Session 74, Using Predictive Analytics to Prescribe Care

SOA Antitrust Disclaimer
SOA Presentation Disclaimer

SOA Health Meeting 2019

Using Predictive Analytics to Prescribe Care

Thinking beyond the traditional risk score

Hans Leida, PhD, FSA, MAAA
Principal and Consulting Actuary, Milliman Inc.

Joe Long
Assistant Actuary and Data Scientist, Milliman Inc.

Joe Dorocak, ASA, MAAA
Manager, Population Health Analytics, Cleveland Clinic

Don McLellan, ASA, MAAA
Director Risk Analytics, Cleveland Clinic

JUNE 25, 2019

Milliman

Cleveland Clinic

Limitations

- This presentation is intended for informational purposes only. It reflects the opinions of the presenter, and does not represent any formal views held by Milliman, Inc. Milliman makes no representations or warranties regarding the contents of this presentation. Milliman does not intend to benefit or create a legal duty to any recipient of this presentation.

Agenda

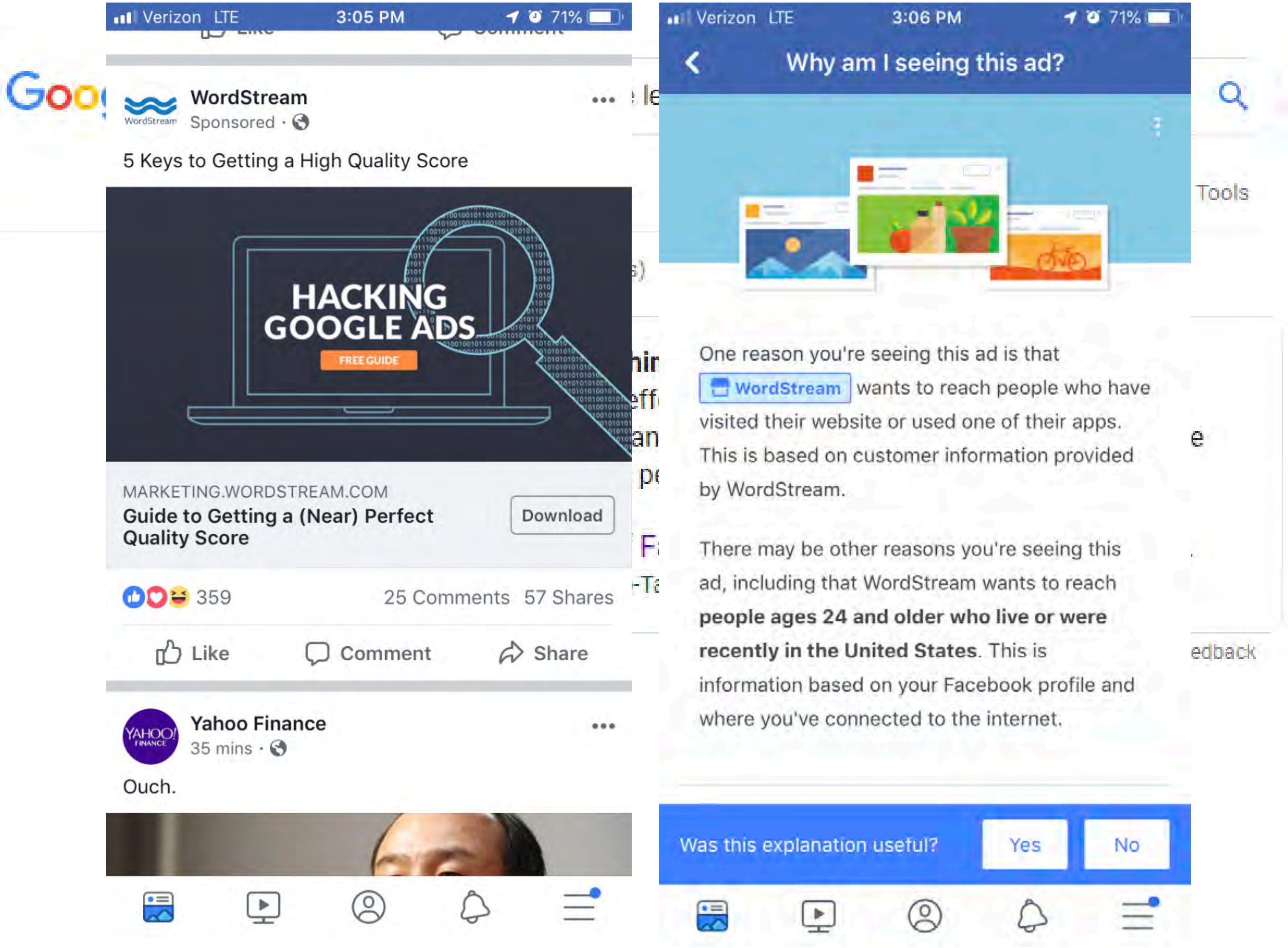
- Personalized marketing
- What does this look like for health care?
- Cleveland Clinic case studies
  - Primary Care Risk Model
  - Recommendation Model
- Discussion on clinical validation of models
- Now and the near future

Traditional marketing

Find the right place to advertise

Personalized marketing

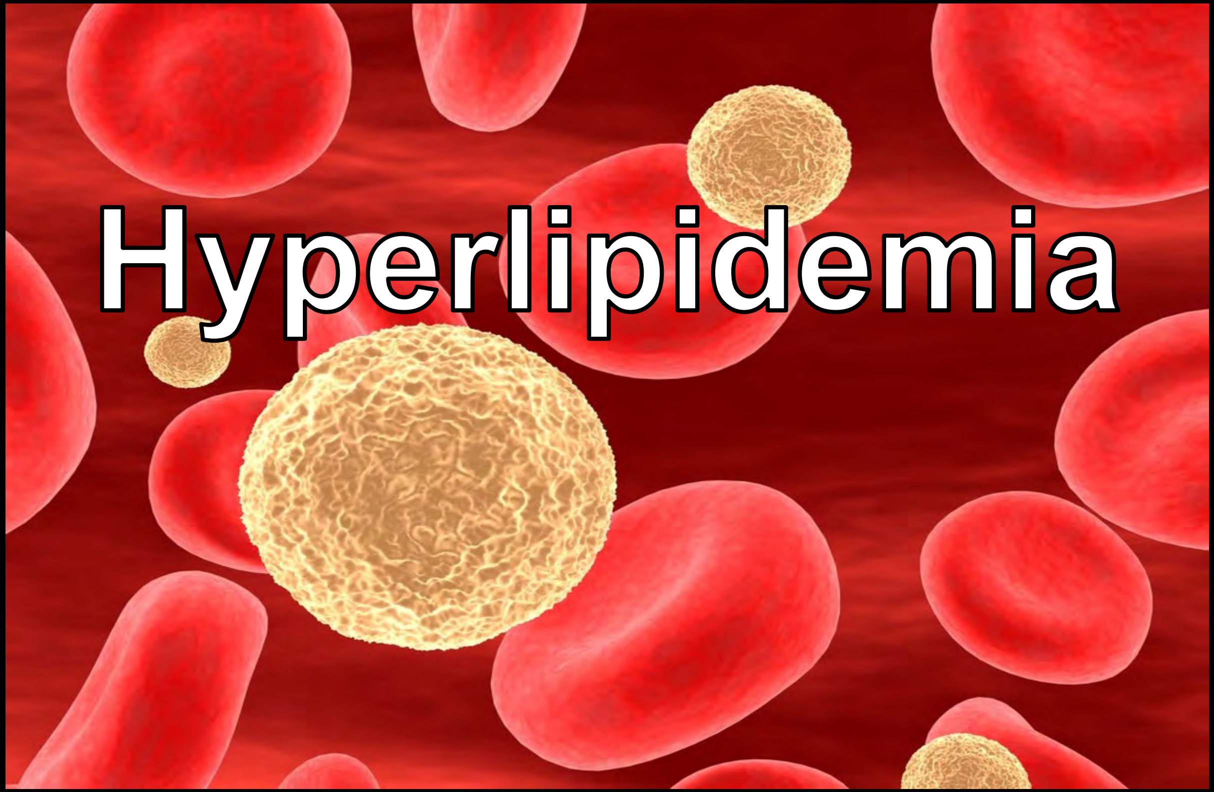
Show the right ad, to the right person, at the right time



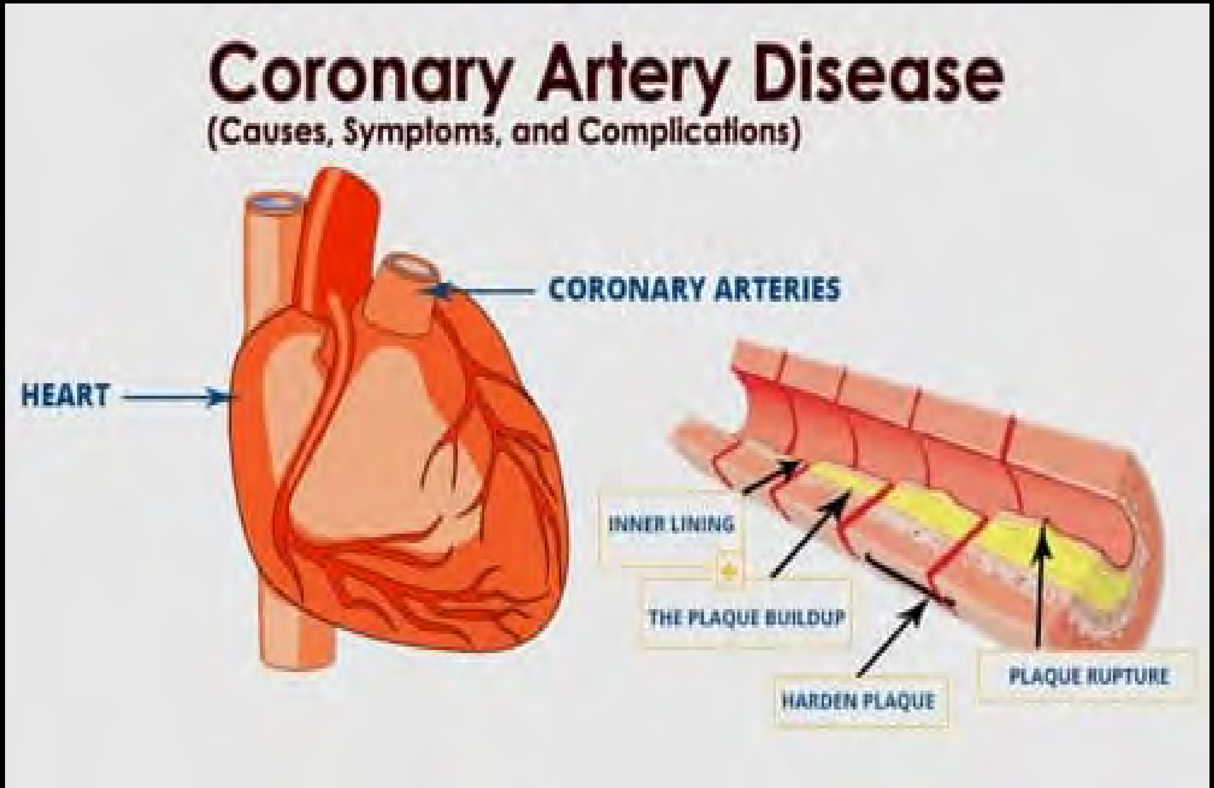
What does this look like for health care?

