# Machine Learning and Predictive Models in Hepatitis C

**Akbar K. Waljee, MD, Msc**
Associate Professor of Internal Medicine
University of Michigan Health Systems
Ann Arbor VA Center of Excellence

VA CCMR

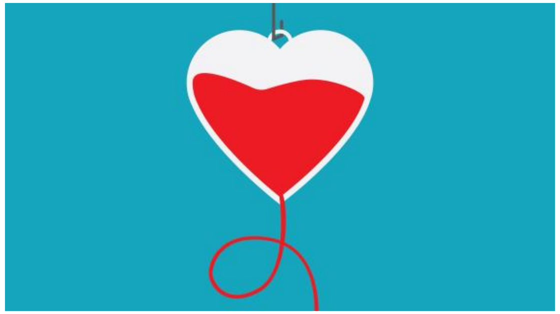
MICHIGAN MEDICINE
UNIVERSITY OF MICHIGAN

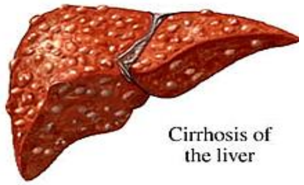
# Goals of the Talk

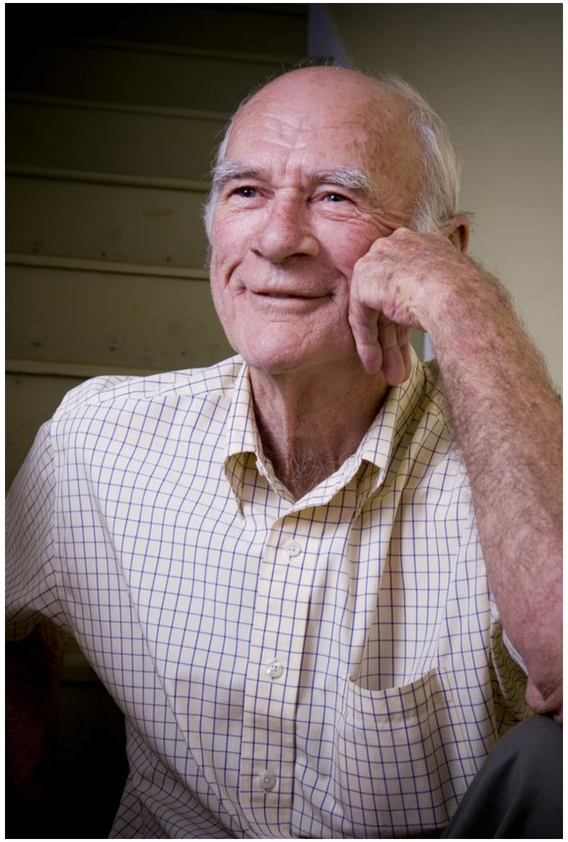
- Introduction to a Clinical Problem
- How can Machine Learning help?
- Examples in HCV:
  - HALT-C/Michigan Medicine Cohorts
  - VA Cohort
- Implications for Medicaid patients

# Mr. S

- 64 yo male
- HCV RNA+
- Received blood transfusion in 1978
- Early stage cirrhosis



