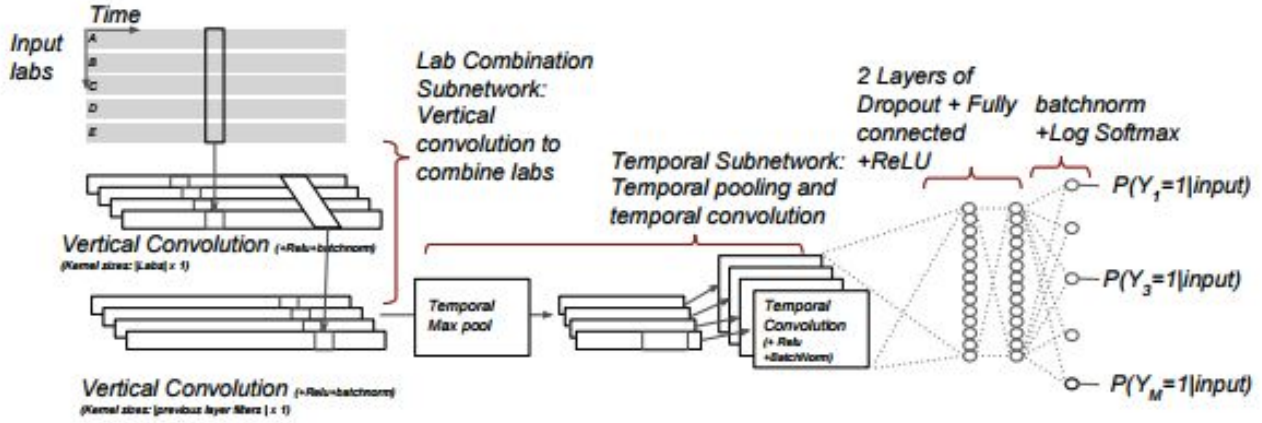
Baseline Models (SVM, Random Forest)





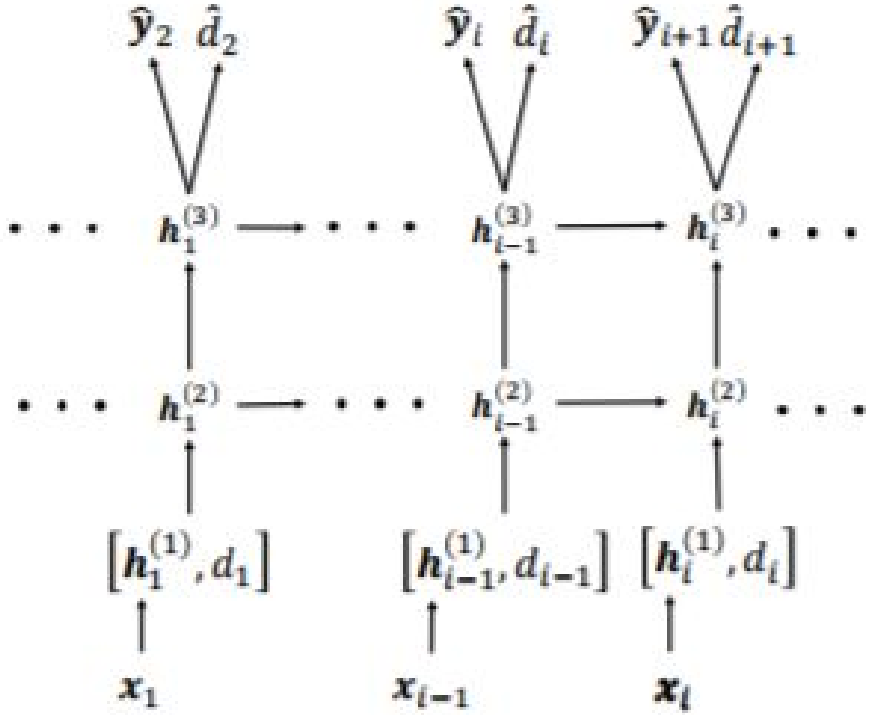
Multi-task prediction of days until death