DISEASEFLIX Browse Kids DVD Search

Because You Have…



You'll Love…

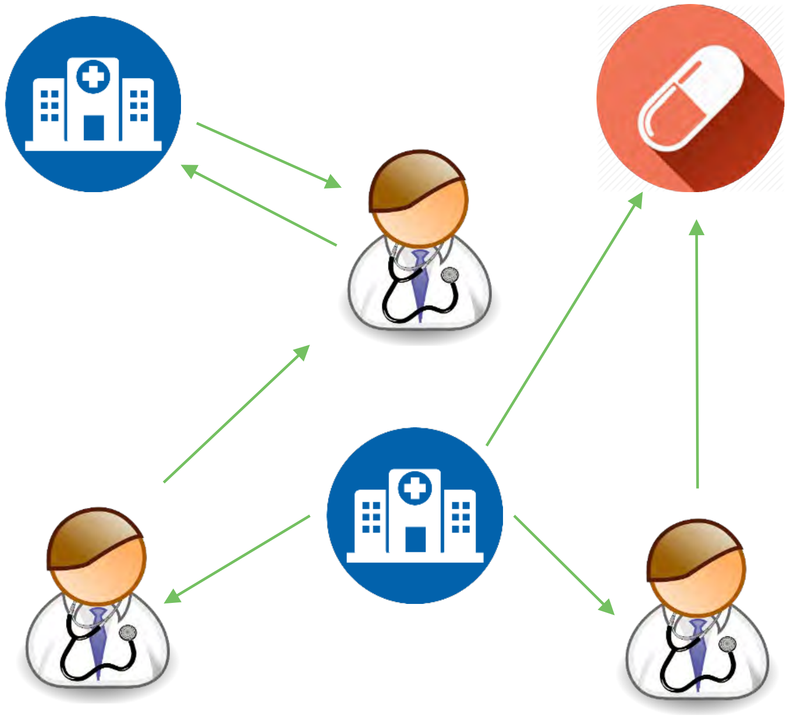


“One new development over the past decade has been the introduction of predictive models that aim to predict more than simple relative risk. For example, some models now produce probabilities of hospitalization as an additional dependent variable.”

—Geof Hileman and Spenser Steele, Accuracy of Claims-Based Risk Scoring Models, 2016

Predict other dependent variables

- Probabilities of events:
  - Hospitalization,
  - Death,
  - Opioid addiction,
  - Readmission,
  - Etc.
- Staffing needs/no shows
- Care management paths



Non-traditional data sources

Social determinants of health (SDoH)

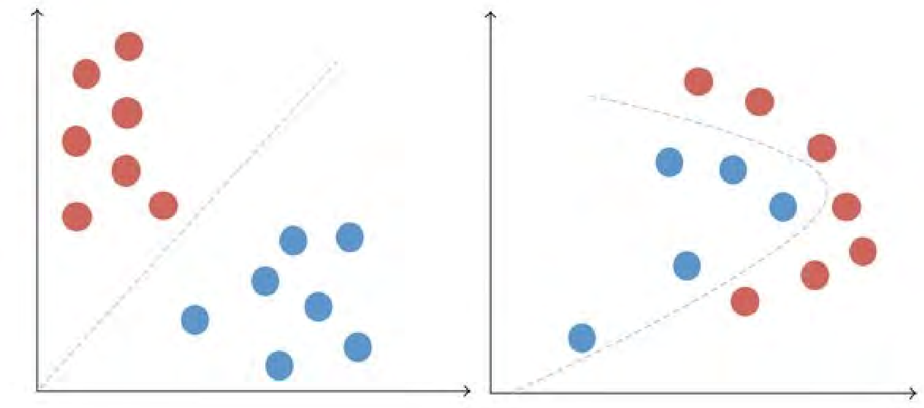
Consumer data

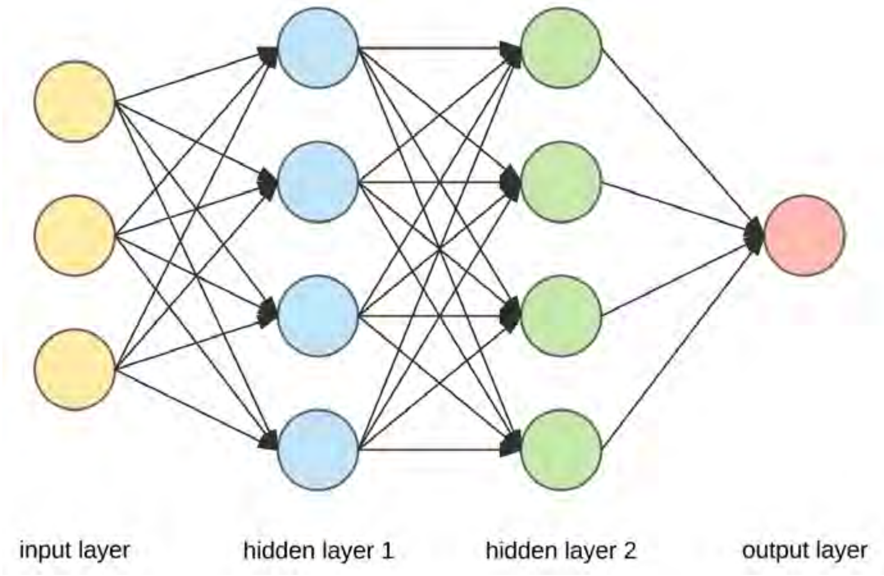
Electronic Health Records (EHR)

Milliman

Moving beyond traditional linear models

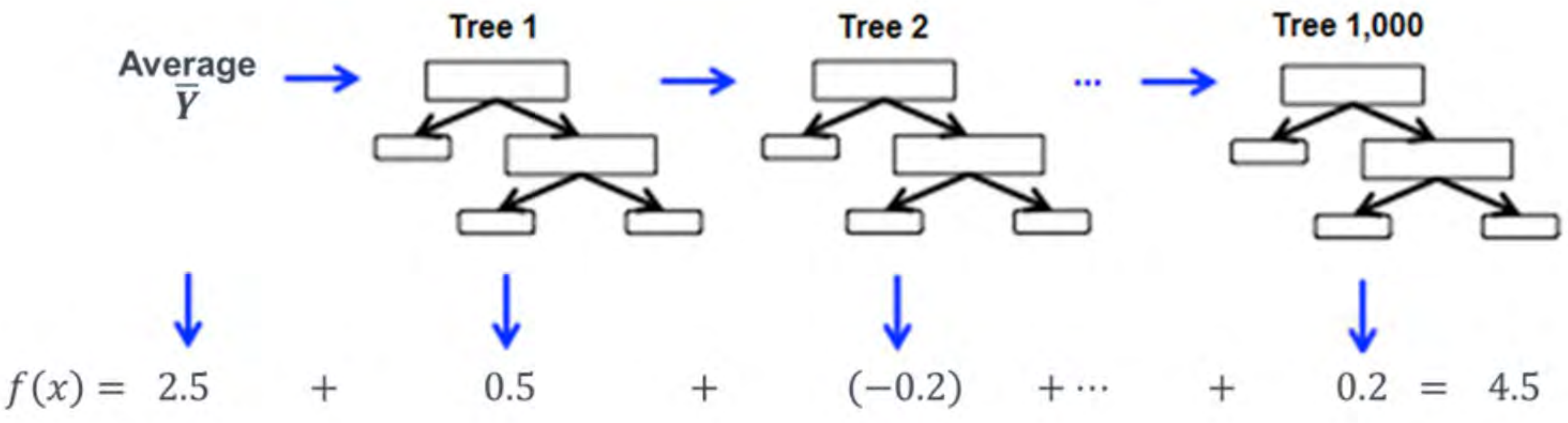
- Methods we'll discuss today
  - Gradient boosting machines (GBM)
  - *k*-nearest neighbors (KNN)
- Other methods to mention
  - Random Forest
  - Neural networks/deep learning
  - Bayesian Networks
  - Penalized GLMs
  - Generalized additive models (GAMs)