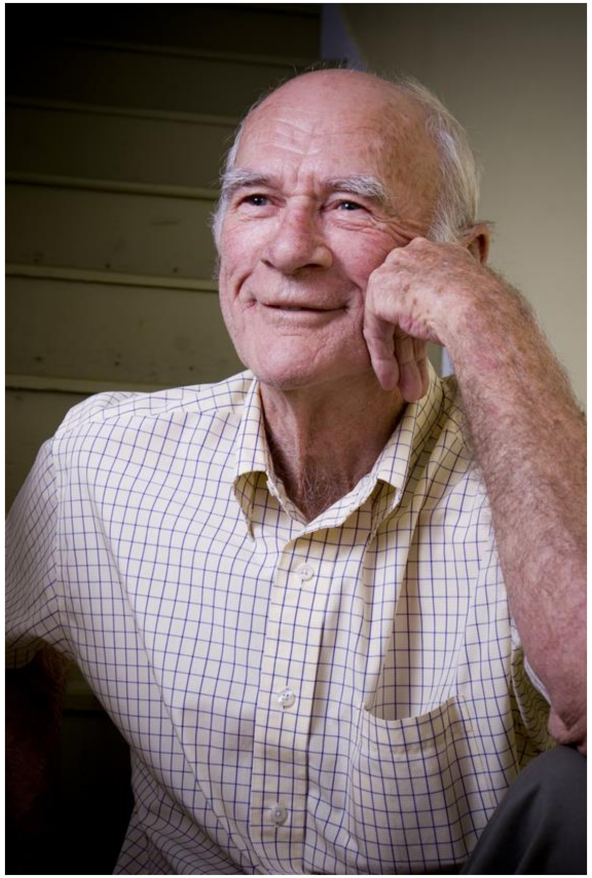


# Ms. J

- 23 yo female
- HCV RNA+
- Active IVDU
- Recently acquired HCV

# Do you treat them the same?



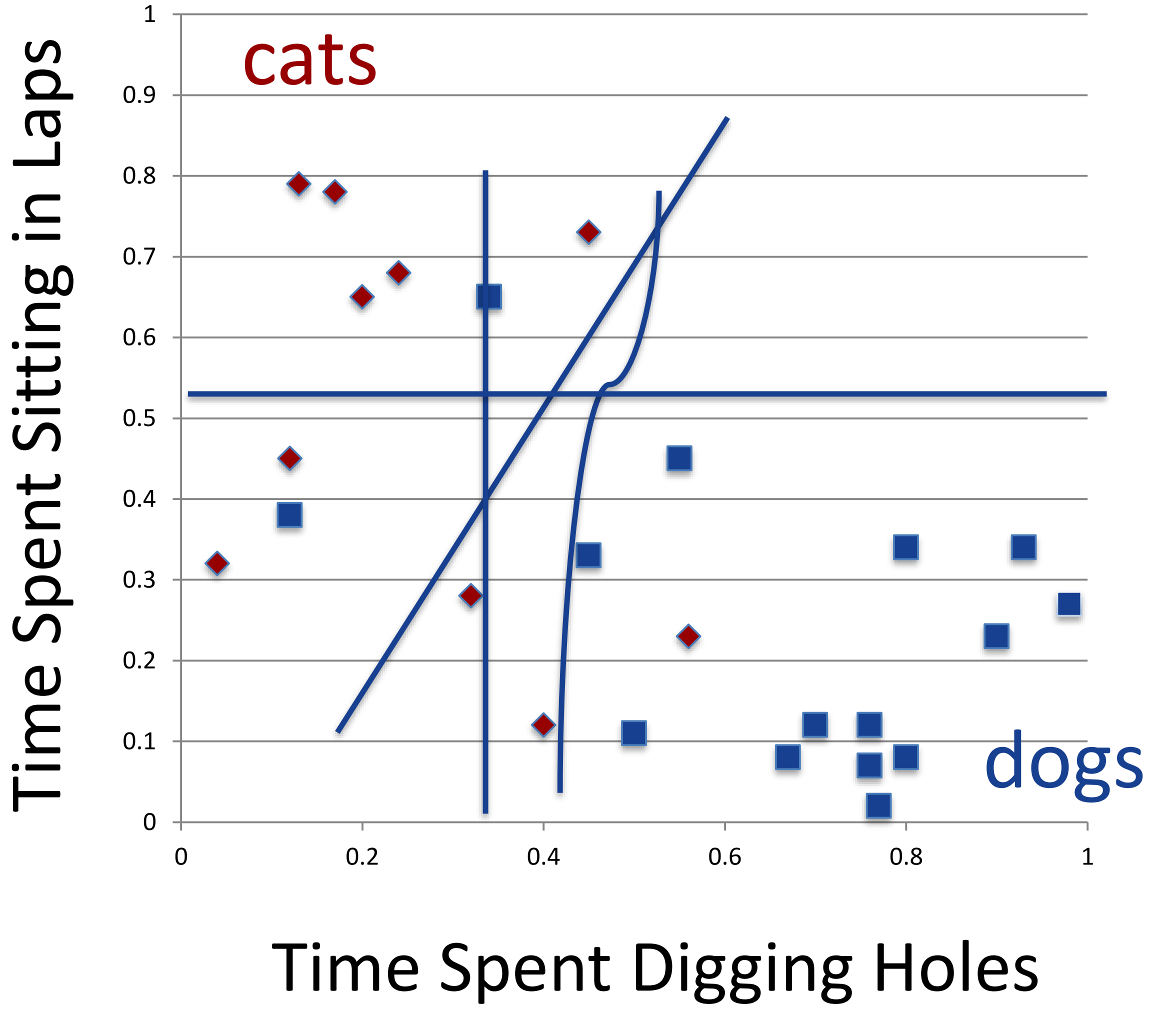
**Individualize Treatme**

# What is Machine Learning?

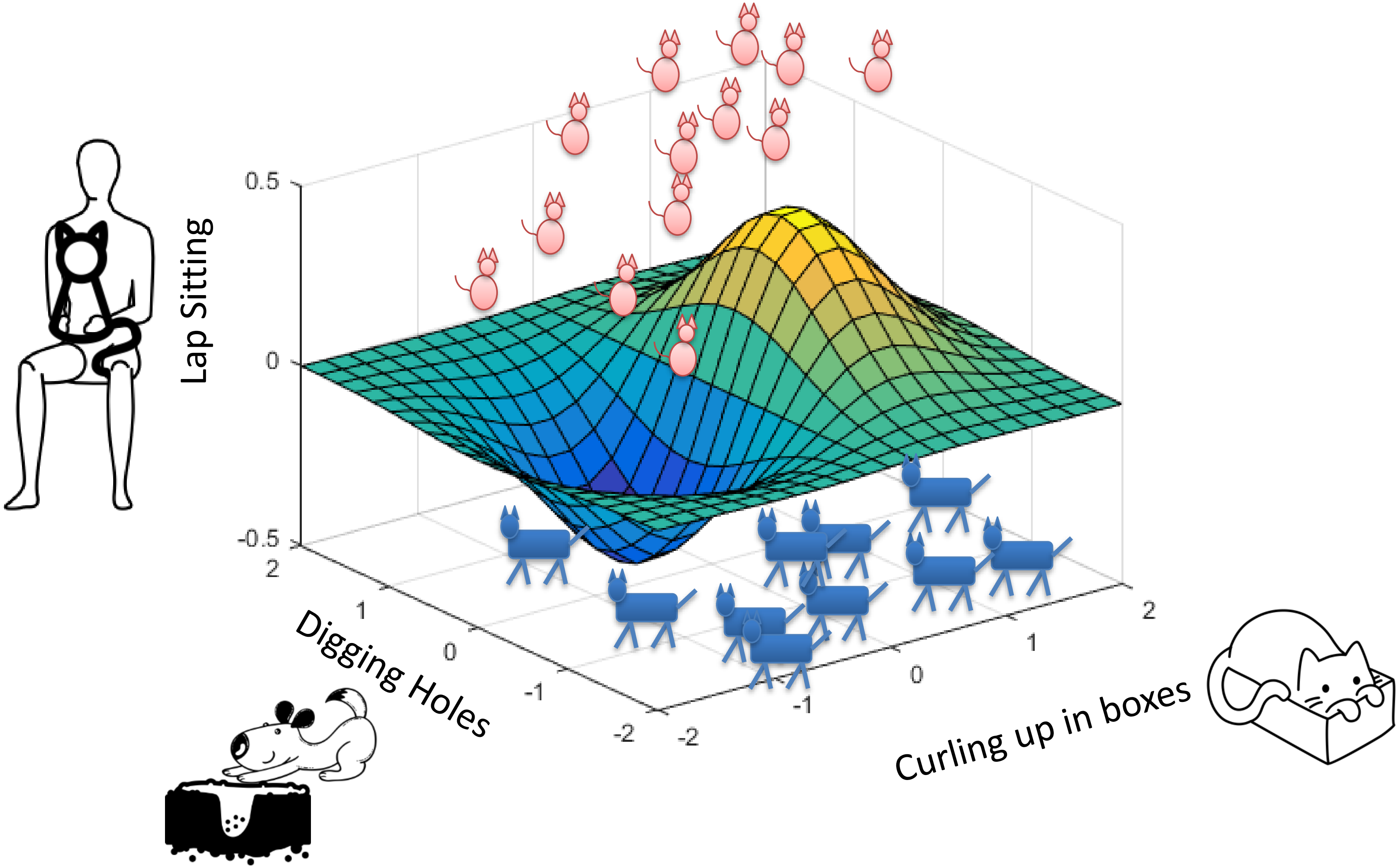
# Real World Applications

- Large digital datasets
- Known outcomes
- Identify patterns in data to predict the future
  - Clicking on an ad
  - Purchasing items
  - Signing up for a credit card
  - Switching cell phone providers