- Convex Fused Sparse Group Lasso [Zhou et al, KDD ‘12]
- CNN [Razavian et al, ‘16]



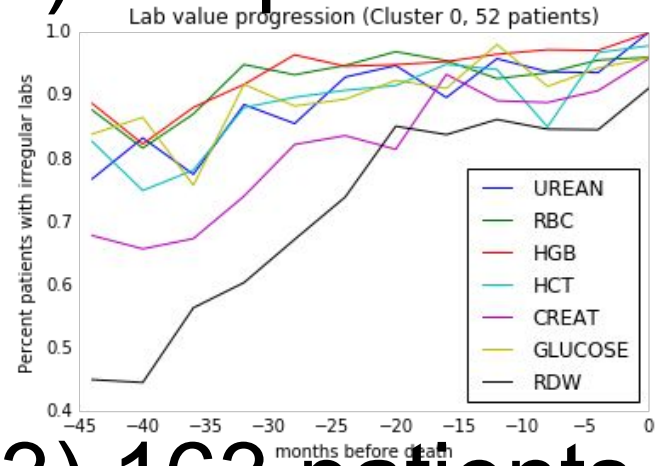
Predicting time until next visit and future diagnoses

- Model
  - RNN
- Input
  - Admissions
  - ICD9 diagnoses

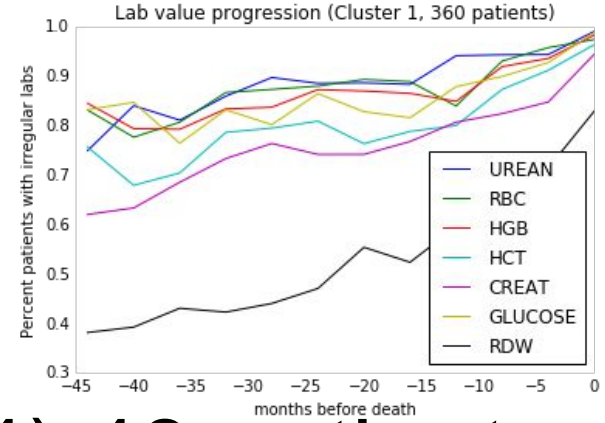


Unsupervised Learning Approaches

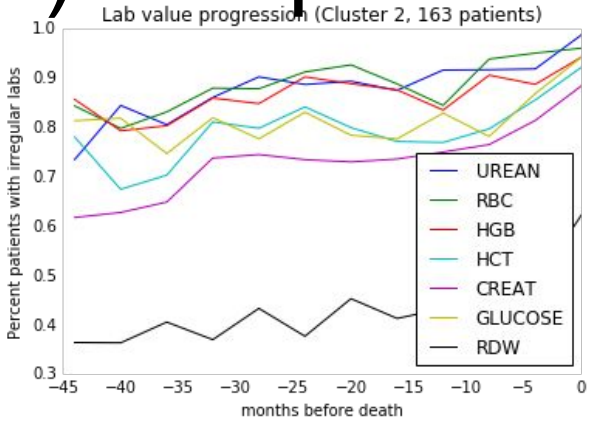
1) 52 patients



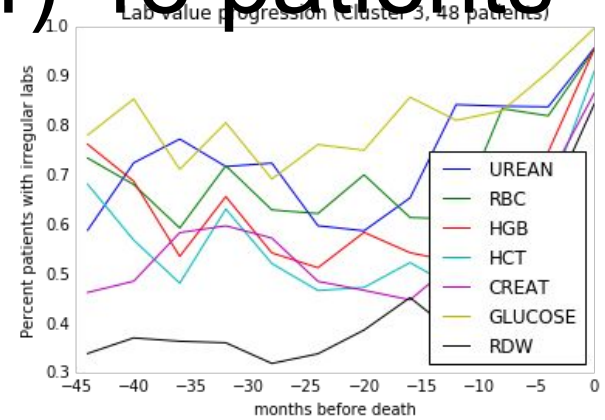