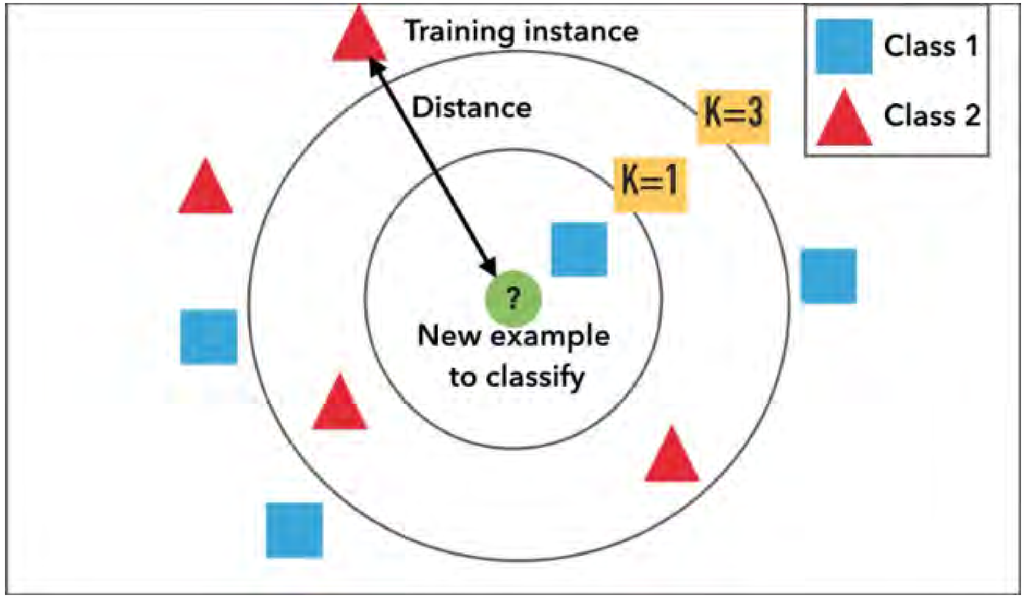
Gradient Boosting Machines

- Boosting
  - Build 100's of decision trees sequentially
  - Final prediction is the sum across all the trees
  - Too many trees can lead to overfitting the training data
- Can be used for regression and classification problems



k-nearest neighbors

- Non-parametric instance-based learning
  - No pre-computed model
  - Predictions are based on the *k* nearest observations in the training data
    - Classification: most frequent class among the neighbors (majority voting)
    - Regression: average of the neighbors
  - The choice of *k* is important
    - Small *k*: low bias and high variance (tends to overfit)
    - Large *k*: high bias and low variance (tends to underfit)



“Risk scores could be helpful if they help clinicians begin candid conversations about the unique circumstances that could make a patient more vulnerable to opioid use disorder, said Yngvild Olsen, a board member at the American Society of Addiction Medicine.”

—Politico

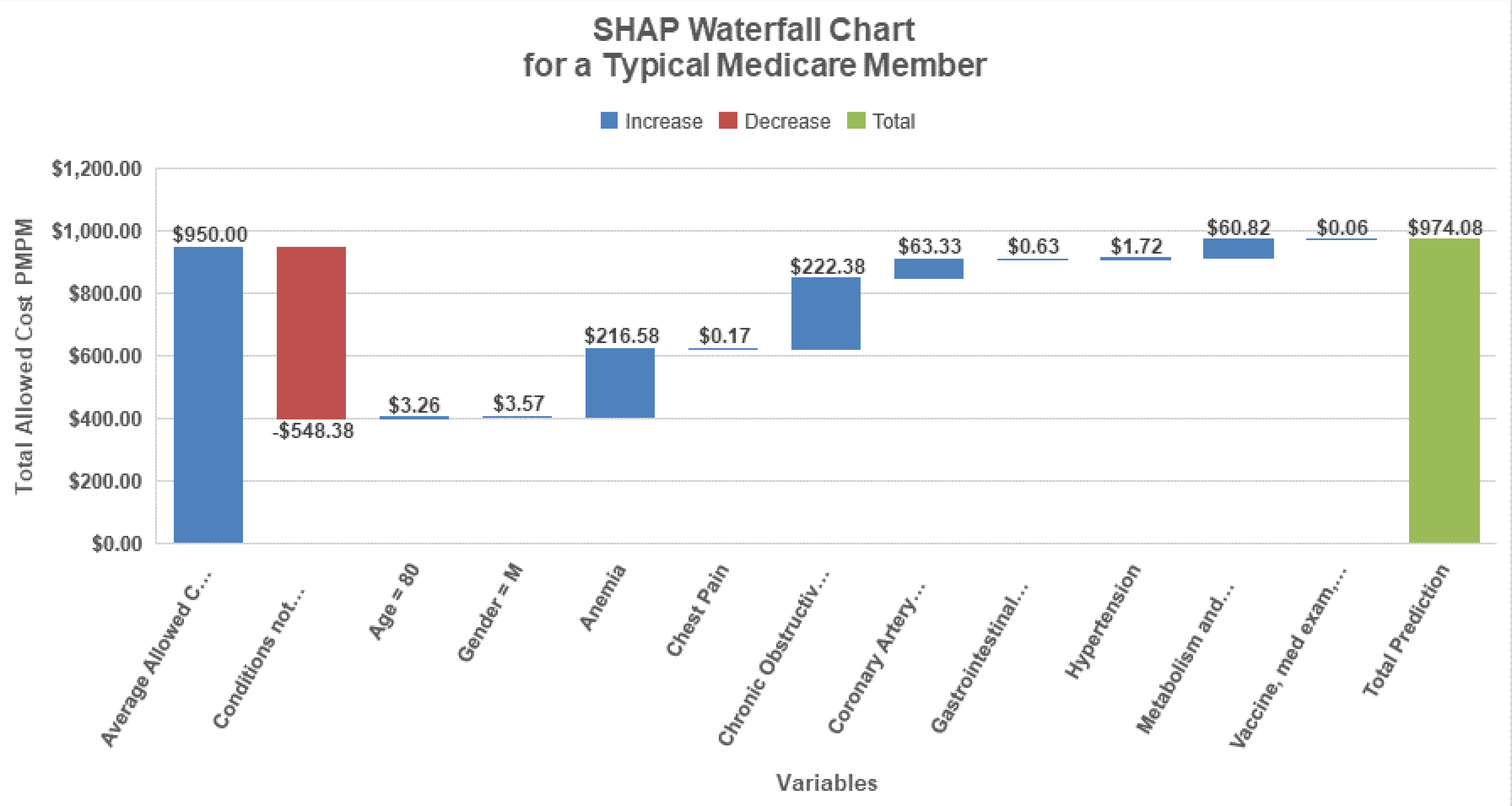
What’s driving the predictions

- A couple (of the many) methods for opening up the machine learning black box
  - Partial dependence plots (PDP)
  - SHapley Additive exPlanations (SHAP)

Medicare GBM SHAP waterfall example

FOR ILLUSTRATIVE PURPOSES ONLY

Typical Medicare Member	
Variables	**PMPM**
Average Allowed Cost PMPM	$950.00
Conditions not flagged	-$548.38
Age = 80	$3.26
Gender = M	$3.57
Anemia	$216.58
Chest Pain	$0.17
Chronic Obstructive Pulmonary Disease	$222.38
Coronary Artery Disease	$63.33
Gastrointestinal Hemorrhage	$0.63
Hypertension	$1.72
Metabolism and Nutrition Problems	$60.82
Vaccine, med exam, other preventive	$0.06
Total Prediction	**$974.08**