# Can you differentiate?
# Cat vs. Dog



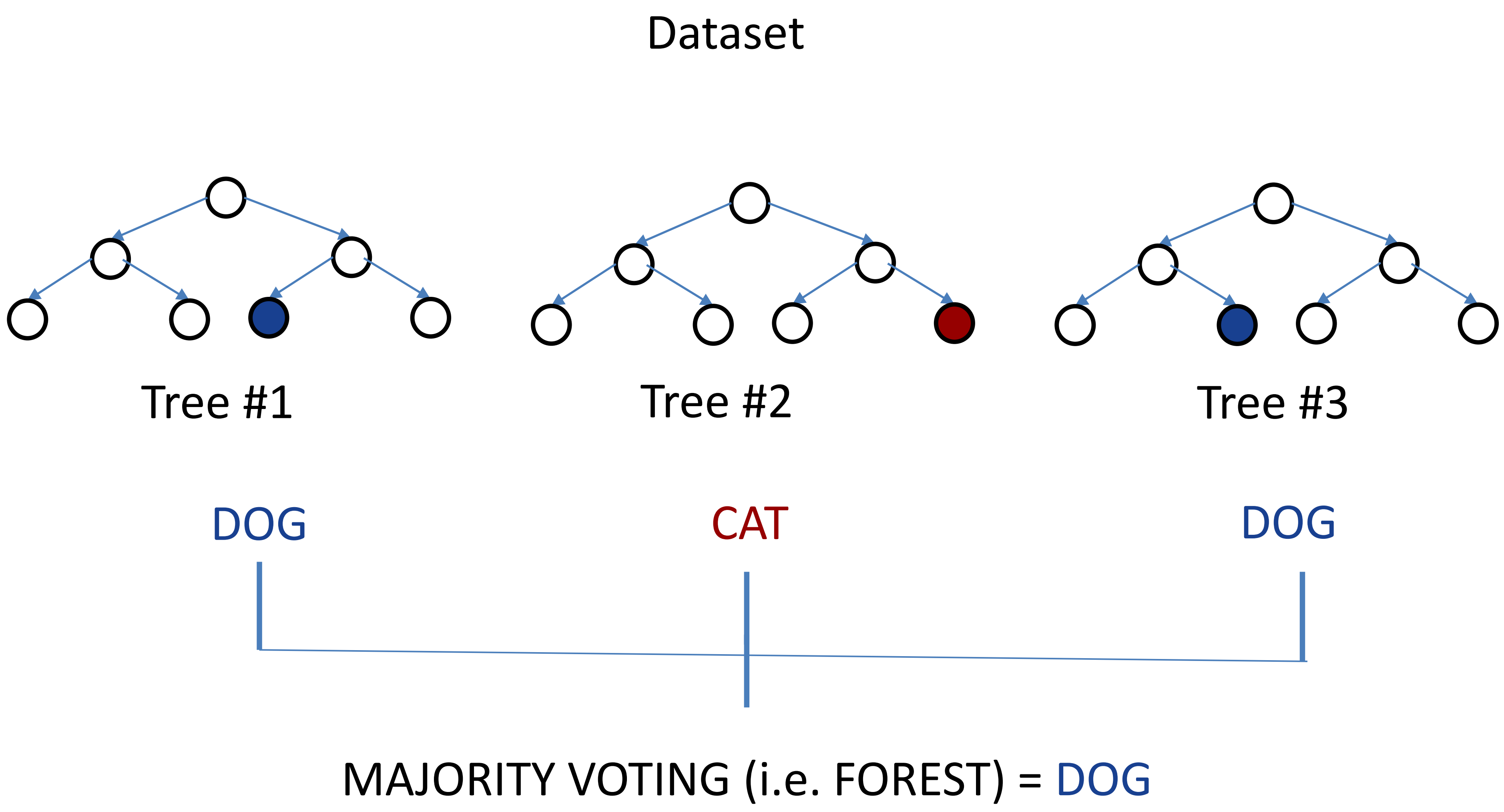

# 3 Variables = 3 Dimensions



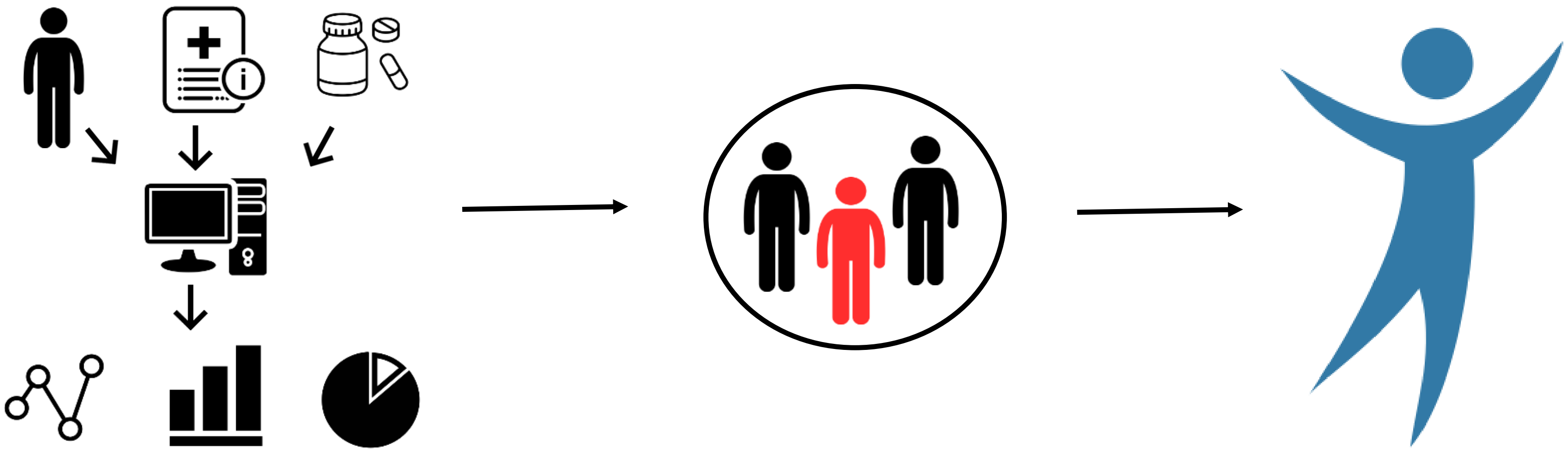
# What is Random Forest?

- A modern Machine Learning method
- Computer-based algorithm which uses decision trees to classify outcomes. e.g. Cat or Dog
- Can incorporate many variables and interactions
  - Identifies most important variables for prediction
  - “No pre-conceived notions”

# What is Random Forest?



# How can we improve care for an individual?



# Example:
# Liver-Related Clinical Outcomes

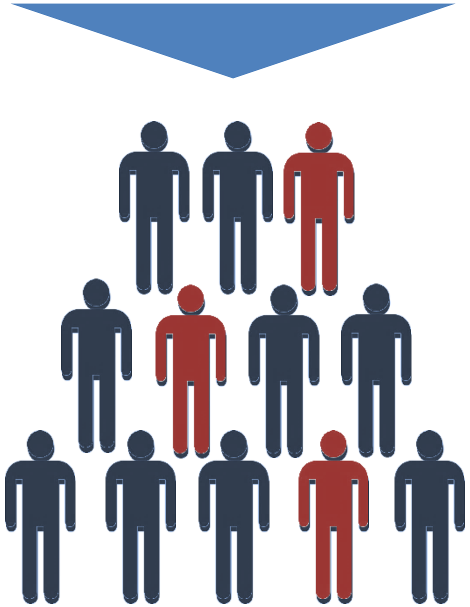
Evaluating liver-related clinical outcomes in Hepatitis C

Konerman, et al. PLOS One 2017

# Topical Research Questions of Interest

**1. Risk Stratification and Prognosis**

**2. Interventions to Decrease Risk of Adverse Outcomes**



Population of Interest

*modifiable*

**Risk Factors**

*Rigid*

**Adverse Clinical Outcome**

**No Clinical Outcome**

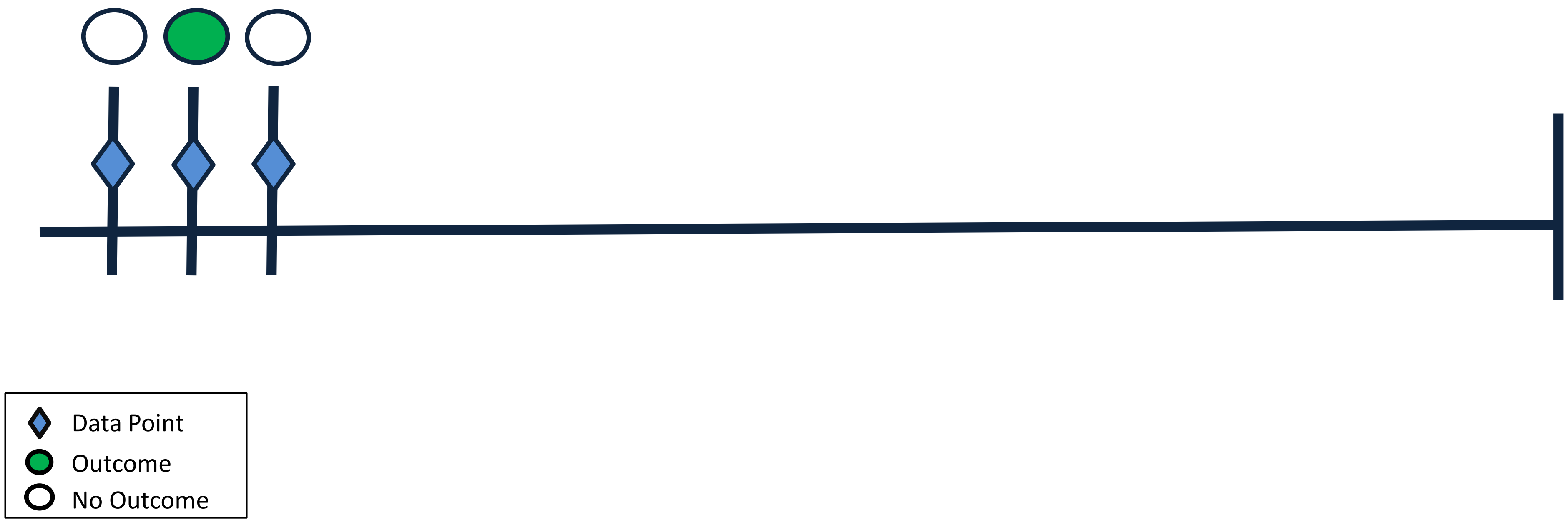
# Predictive Models for Risk of Clinical Outcomes

- **Methods:**
  - **Develop Longitudinal Models on HALT-C for a composite clinical outcome**
  - **Validate 1007 HCV patients at Michigan Medicine**
  - **Predict outcomes at 1 and 3 years**

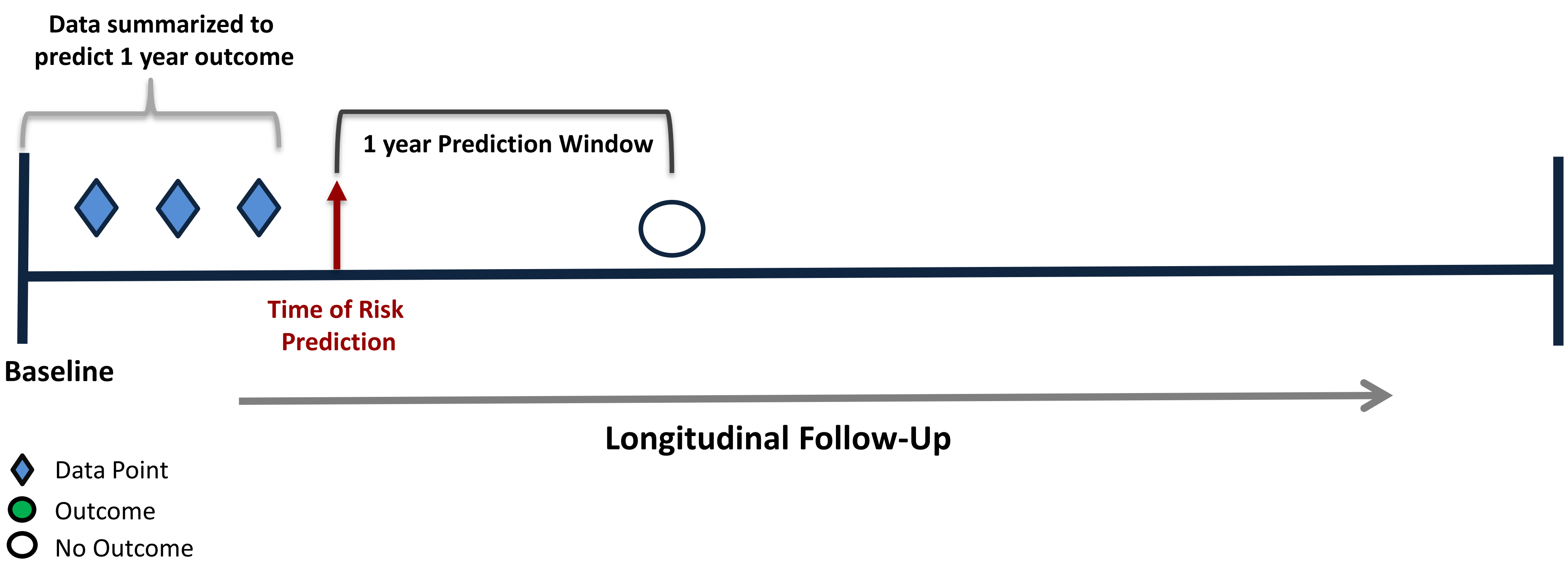
# Patient Demographics

| Variable | Summary statistics | |
|---|---|---|
| | **HALT-C Cohort** (N=1050) | **Michigan Medicine Cohort** (N=1007) |
| Age (median, IQR) | 49 (46-54) | 49.4 (44.3-54.3) |
| Male | 745 (71%) | 612 (61%) |
| Race (% White) | 752 (71.6%) | 636 (80.1%) |

