The FIND FH® Initiative

www.theFHfoundation.org

The FIND FH® Initiative
Finding Undiagnosed Familial Hypercholesterolemia Patients in the U.S.

Joshua W. Knowles, MD, PhD
Stanford University and the FH Foundation
knowlej@stanford.edu

Kelly D. Myers
The FH Foundation
KM@theFHfoundation.org

Our Mission

The FH Foundation is a patient-centered nonprofit organization, dedicated to **education**, **advocacy,** and **research** of Familial Hypercholesterolemia (FH).

Our mission is to raise awareness of FH and **save lives** by increasing the rate of early diagnosis and encouraging proactive treatment.
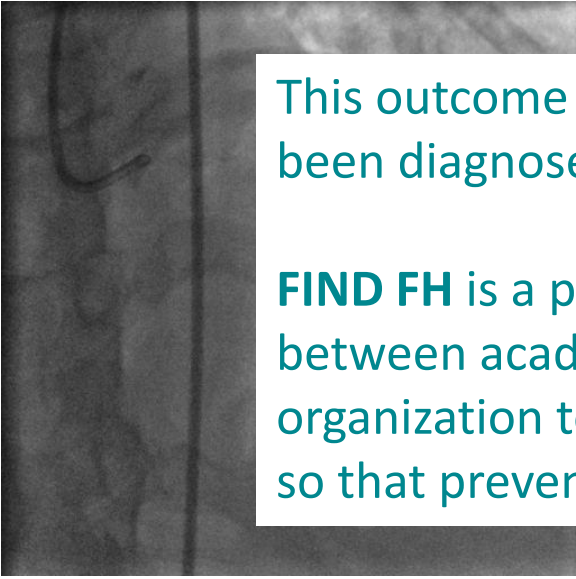
Case Presentation

38-year-old female

- Normal childhood. In high school told she had elevated cholesterol but no therapy ("too young").

- Married 2004, first child 2009

- Early 2010: began to have chest pain

- Summer 2010: PCP checked cholesterol -> very elevated -> referred to cardiologist -> stress test markedly positive -> scheduled for angiogram

- While sleeping on morning of scheduled cath began to groan, act in distress, unresponsive

- Husband began CPR -> EMS -> Vfib arrest -> resuscitated -> cooling protocol -> Troponin 10, EF 20% -> Transferred to Stanford
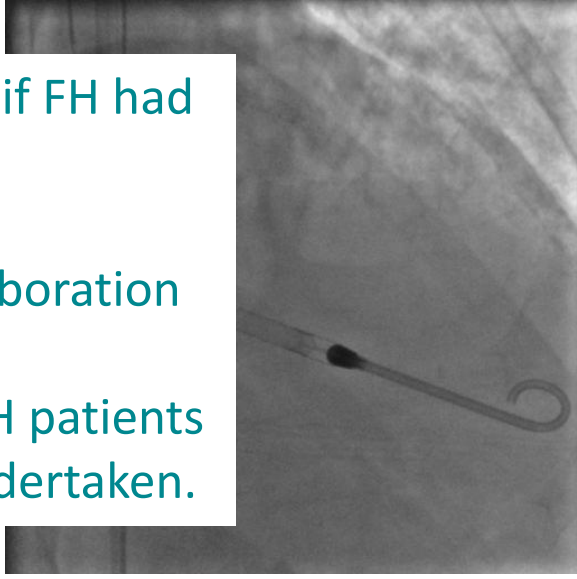
Angiogram