Case Studies

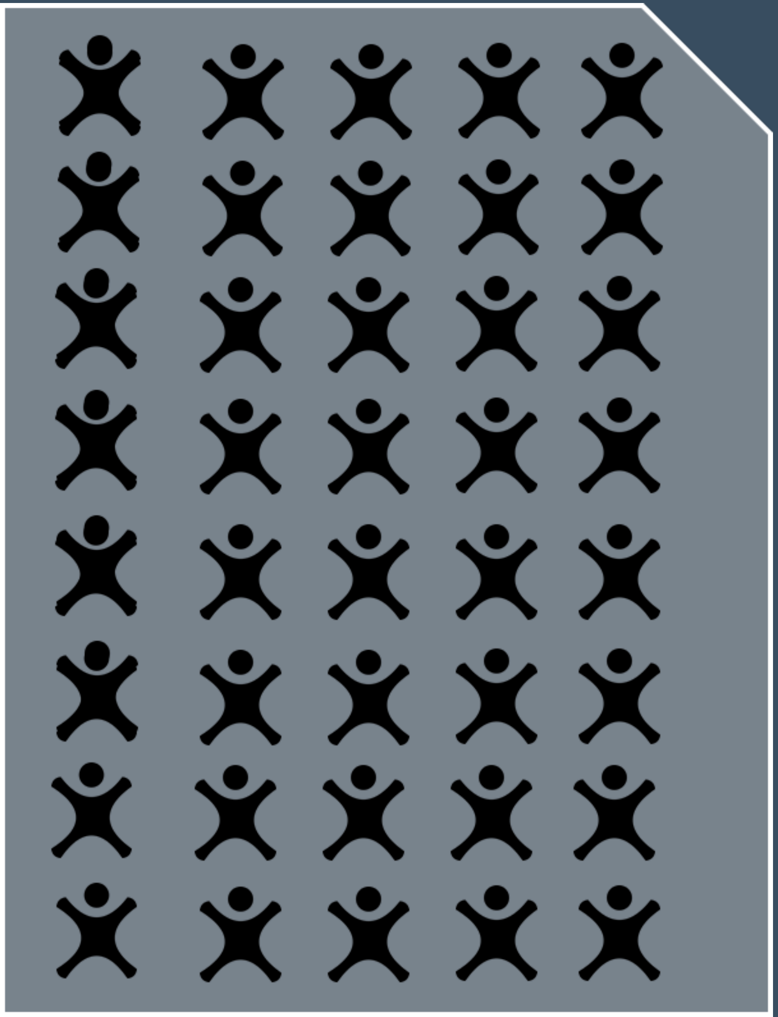
Cleveland Clinic

Care Coordination at Cleveland Clinic

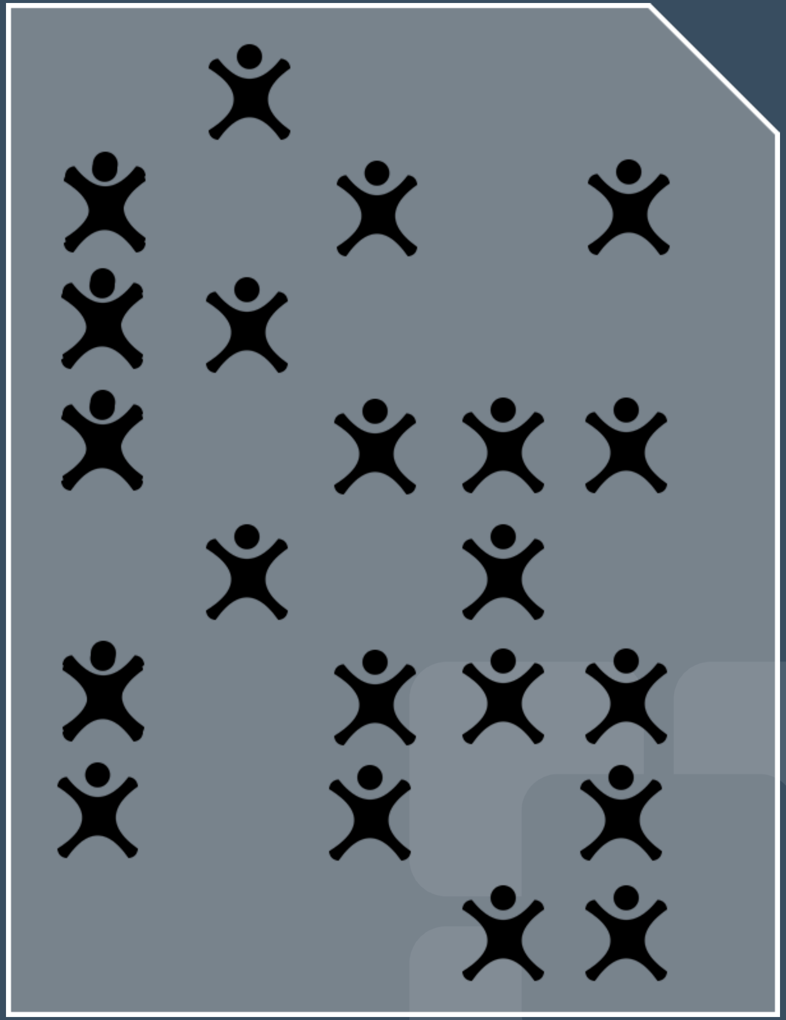
All patients at the Cleveland Clinic can qualify to be placed on a care gap registry which is live in the electronic medical record

Only a subset of those patients are placed on the High Risk Registry and actively outreached by care coordinators

Care Gap Registry



High Risk Registry



The Analytics Journey

How does a patient qualify for the High Risk Registry?

Initial Cohort

All of:

- Risk contract
- Employed or unassigned PCP
- 11 county area

175,000

+

Chronic Disease

One or more of:

Asthma, Atrial Fib, CAD, CHF, COPD, DM, Dyslipidemia, HTN

100,000

−

Exclusions

- Hospice

95,000

+

Utilization & Care Gaps

One or more of:

- >80^{th} %ile admit risk (CHF, DM, Asthma, COPD)
- >80^{th} %ile admit risk (generic, claims model)
- 4 or more ED visits in the past 1 year
- 1 or more evidence-based care gaps

80,000

+

Comorbidities

- Charlson Score $\geq$ 5

7,000

✓

The Analytics Journey

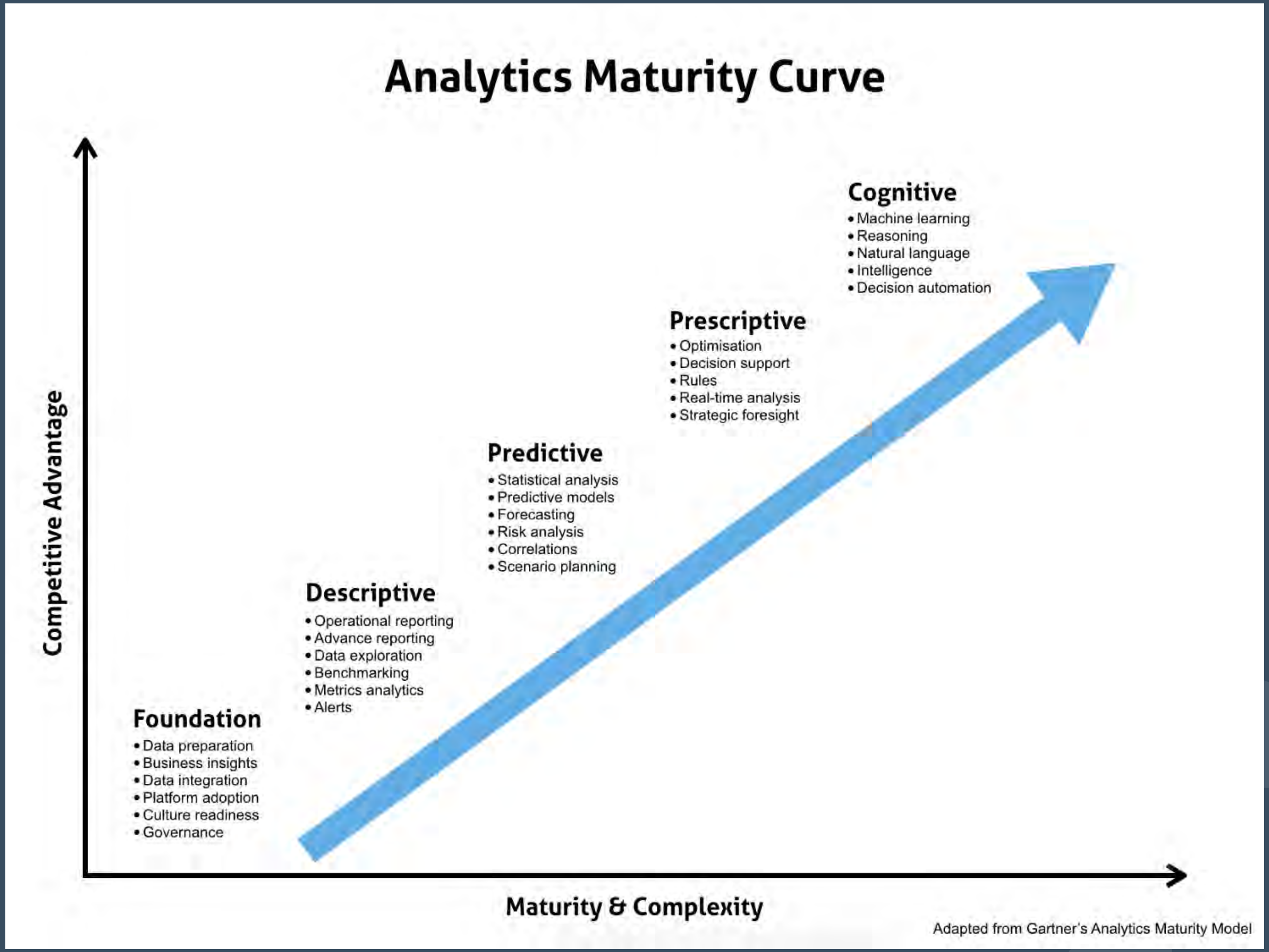
How does a patient qualify for the High Risk Registry?

1. Relied heavily on physician intuition
2. Lacked differentiation between patients
3. Difficult to apply across an entire population

The Analytics Journey

How does a patient qualify for the High Risk Registry?

To solve for these pitfalls of the original High Risk Registry criteria, Cleveland Clinic adopted machine learning processes to move from descriptive to predictive

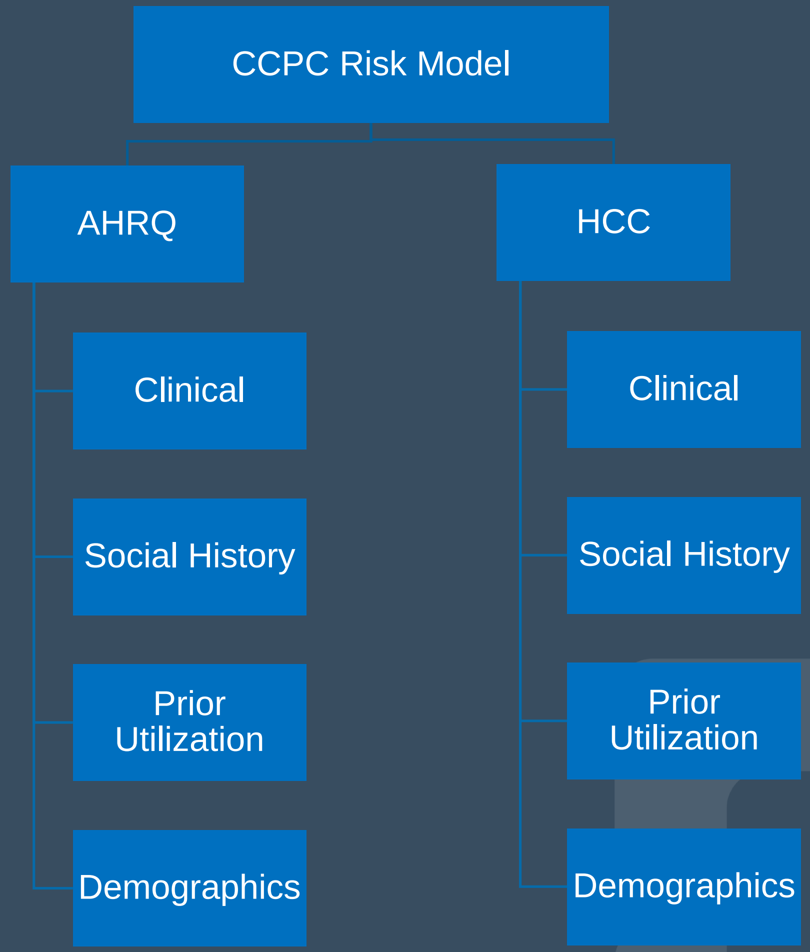


Modeling Objective

Cleveland Clinic Patient Care Risk Model

- Calculate a risk score for each patient in the target population for clinical care coordination by predicting future Cleveland Clinic based direct cost of care resource utilization.
- Target variable: **<u>Direct Cost</u>**, The patient predicted direct cost will determine the overall Risk score for that patient.
- Predictive variables: 300+ variables considered for the model.
- Total number of patients receiving a score: 798,586

Two different diagnosis groupers (AHRQ, HCC) will be tested to determine the best fitting models.