# Approach

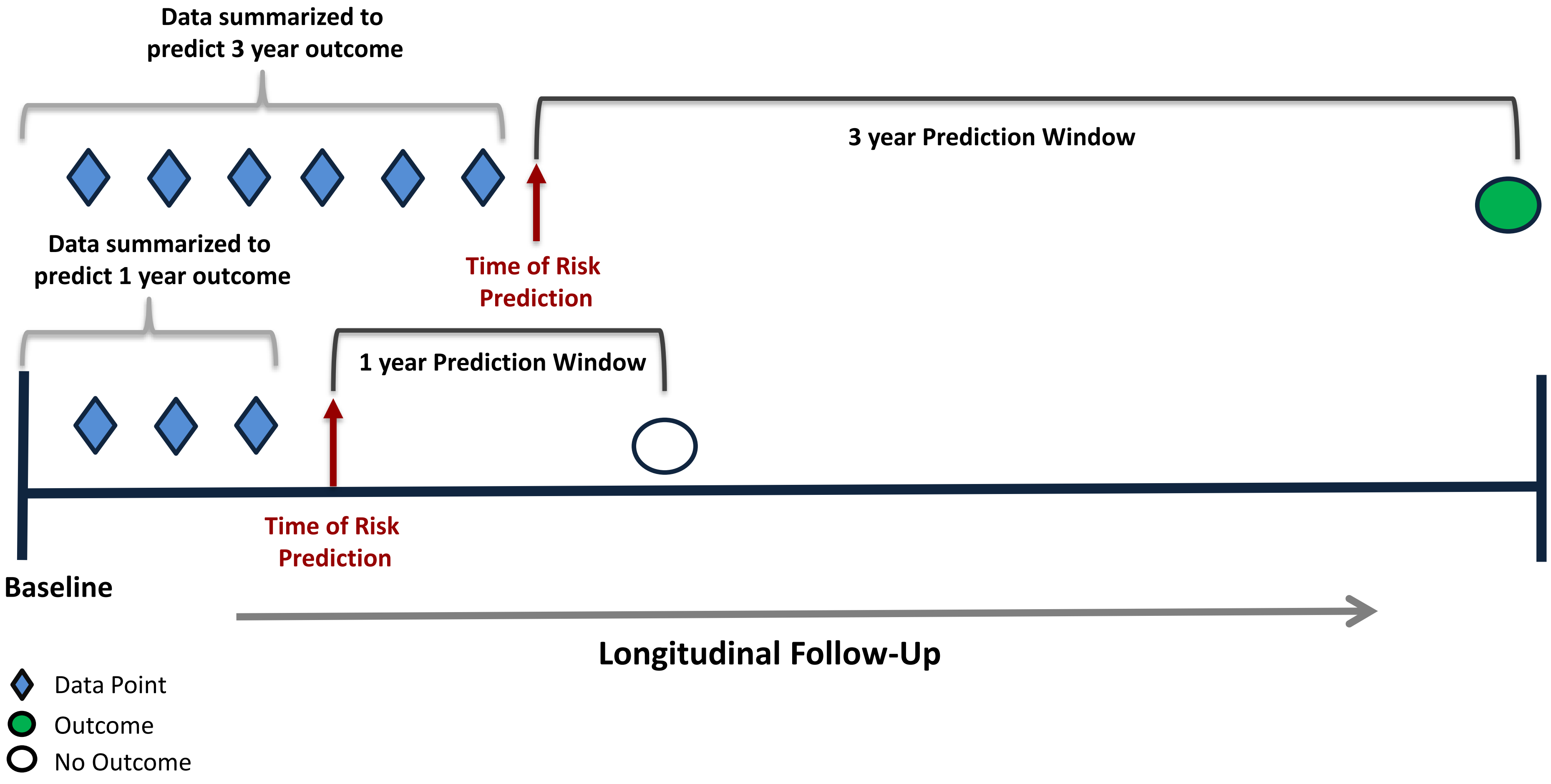


# Approach



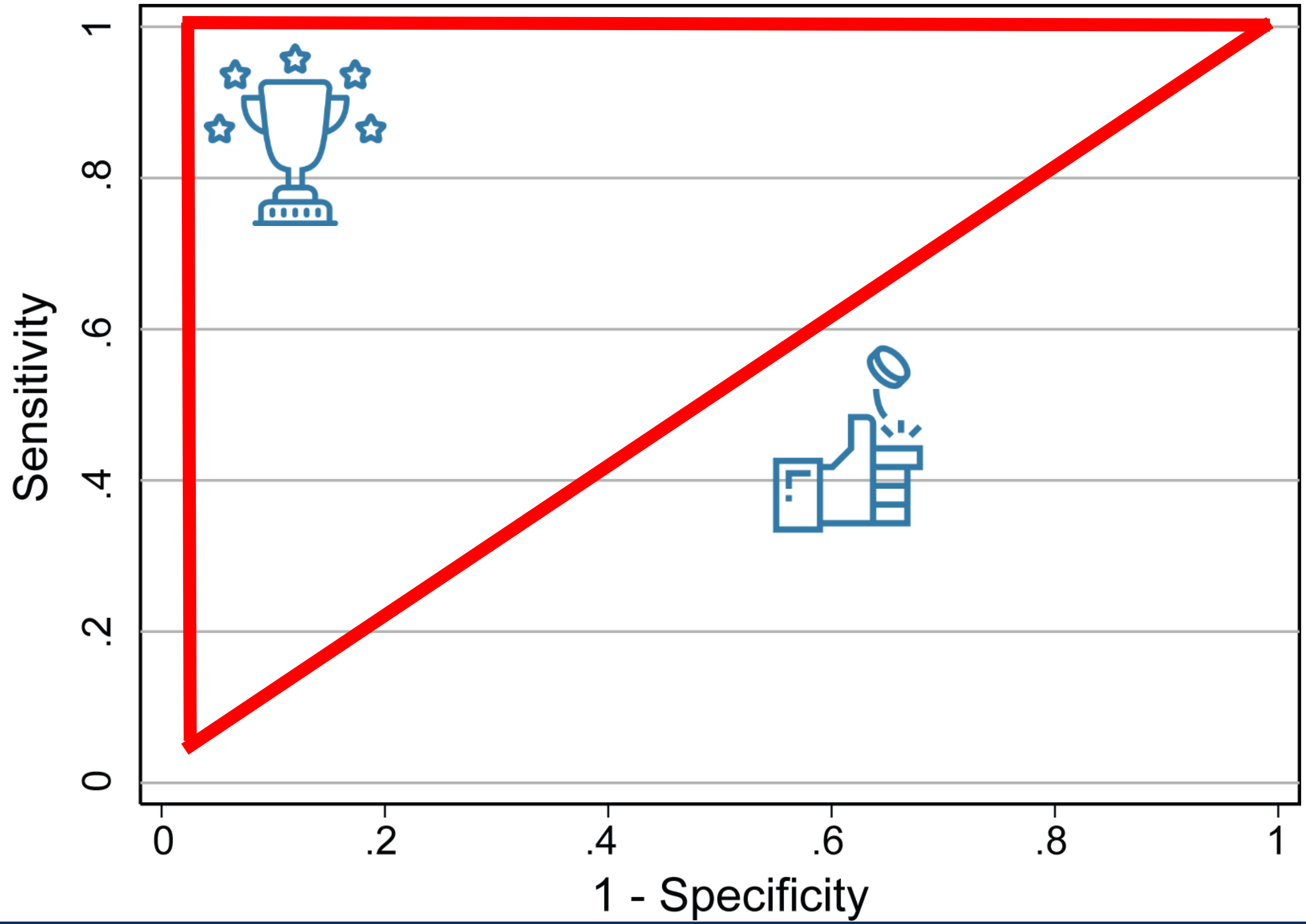
Data summarized to predict 1 year outcome
1 year Prediction Window
Time of Risk Prediction
Baseline
Longitudinal Follow-Up
Data Point
Outcome
No Outcome