Severe 3 vessel coronary disease

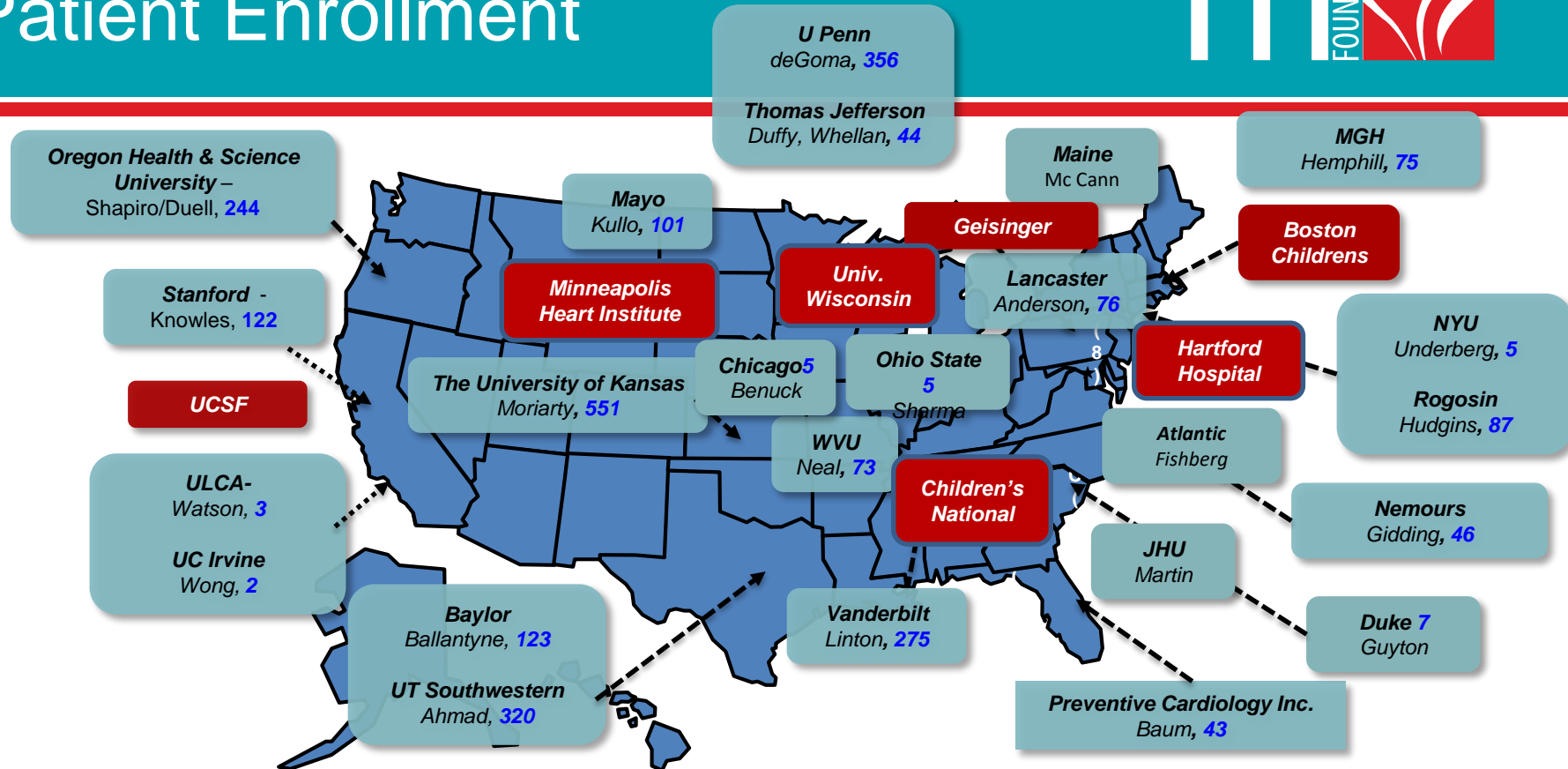Placement of stent and ventricular assist device