Data Sources and Condition Groupers

- All models use data from the Cleveland Clinic Enterprise Data Vault

- Historical diagnosis and cost info from internal billing

- CMS-HCC Model: 79 Hierarchical Condition Categories from the Centers for Medicare & Medicaid Services
  - All CMS categories included as Binary Variables
  - Count of all HCC diagnoses variable included

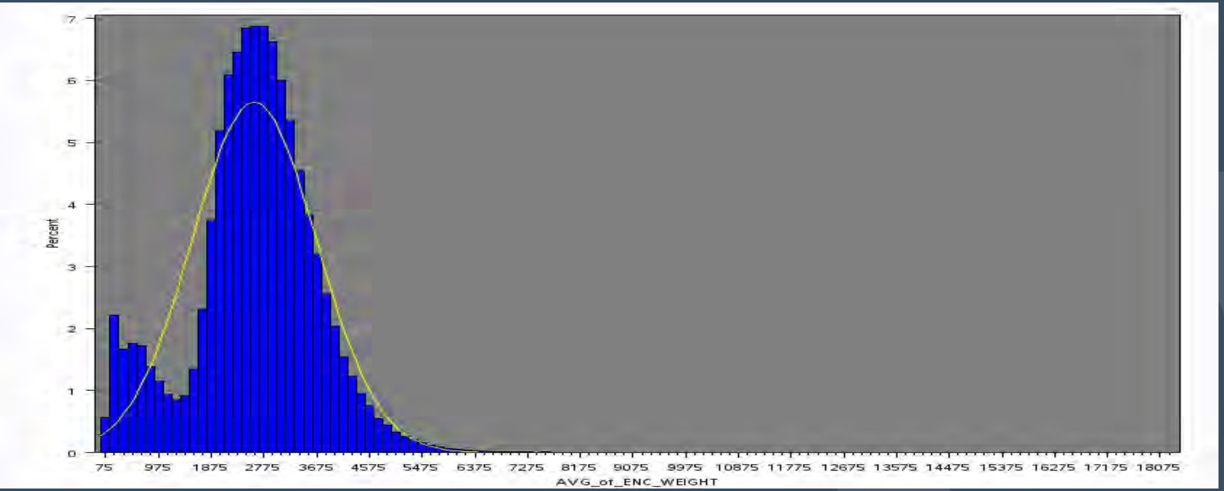
- Agency of Healthcare Research and Quality Model: 278 categories
  - All categories included as Binary Variables
  - Count of all AHRQ diagnoses variable included

Example of Variables Used in the Model

Alcohol User	Employment Status	Systolic Blood Pressure (SBP)	Diastolic Blood Pressure (DBP)	Illicit Drug User	Calcium Level
Creatinine Score	Body Mass Index (BMI)	Body Surface Area (BSA)	Potassium Level	LDL	Sodium Level
Weight	Glucose Level	Hemoglobin	Log-Transformed Allowed Amount	Patient County	Patient Language
Platelet Count	Red Blood Cell Count	White Blood Cell Count	Patient Age Bins	Under 18	Over 65
Marital Status	Sexually Active	Tobacco User	Cholesterol	Living in Florida	Market Area
Global CCHS Group	Prior Year Direct Cost	Number of Visits			

Replacement and Imputation

- Each Clinical Variable included 4 variations, for example:
  - Average of Encounter Weight
  - Minimum of Encounter Weight
  - Maximum of Encounter Weight
  - Most Recent Encounter Weight
- Outlier values were removed and missing values were imputed
- Some continuous variables were transformed to categorical
  - E.g. A cholesterol reading of 200 or greater was more predictive than the continuous value

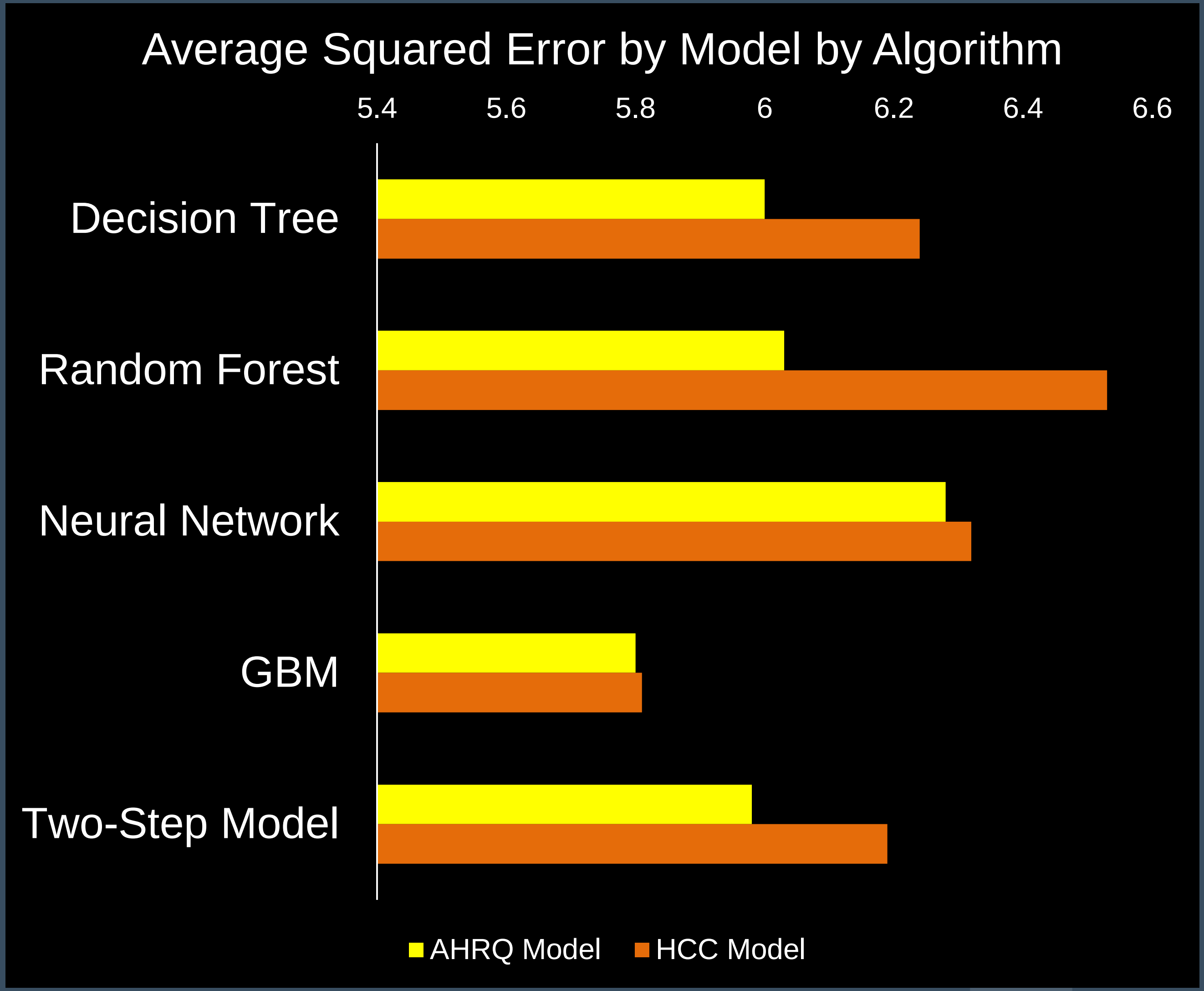


Model Validation

Several machine learning algorithms were tested in the creation of the Cleveland Clinic Primary Care Risk Score

In all cases, the AHRQ diagnosis classification performed better than the HCC grouper

Eventually a gradient boosted machine proved to have the best validation average squared error



Most Important Variables – AHRQ GBM

AHRQ Variable	Rules	Importance
Days from Most Recent Visit	236	1.000
Direct Cost (Lag 1 Year)	108	0.880
Distinct AHRQ Dx Categories	53	0.852
Direct Cost (Lag 2 Year)	56	0.489
Direct Cost (Lag 3 Year)	43	0.444
Discharge Status	130	0.350
CC PCP Indicator	38	0.163
Patient County	198	0.162
CCHS Payer Group	316	0.152
Patient Age	129	0.146
Prior Admits	3	0.103
Prior ED Visits	32	0.077
Minimum LDL	6	0.074

The resulting Gradient Boosted Machine only used a subset of the initial variables that were tested

Variables that led to node splits in the tree ranged from cost and utilization metrics, to lab values and patient demographic info