2) 360 patients



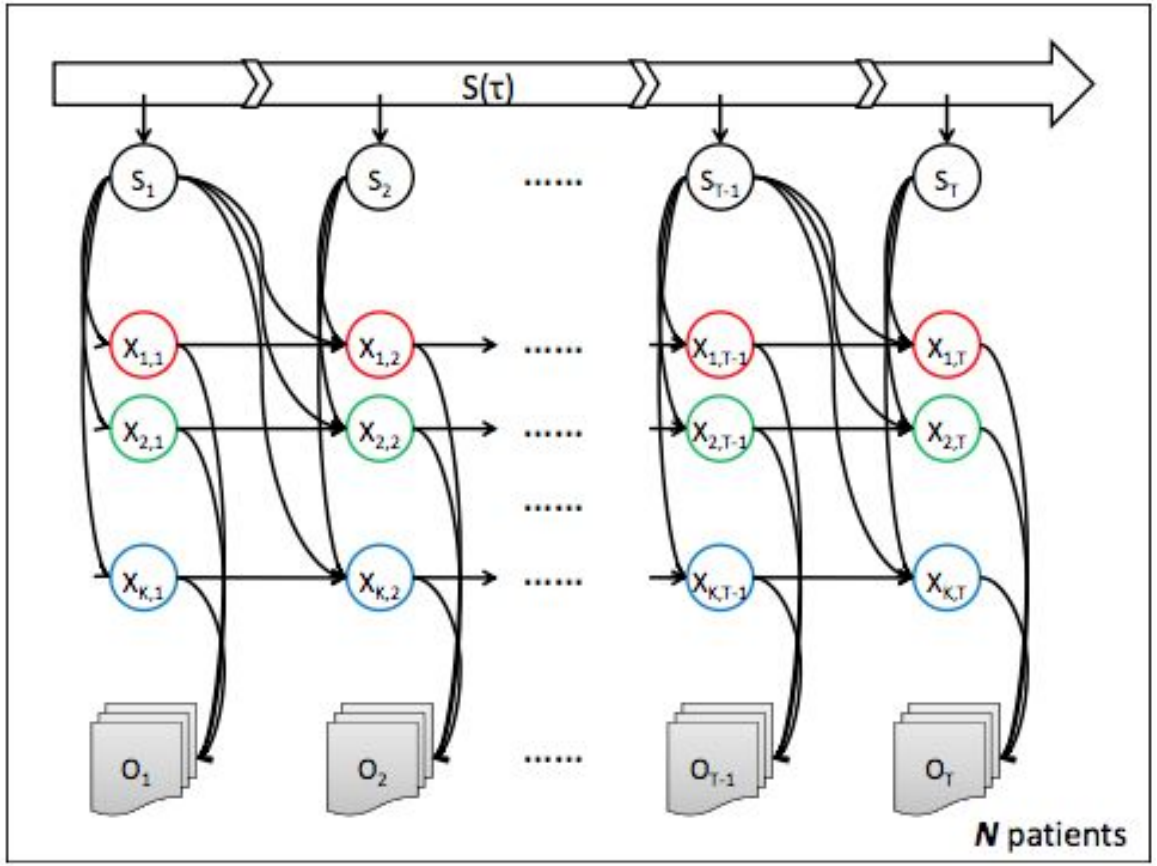
3) 163 patients



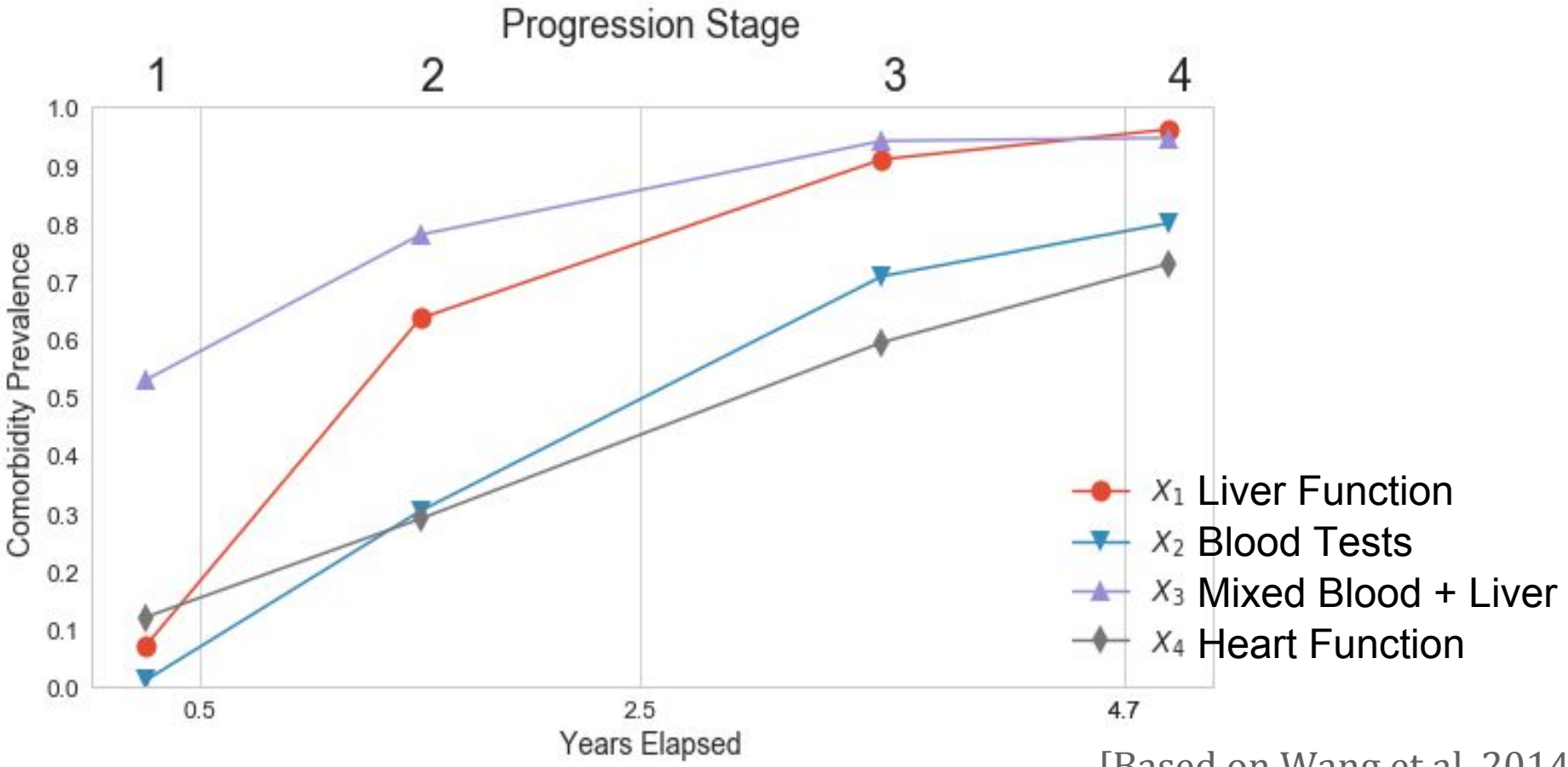
4) 48 patients



[Based on Haldar et al, 2008]



[Wang et al, 2014]



[Based on Wang et al, 2014]