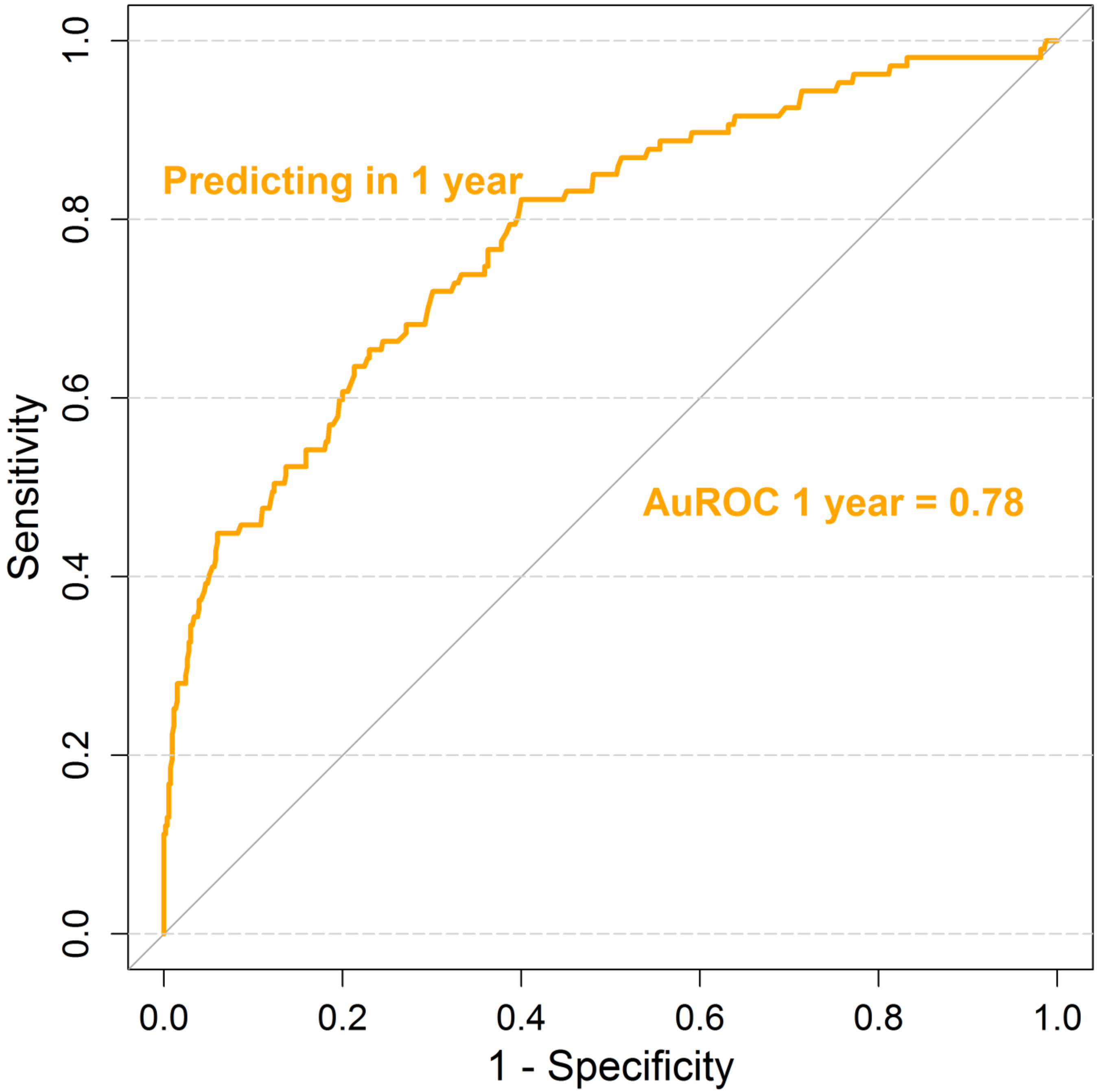
# Approach

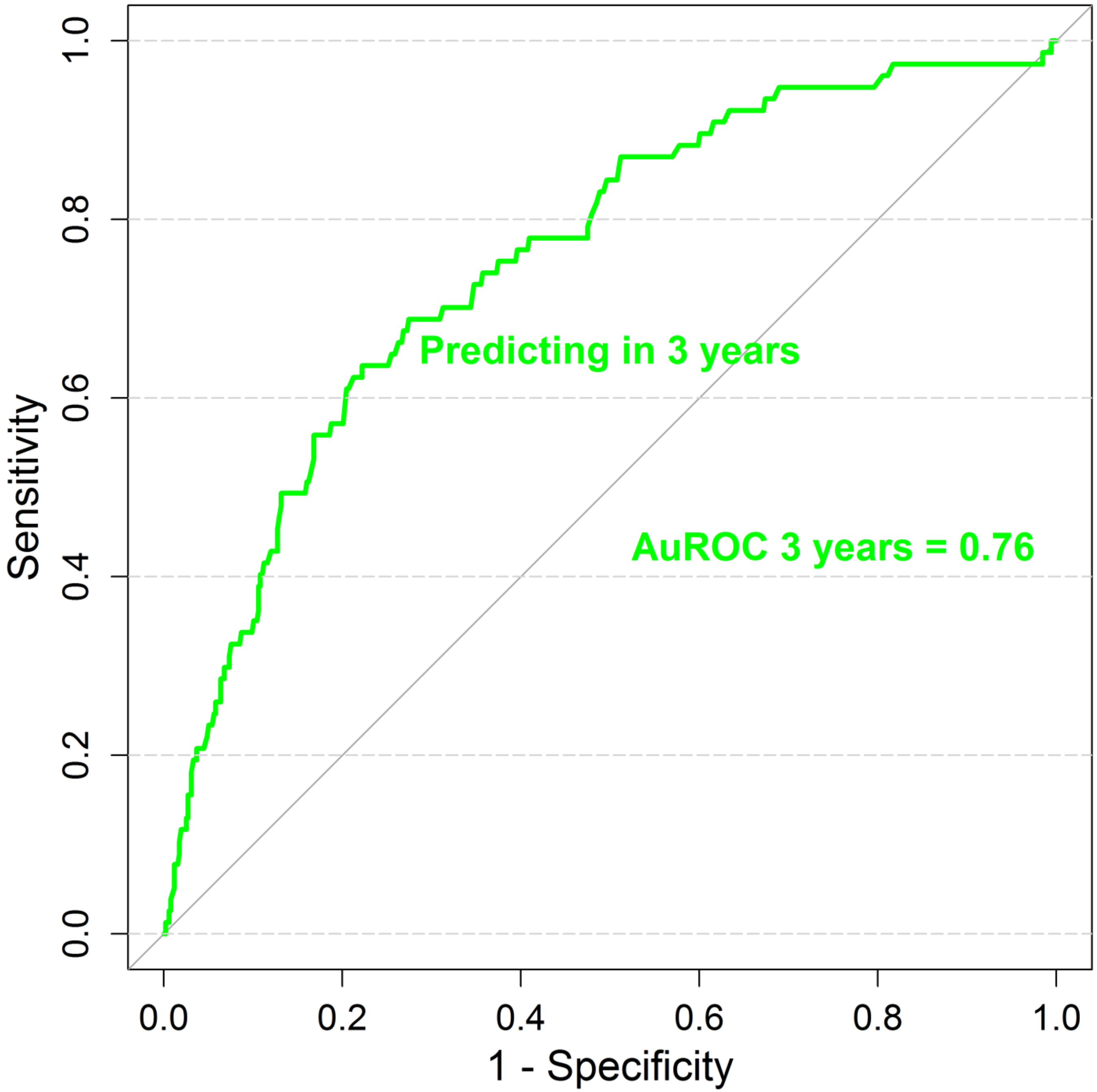


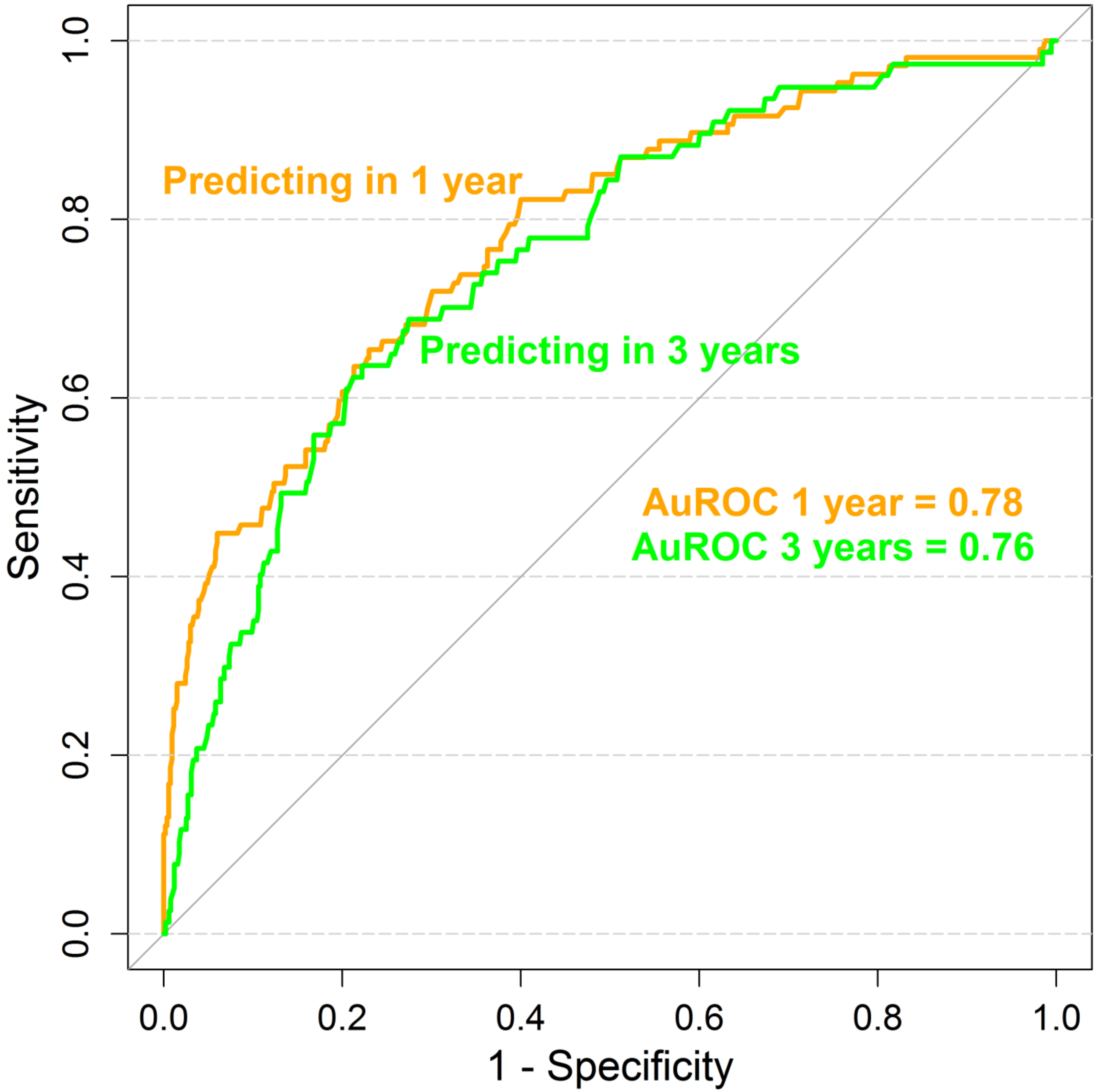
Data summarized to predict 3 year outcome
3 year Prediction Window
Data summarized to predict 1 year outcome
Time of Risk Prediction
1 year Prediction Window
Time of Risk Prediction
Baseline
Longitudinal Follow-Up
Data Point
Outcome
No Outcome