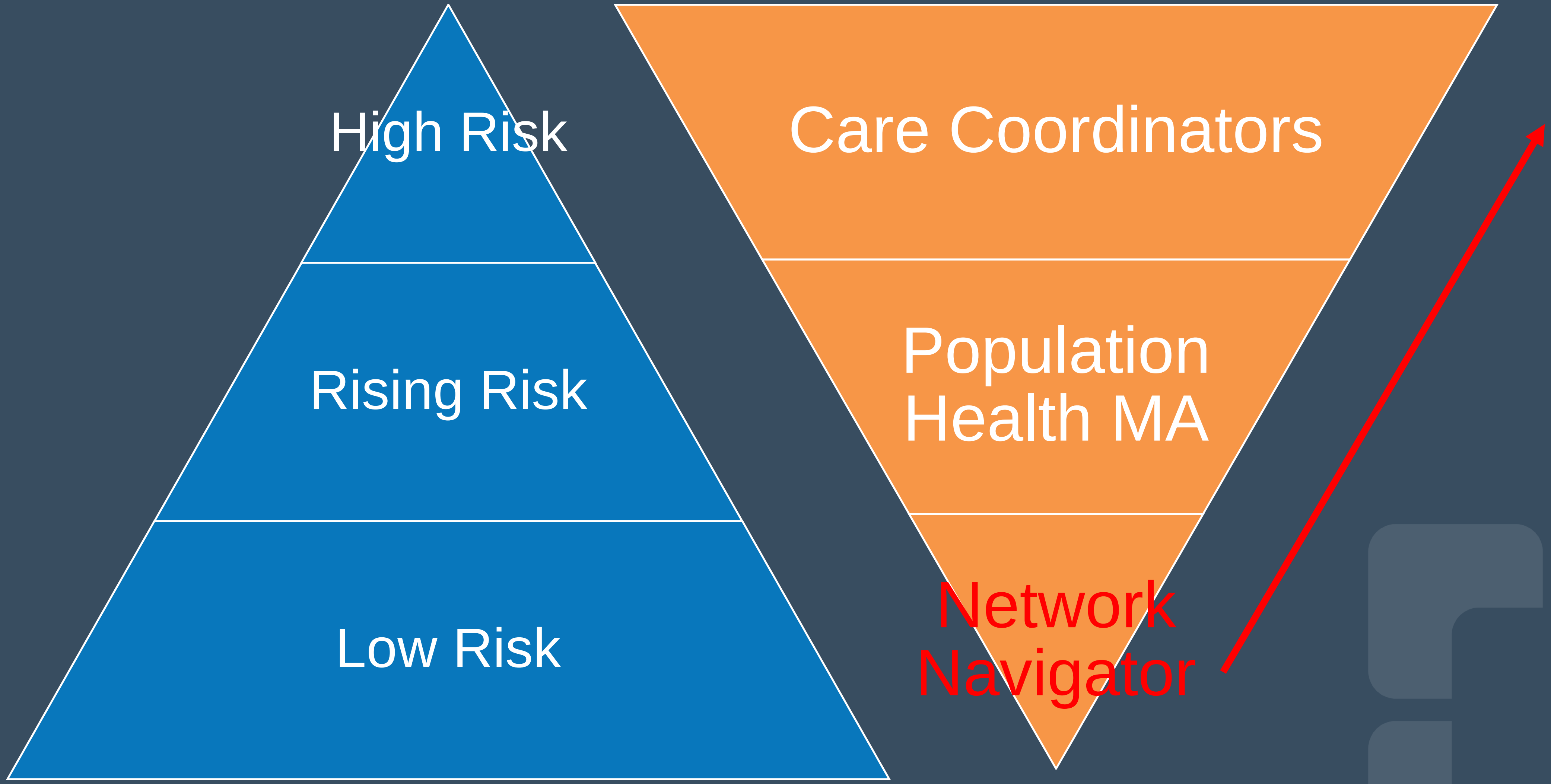
Three years of historical cost and diagnosis information proved to be some of the most important data in the model

Model performance was consistent with expectations

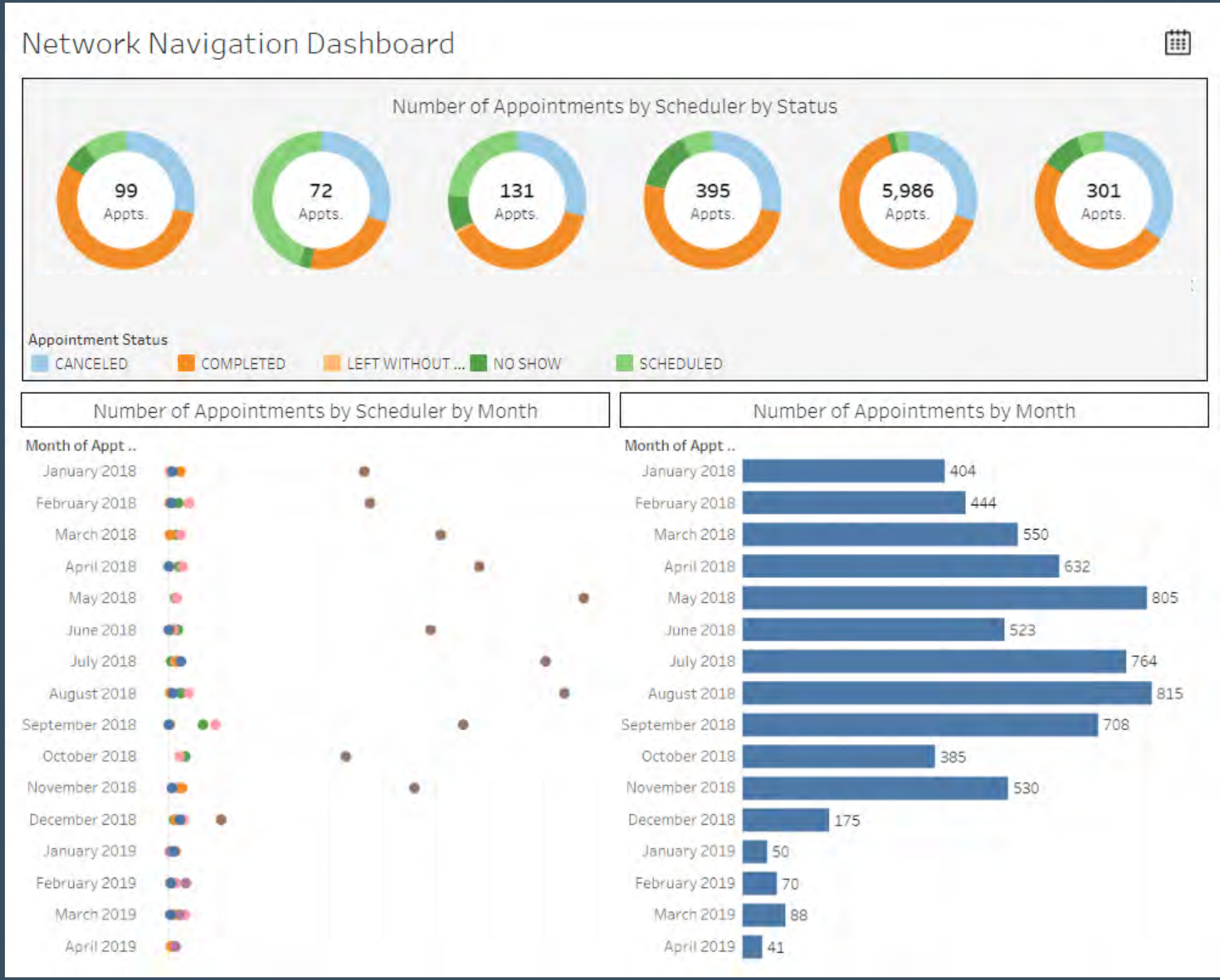
Prediction Decile	Average AHRQ Prediction	Average Number of Admits	Average Number of Visits	PATIENT AGE	Distinct AHRQ Category Count	Average Number of Encounters	Average 2017 Direct Cost	Average 2016 Direct Cost	Average 2015 Direct Cost
0	8.6	1.0	2.5	61.5	15.2	53.3	15,480.3	8,647.4	6,401.8
1	7.4	0.3	1.2	51.1	8.9	24.7	3,624.6	2,380.8	2,040.4
2	6.6	0.2	0.9	45.5	6.4	16.0	1,970.8	1,424.9	1,271.3
3	5.9	0.1	0.7	43.1	4.8	11.5	1,110.0	1,113.1	970.2
4	5.3	0.1	0.6	42.2	3.5	8.2	665.3	868.1	714.9
5	4.6	0.1	0.5	41.9	2.6	5.9	327.5	743.0	542.7
6	4.0	0.1	0.4	41.0	1.9	4.3	164.7	626.3	457.4
7	3.4	0.1	0.5	38.0	1.7	3.7	65.0	707.4	541.5
8	2.4	0.1	0.5	40.9	1.7	3.8	2.9	657.1	1,030.3
9	1.2	0.1	0.4	43.3	1.4	2.9	0.3	206.0	1,637.5

*Prediction Decile is based on the AHRQ Prediction. The highest predictions are in Decile 0.

Population Health Management Approach



Network Navigation



Network Navigator Responsibilities

1) Appointment Scheduling
2) Orders for Evidence Based Medicine
3) Outreach Calls
4) PCP Transitioning for Patients

Network Navigation and Care Coordination

Patient's CCPC Score is Calculated

Patient directed to care coordinator's outreach list

Patient has typical office visit or interaction with health system

Now What?

Clinical judgment

What are the options

Lots of data, but what's most important

Recommendation Model

As we identify patients who are high risk (CCPC Model), how do we then direct patients to the next point of care?

High Risk Patient

E7777777	
Diabetes	1
ER Visits	3
HbA1c	9.1

Distance from Patient One in Library
$= \mathrm{sqrt}((1-0)^2 + (3-0)^2 + (9.1-7.0)^2)$
$= 3.8$

Library of Historic Patients

Title	Diabetes	ER Visits	HbA1c	
E1111111	0	0	7.0	3.8
E2222222	0	10	8.0	7.2
E3333333	0	0	6.4	4.2
E4444444	1	2	9.4	1.0
E5555555	1	1	6.5	3.3
E6666666	1	0	8.9	3.0

Recommendation Model

As we identify patients who are high risk (CCPC Model), how do we then direct patients to the next point of care?

High Risk Patient

E7777777	
Diabetes	1
ER Visits	3
HbA1c	9.1

Distance from Patient One in Library
$= \text{sqrt}((1-0)^2 + (3-0)^2 + (9.1-7.0)^2)$
$= 3.8$

Library of Historic Patients

Title	Diabetes	ER Visits	HbA1c	
E1111111	0	0	7.0	3.8
E2222222	0	10	8.0	7.2
E3333333	0	0	6.4	4.2
E4444444	1	2	9.4	1.0
E5555555	1	1	6.5	3.3
E6666666	1	[illegible]	[illegible]	[illegible]

The nearest neighbor was recommended for a visit to a nephrologist, so we might suspect a similar course of action for the high risk patient

Recommendation Model

Two important questions emerge….