This outcome could have been averted if FH had been diagnosed and treated earlier.
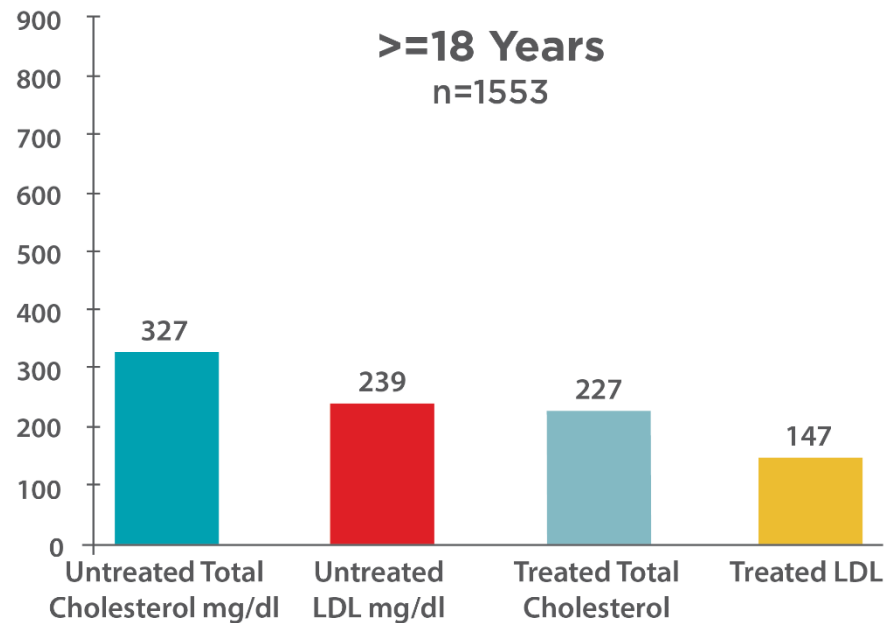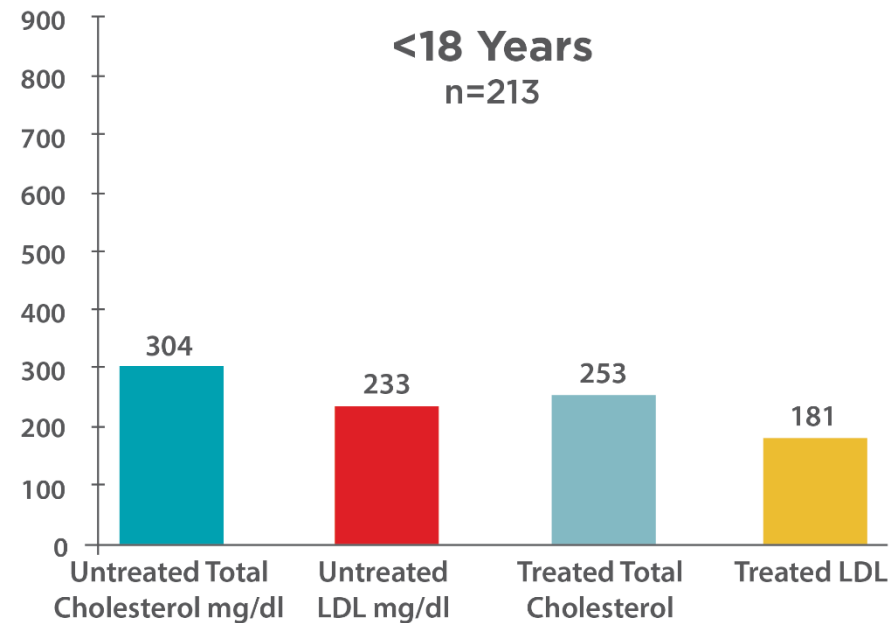
FIND FH is a project designed as a collaboration between academia and a patient-led organization to identify undiagnosed FH patients so that preventive measures can be undertaken.