# Results

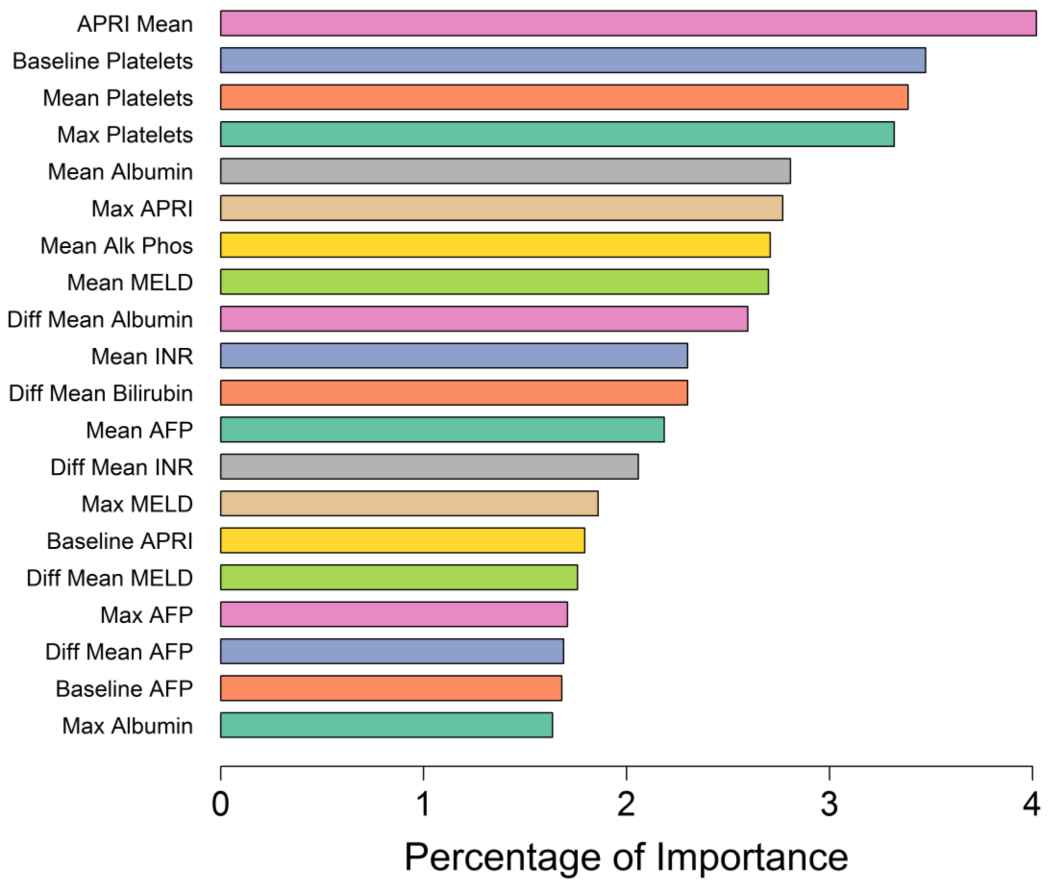



Predicting in 1 year
AuROC 1 year = 0.78
Sensitivity
1 - Specificity
0.0
0.2
0.4
0.6
0.8
1.0


Predicting in 3 years
AuROC 3 years = 0.76
Sensitivity
1 - Specificity
0.0
0.2
0.4
0.6
0.8
1.0


Predicting in 1 year
Predicting in 3 years
AuROC 1 year = 0.78
AuROC 3 years = 0.76
Sensitivity
1 - Specificity
0.0
0.2
0.4
0.6
0.8
1.0