1. How accurate is this prediction?
2. How accurate does the model need to be?

Recommendation Model

Approach #2:

Develop a classification model that would identified the proper recommendation for patients using order sets defined in the electronic medical record

Recommendation Areas

Referral to Palliative Care

Consult to Primary Care Coordination

Consult to Pharmacy

Referral to Primary Care

Consult to Specialist

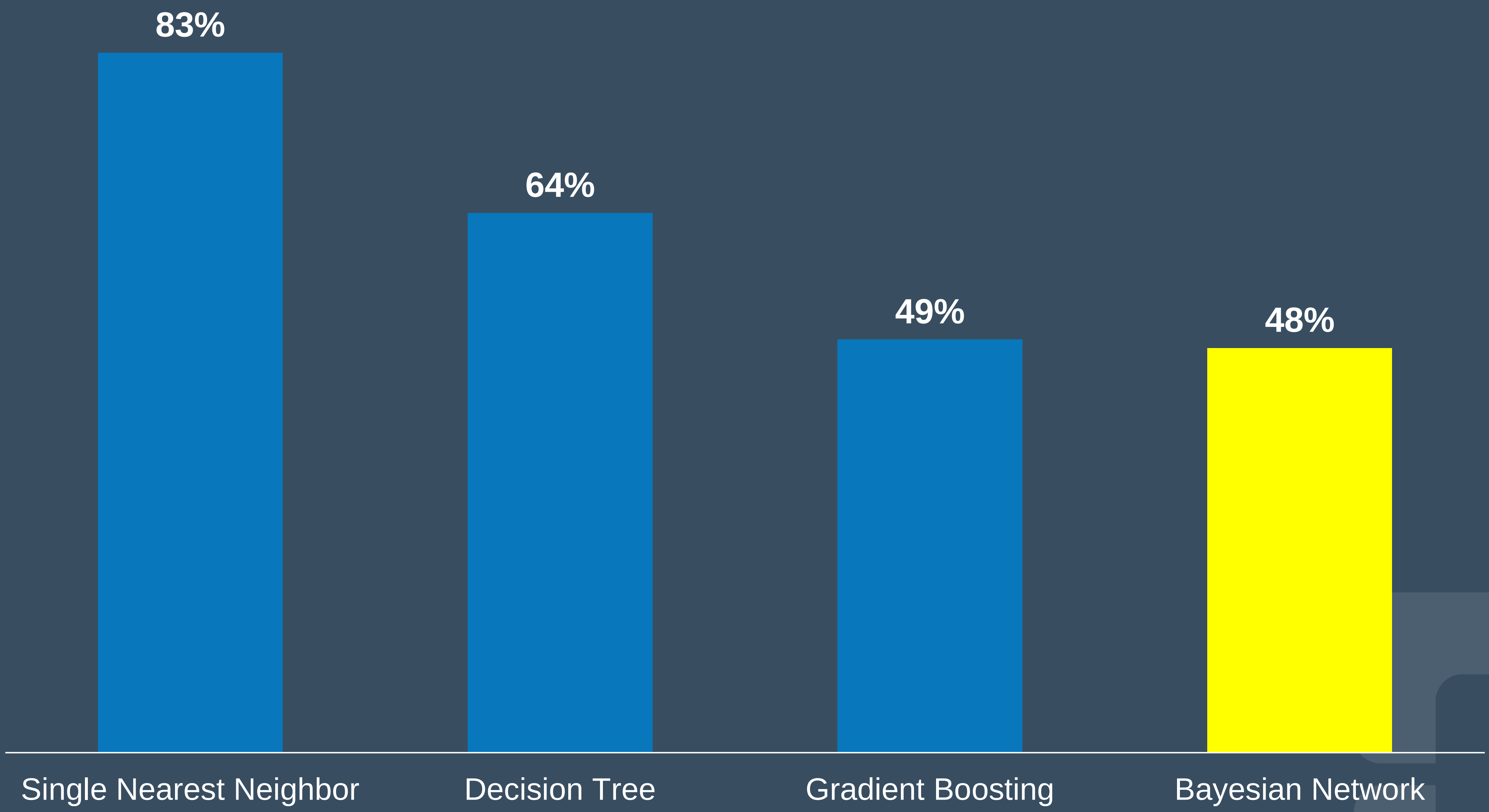
Referral to Home Health

Modeling Dataset

Variables	
Recommendation / Order Set (Target)	
Patient Age	▲
Patient Gender	▲
Last Encounter Diagnosis Category	▲
Historical Diagnosis Information	
Last Encounter Clinical Measures	
Last Encounter Physician Specialty	▲
Lab Result (HbA1c, LDL, eGFR)	
Admit Utilizations (Last 12 Months)	
ER Utilizations (Last 12 Months)	▲

Five most important variables identified with a triangle

Misclassification Rate by Algorithm



Open Questions

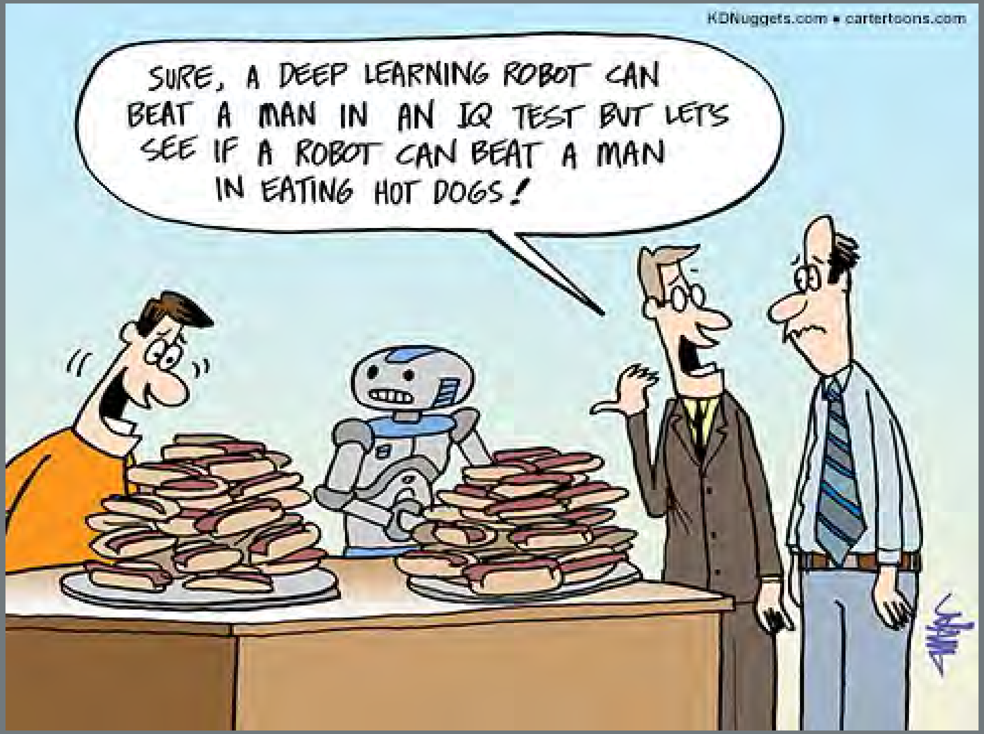
1. What’s an acceptable misclassification rate?

2. Is there more utility in a single prediction or multiple suggestions?

3. What is the tradeoff between <u>supplementing</u> and <u>focusing</u> clinical judgment?

“But the algorithms could be relying on inaccurate public data, and they may disempower patients, leaving them in the dark about the Big Brotherish systems rating them. Another key challenge, says Case Western’s Hoffman, is ensuring that the predictions don’t override a clinicians’ instinct or reinforce biases.”

—Politico

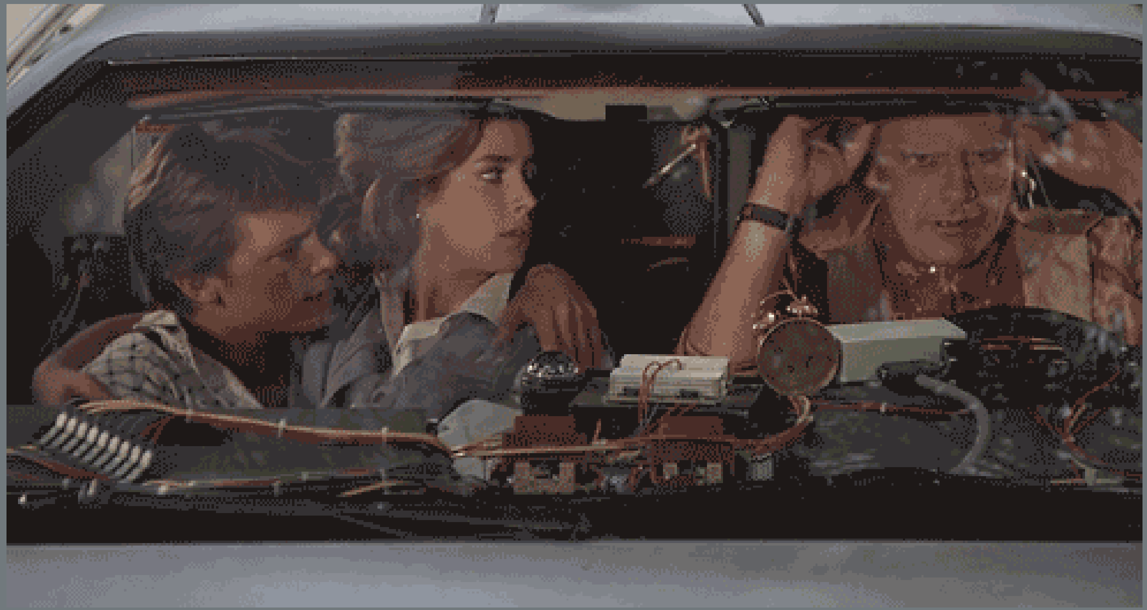


Clinical validation of results

- Data collection/sources
  - Is your training data appropriate for the task?
- Feature engineering
  - Do your drug and condition groupers make sense?
  - Should constraints be placed on variables?
- Feature selection
  - Are some variables outdated or proxies for something else?
- Face validity of the output
  - Do the predictions and their drivers make sense?
  - How do wrong predictions affect people?