Patient Enrollment

U Penn
deGoma, **356**

Thomas Jefferson
Duffy, Whellan, **44**

Oregon Health & Science University – Shapiro/Duell, **244**

Maine
Mc Cann

MGH
Hemphill, **75**

Mayo
Kullo, **101**

Geisinger

Boston Childrens

Stanford - Knowles, **122**

Minneapolis Heart Institute

Univ. Wisconsin

Lancaster
Anderson, **76**

Hartford Hospital

NYU
Underberg, **5**

UCSF

The University of Kansas
Moriarty, **551**

Chicago 5
Benuck

Ohio State
5
Sharma

8

Rogosin
Hudgins, **87**

ULCA-
Watson, **3**

WVU
Neal, **73**

Atlantic
Fishberg

Nemours
Gidding, **46**

UC Irvine
Wong, **2**

Children's National

JHU
Martin

Baylor
Ballantyne, **123**

Vanderbilt
Linton, **275**

Duke 7
Guyton

UT Southwestern
Ahmad, **320**

Preventive Cardiology Inc.
Baum, **43**

Lipids for HeFH Patients

<18 Years
n=213

	Value
Untreated Total Cholesterol mg/dl	304
Untreated LDL mg/dl	233
Treated Total Cholesterol	253
Treated LDL	181

>=18 Years
n=1553

	Value
Untreated Total Cholesterol mg/dl	327
Untreated LDL mg/dl	239
Treated Total Cholesterol	227
Treated LDL	147

Cardiovascular Disease

ASCVD, n=1273 .... **38%**

    Age at onset, years **52**

CHD, overall cohort **36%**

    Age at onset, years **51**

CHD, men **47%**

    Age at onset, years **47**

CHD, women **29%**

    Age at onset, years **55**

Stroke or TIA, n=1282 .... **5%**

Aortic valve disease, n=1284 .... **3%**

Treated LDL-C Values & Reduction

Treated LDL-C[*]		Statin-treated	Not statin-treated
		n=959	n=125
<70 mg/dl		58 (6%)	5 (4%)
70-99 mg/dl		194 (20%)	11 (9%)
100-129 mg/dl		238 (25%)	7 (6%)
130-159 mg/dl		153 (16%)	35 (28%)
160-189 mg/dl		113 (12%)	22 (18%)
≥190 mg/dl		203 (21%)	45 (36%)
LDL-C reduction[*]		n=576	n=76
≥50%		257 (45%)	9 (12%)