# Variable Importance



# Example:
# Progression of Disease

Evaluating progression to Cirrhosis
in Veterans

Konerman, et al. PLOS One 2019

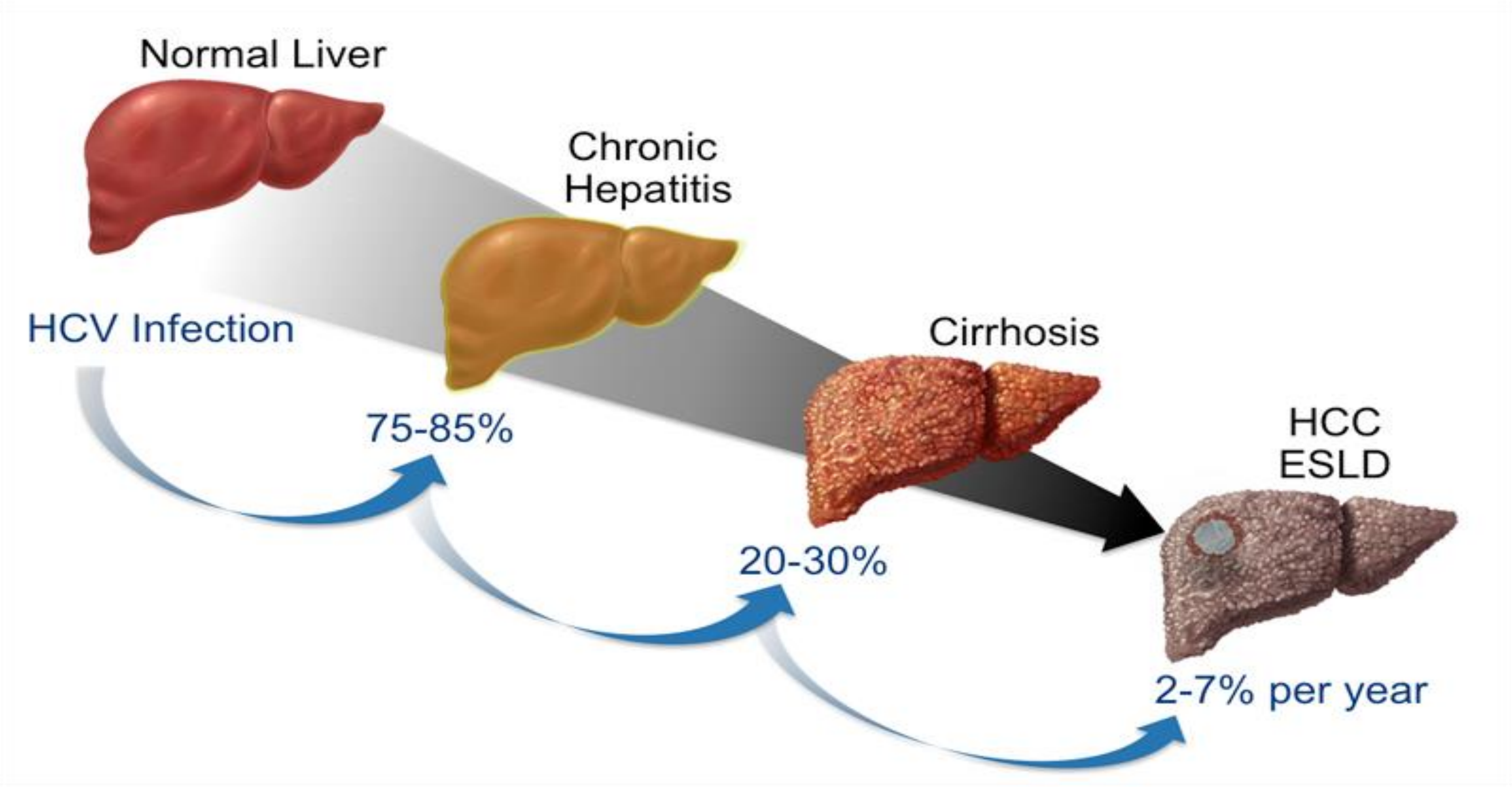
# HCV Models in VA Data

**250,000 Veterans**
in 2016

**3** Million

**170**
Million

# Progression



# Cohort

280,475 VHA veterans with a positive HCV RNA test

↓

232,668 with at least 2 APRI scores

↓

214,274 no history of HCC and cirrhosis

↓

203,188 without elevated baseline APRI

↓

156,400 with at least 1 year follow-up

↓

86,340 without any treatment

↓

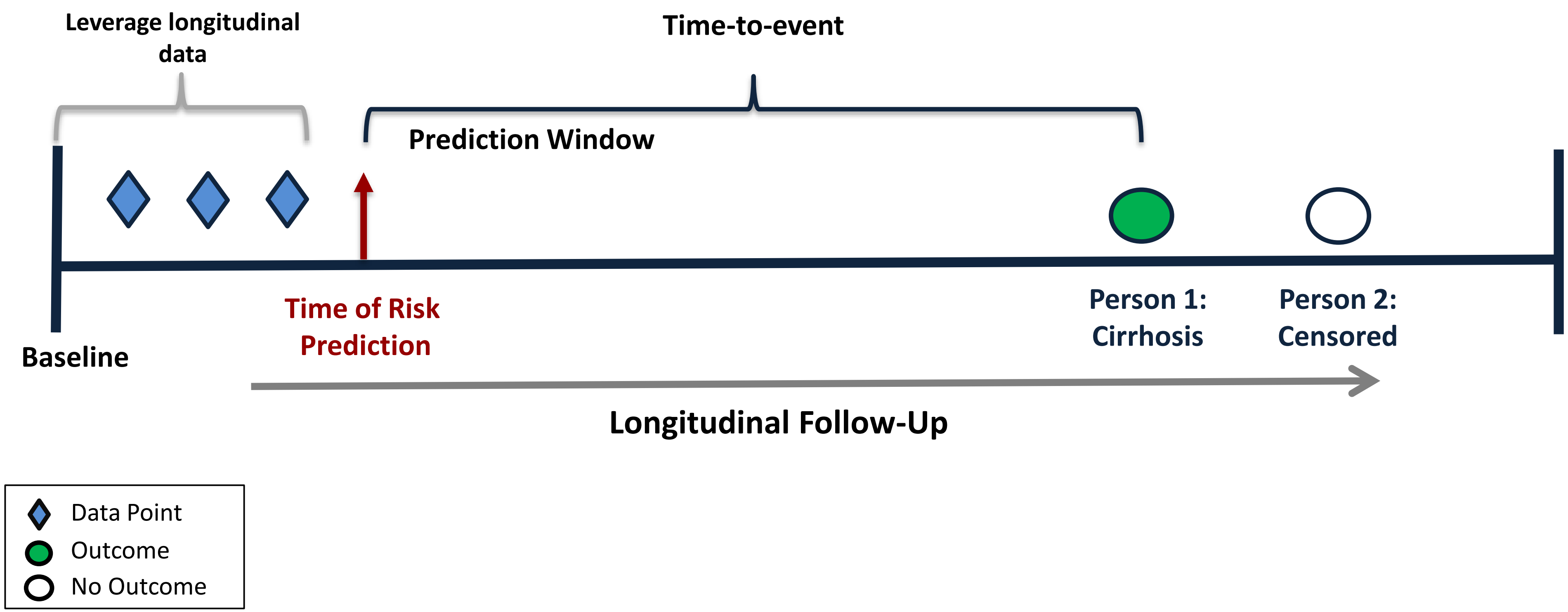
72,683 with valid lab data within 2 years of enrollment

# Patient Demographics

- Large heterogeneous population, primarily male

| Variable | Summary statistics |
|---|---|
| Age (mean, sd) | 52.84 (8.74) |
| Male | 70,377 (96.8%) |
| Race (% White) | 35,216 (52.9%) |

# Approach



Leverage longitudinal data
Time-to-event
Prediction Window
Baseline
Time of Risk Prediction
Person 1: Cirrhosis
Person 2: Censored
Longitudinal Follow-Up
Data Point
Outcome
No Outcome

# Results

| Time | N | % Events | Model | AuROC | SN | SP | PPV | NPV |
|---|---|---|---|---|---|---|---|---|
| 1 year | 18896 | 0.036 | CS Cox | 0.807 | 0.79 | 0.71 | 0.11 | 0.99 |
| | | | CS Boosting | 0.817 | 0.77 | 0.73 | 0.11 | 0.99 |
| | | | LGT Cox | 0.828 | 0.75 | 0.76 | 0.10 | 0.99 |
| | | | LGT Boosting | 0.838 | 0.76 | 0.77 | 0.11 | 0.99 |
| 3 years | 14605 | 0.112 | CS Cox | 0.784 | 0.73 | 0.72 | 0.25 | 0.95 |
| | | | CS Boosting | 0.799 | 0.76 | 0.71 | 0.27 | 0.95 |
| | | | LGT Cox | 0.804 | 0.75 | 0.74 | 0.27 | 0.96 |
| | | | LGT Boosting | 0.815 | 0.76 | 0.73 | 0.28 | 0.96 |
| 5 years | 11334 | 0.206 | CS Cox | 0.775 | 0.74 | 0.70 | 0.41 | 0.90 |
| | | | CS Boosting | 0.790 | 0.75 | 0.70 | 0.42 | 0.91 |
| | | | LGT Cox | 0.794 | 0.75 | 0.71 | 0.42 | 0.91 |
| | | | LGT Boosting | 0.805 | 0.73 | 0.74 | 0.41 | 0.92 |

# HCV Treatment Approaches

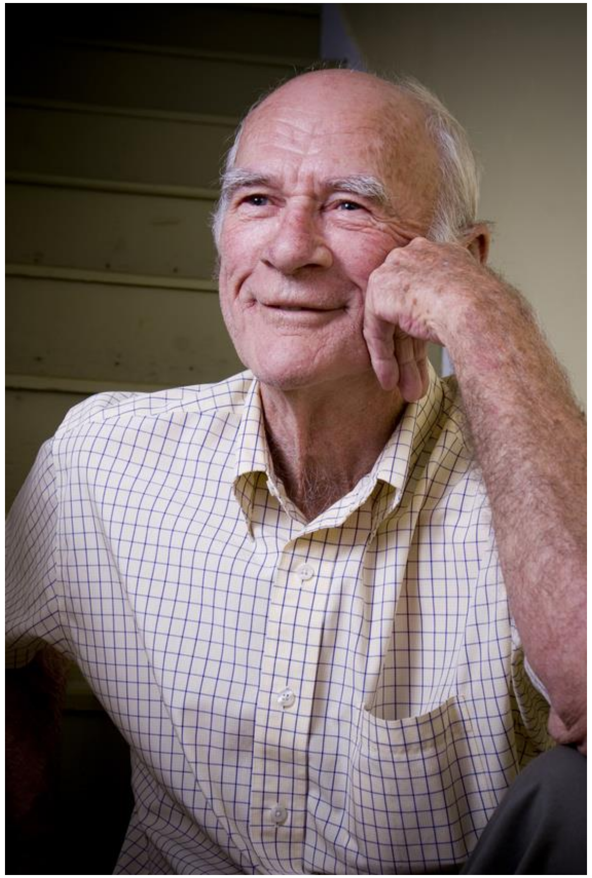
**Population Health Approach**

- Population Level
- Untargeted
- Only need high level patient data
- Quickly scalable (available data infrastructure?)
- Economical in resource limited setting?