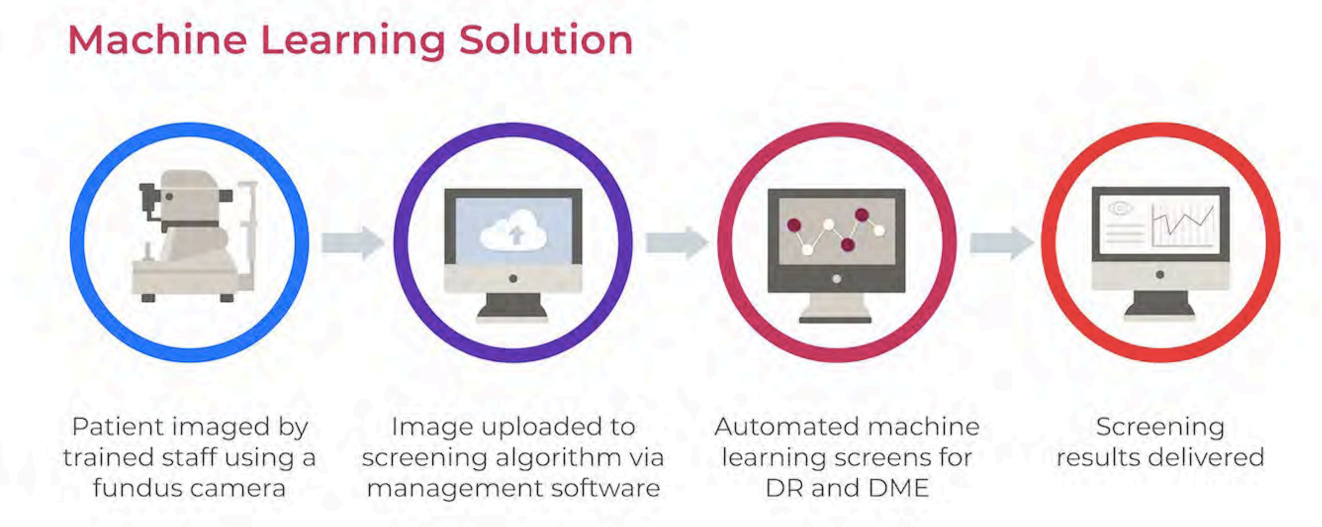
“Roads? Where we’re going, we don’t need roads.”

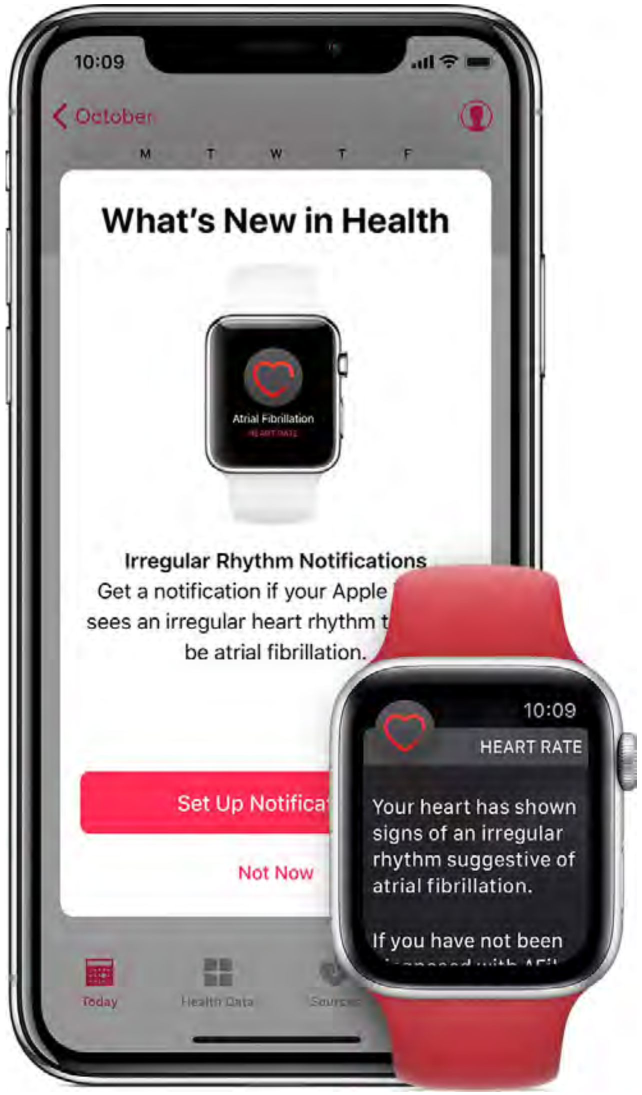
—Dr. Emmet Brown



Now and the near future

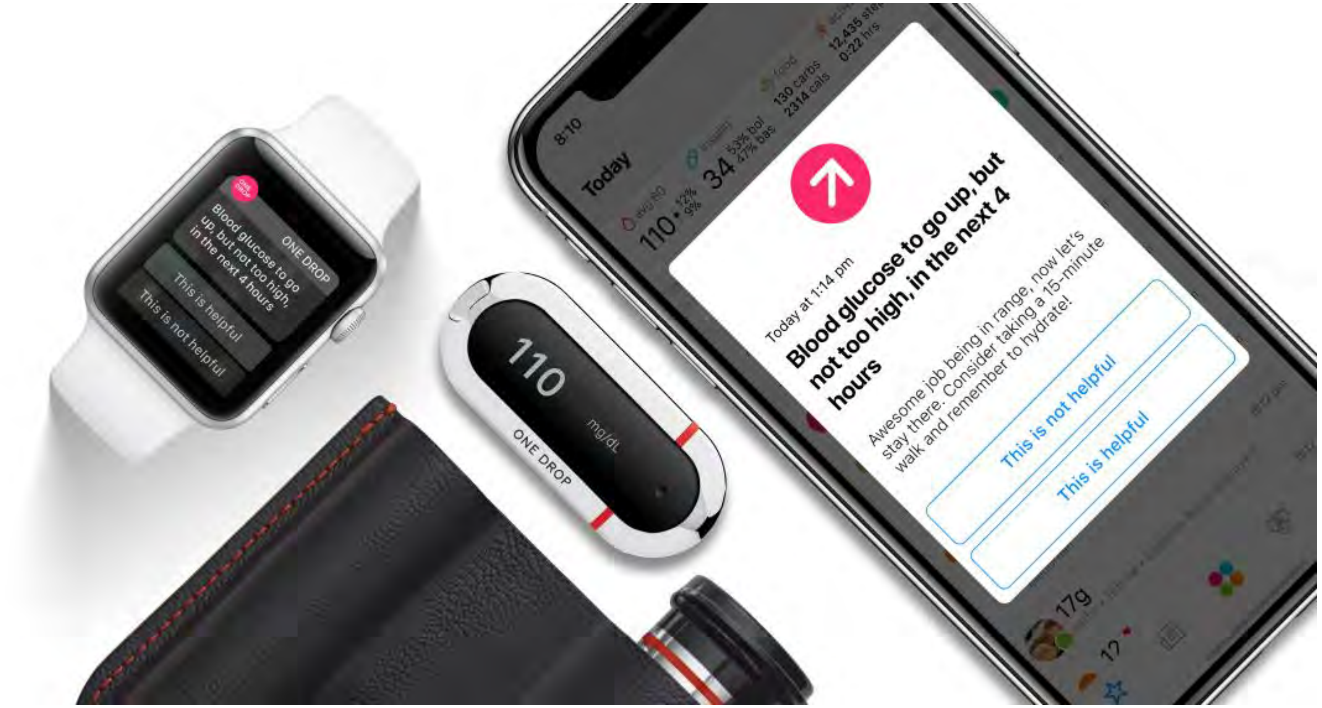
- Detecting early warning signs
  - The new Apple Watch looks for signs of atrial fibrillation
  - Google and Verily using AI for clinical use
    - Image recognition to diagnose diabetes-related eye conditions
    - Early detection can prevent blindness





Now and the near future

- What's next
  - Wearable tech + EHR + SDoH + Consumer Data
  - "Live" risk scores and other predictions
  - Streaming preventative health recommendations
    - Dietary needs
    - Exercise routine
    - Doctor visits
    - Medication changes



Sources

- Accuracy of Claims-Based Risk Scoring Models
  - https://www.soa.org/Files/Research/research-2016-accuracy-claims-based-risk-scoring-models.pdf
- XGBoost
  - https://github.com/dmlc/xgboost
- SHapley Additive exPlanations (SHAP)
  - https://arxiv.org/pdf/1705.07874
- Google and Verily using AI for clinical use
  - https://blog.verily.com/2019/02/launching-powerful-new-screening-tool.html
- Apple watch heart rate notifications
  - https://support.apple.com/en-us/HT208931
- Collaborative Filtering for Medical Conditions
  - https://www.soa.org/Library/Newsletters/Predictive-Analytics-and-Futurism/2016/december/paf-iss14-parkes-copeland.pdf

Collaborative Filtering Example

Conditions to Consider					Latest Claim Info				
Condition ID	Condition Description	Chronic	Probability	Risk Score Contribution	Date	Claim Detail	Provider Name	Provider Specialty	Provider ID
hcc96	HCC96 - Specified Heart Arrhythmias	Y	29%	0.27					

Primary Reasons to Consider		Latest Claim Info				
Reason Description	Contribution	Date	Claim Detail	Provider Name	Provider Specialty	Provider ID
CCS106 - Cardiac dysrhythmias		2018-07-16	R002: Palpitations	Santacruz, David	Certified clin...	2QK78WPXT...
CCS105 - Conduction disorders		2017-06-07	Z950: Presence of cardiac pacemaker	Apollo Hospital	Hospital	PNNK7E45...
CCS108 - Congestive heart failure; nonhypertensive		2018-06-11	I509: Heart failure, unspecified	Bruder, Flossie	Nephrology	FWRGY5J27...

Milliman

Hans Leida and Joe Long

hans.leida@milliman.com
joe.long@milliman.com

Cleveland Clinic

Joe Dorocak and Don McLellan

dorocaj@ccf.org
mclelld2@ccf.org

Thank You