X_1: Electrolyte Tests for Liver Function			X_2: Blood Count and Tests		
0.51	Urea Nitrogen (BUN)	Kidney and Liver	0.21	Red Cell distribution width (RDW)	Anemia
0.20	Potassium	Kidney	0.20	Lymphocytes	Heart Failure
0.19	Chloride	Kidney	0.18	Neutrophils	Heart Failure
0.12	Phosphate	Kidney	0.17	International Ratio of Prothrombin Time	Blood Clotting
0.09	pH	Kidney	0.16	Prothrombin Time	Blood Clotting
X_3: Mixed Bag			X_4: Heart Function		
0.42	Red Blood Cell	Anemia	0.15	Aspartate Aminotransferase	Heart and Liver
0.13	Hyaline	Dehydration	0.14	B-type Natriuretic Peptide	Heart Failure
0.13	Partial Thromboplastin Time	Blood Clotting	0.14	Hematocrit	Anemia or Heart
0.10	Cardiac-specific Troponin	Heart Attack	0.11	Hemoglobin	Anemia
0.07	Creatine Kinase	Kidney	0.08	Platelet Count	Heart Disease

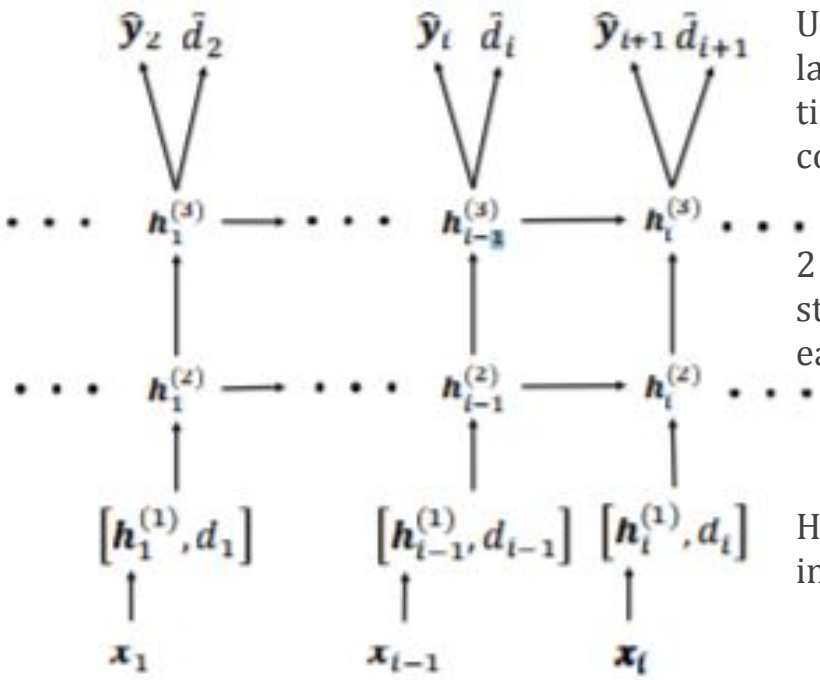
Leak terms: Hemoglobin (0.76), Creatinine (0.75), Urea Nitrogen (0.74), Red Blood Cell (0.72), Hematocrit (0.70), Glucose (0.64), Neutrophils (0.45), White Blood Cell (0.43), ...

[Based on Wang et al, 2014]

1. CHF is a **complex disease** that is difficult to stage quantitatively.
2. We can use **death and future hospital admissions** as proxies for CHF state.
3. Disease progression can be learned without labels using **abnormal lab values**.

Predicting time until next visit and future diagnoses

- Model
  - RNN
- Input
  - Admissions
  - ICD9 diagnoses
- Output
  - Time until next visit
  - Medical codes that occur in next visit



Multi-task prediction of time until death

- Input
  - Features
- Output
  - Days until death (or -1)
  - Or binary state
    - [0, 0, 1, 1, 1, 1] if dead between 2nd and 3rd time points

- Models
  - Convex Fused Sparse Group Lasso [Zhou et al, KDD ‘12]
  - CNN [Razavian et al, ‘16]