**Precision Health Approach**

- Patient Level
- Targeted
- Need granular/sparse patient data
- Scalable but need to build data infrastructure
- Optimizes targeted health and value for patients and payers

# Value of Precision Health



**Targeted Treatment**

# Value of Precision Health

Potential alternative treatment approaches in HCV Medicaid patients:

- **Limited Resources**
  - » Navigation for those not seeking care
  - » IVDU
- **High-Risk HCV Targets**
  - » Non-adherence, more intensive monitoring
  - » Reinfection
- **Post Treatment Monitoring**
  - » Improve treatment transition

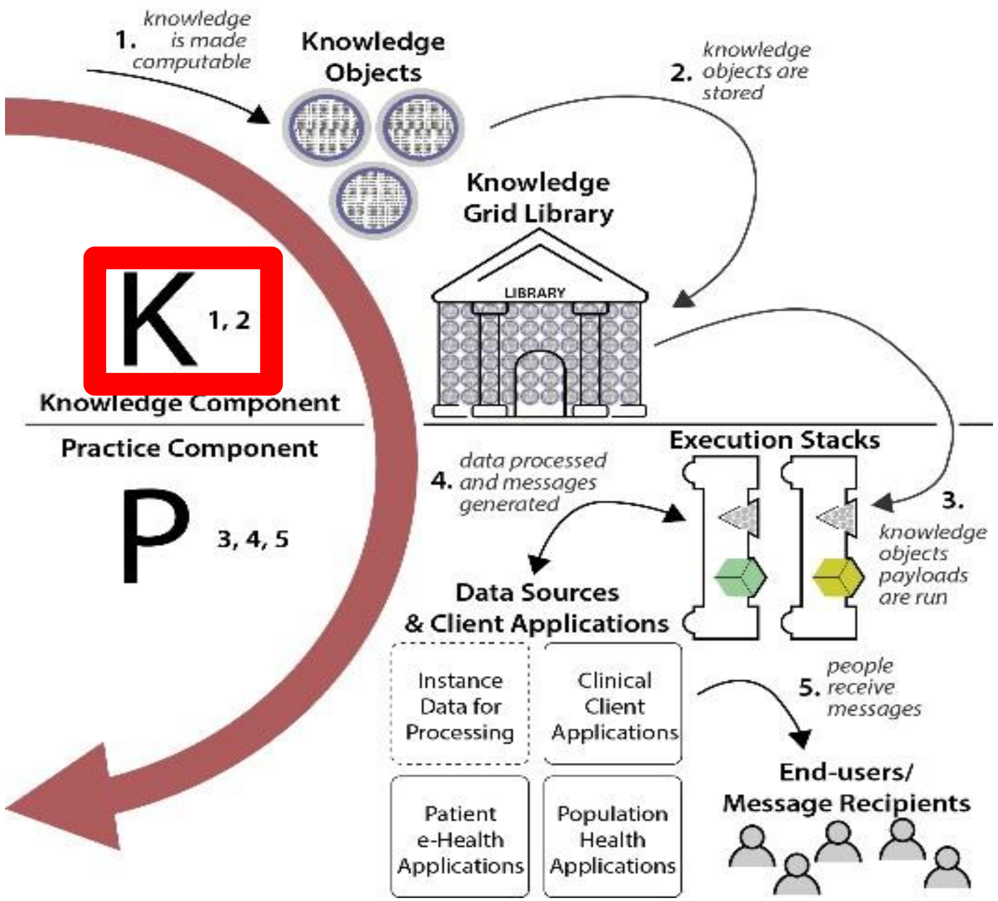
# MDHHS and University of Michigan

The Collaboration

David Neff, DO
Chief Medical Director,
MDHHS

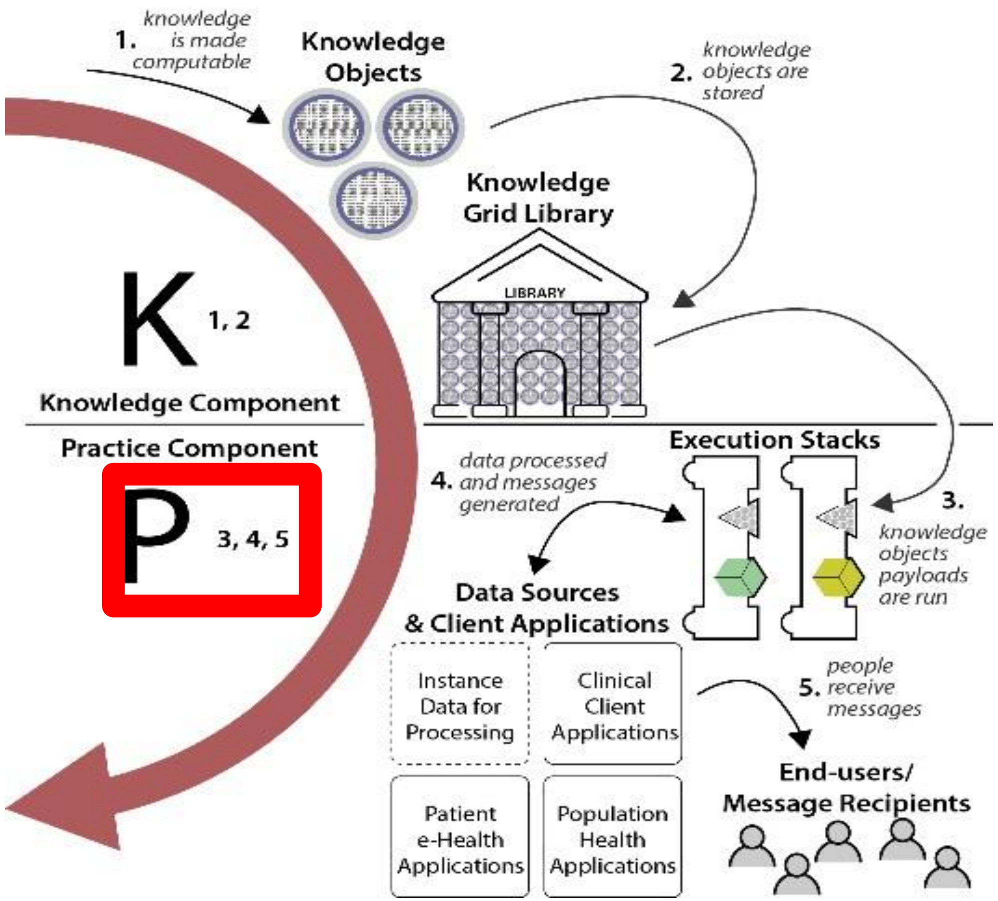
# K-TOP

## K-GRID is the “K” Component



# K-TOP

## MTOP is the "P" Component



MDHHS

# K-TOP

UNIVERSITY OF MICHIGAN

## Proposed Focus Areas to Develop Use Cases

1. Opioids
2. Hepatitis C
3. Rare Disease Registries (ie, spinal muscular atrophy, aka SMA)
4. Social Determinants and Adverse Childhood Experiences (ACE’s)
5. Superutilizors
6. Diabetes
7. Heart Disease

# Thoughts for the group:

- What barriers do you foresee with these approaches?
  - Data Access
  - How to deliver care to those not seeking care
  - Getting Prediction models into practice
  - What do patients think of treatment policies
- What alternatives should we be considering?

# THANK YOU

MICHIGAN MEDICINE
UNIVERSITY OF MICHIGAN

VA CCMR