High Prevalence of CHD
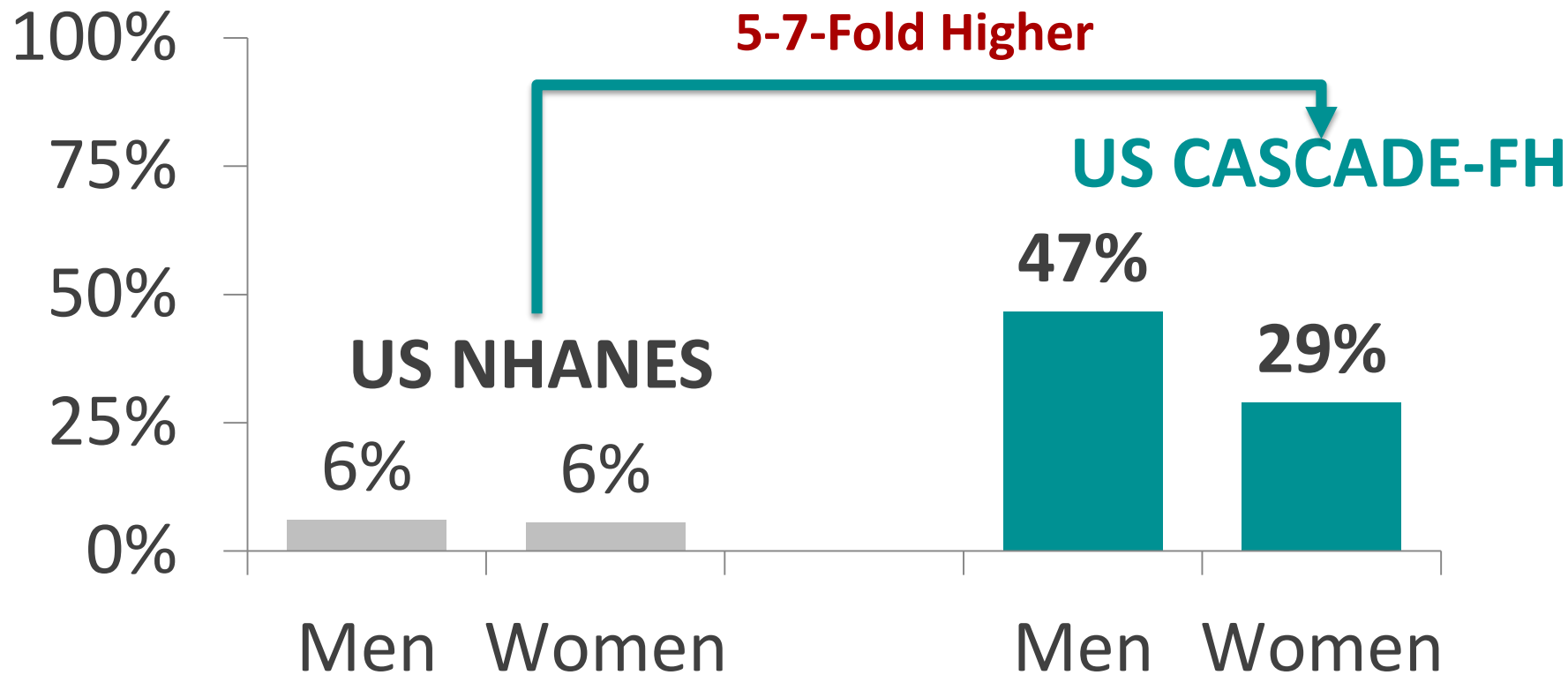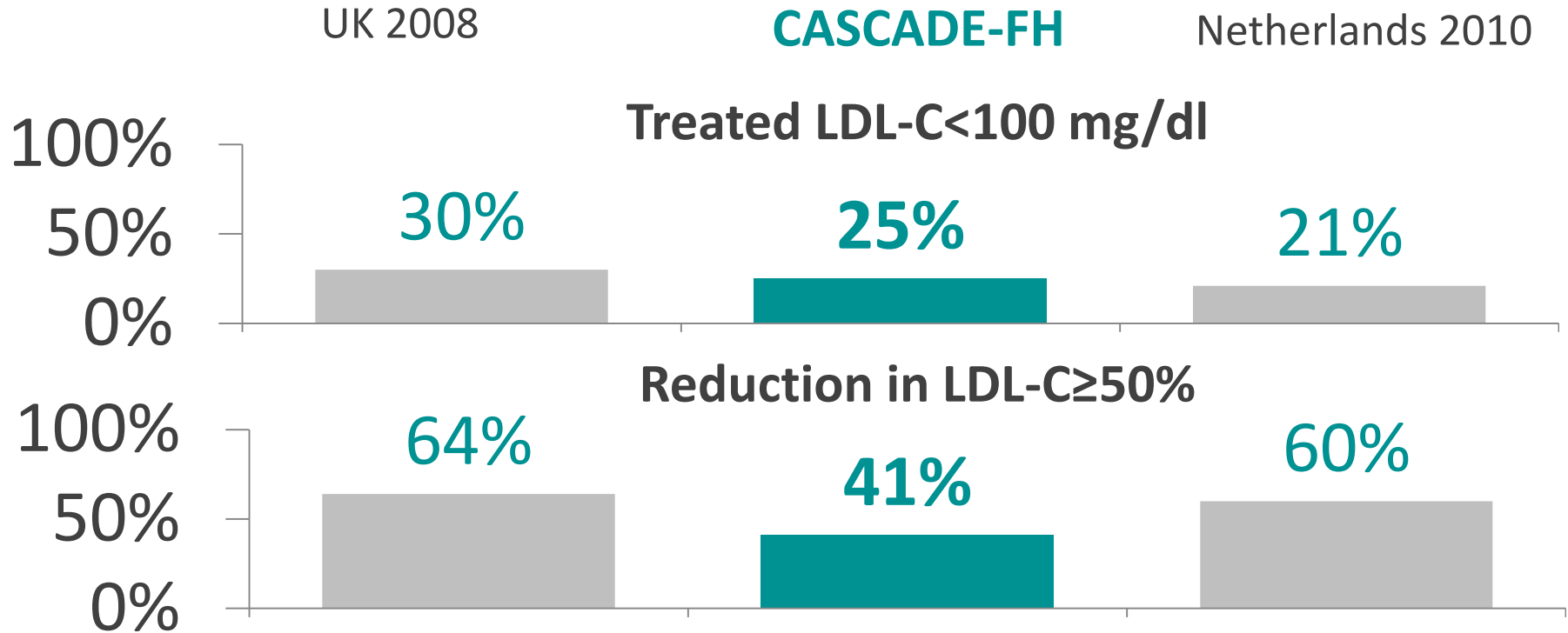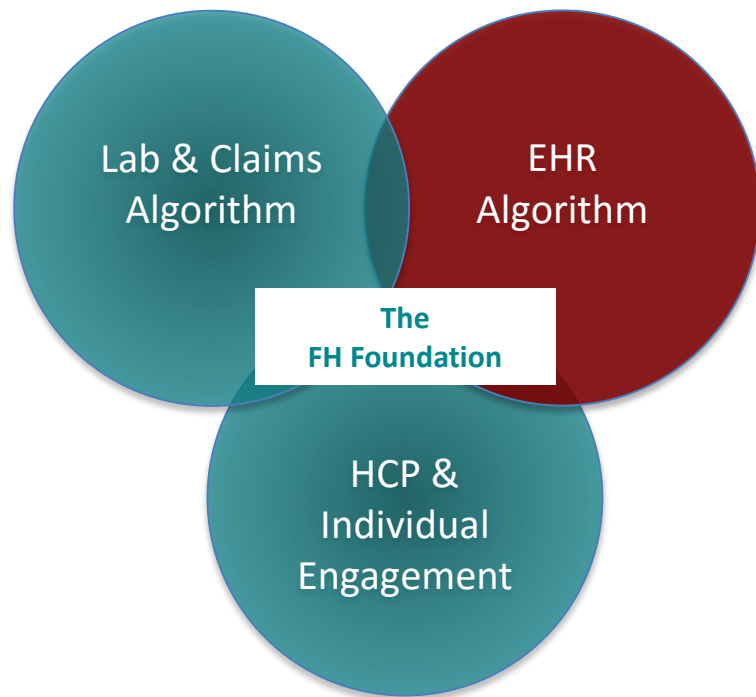
100%

5-7-Fold Higher

75%

US CASCADE-FH

50%

47%

US NHANES

29%

25%

6%   6%

0%

Men   Women   Men   Women

FH Foundation International Summit. September 24-25, 2016. Los Angeles, CA.

Goal Attainment

UK 2008   **CASCADE-FH**   Netherlands 2010

Treated LDL-C<100 mg/dl

100% | 30% | **25%** | 21%
50%
0%

Reduction in LDL-C≥50%

100% | 64% | **41%** | 60%
50%
0%

Pijlman et al. Atherosclerosis. 2010;209:189; Hadfield et al. Annals Clin Biochem. 2008;45:199.

FIND FH®

A multiyear screening and engagement initiative to identify and encourage the diagnosis and treatment of FH

Lab & Claims Algorithm

EHR Algorithm

The FH Foundation

HCP & Individual Engagement

Lab & Claims Data Mining

- Healthcare encounter data on 89 million Americans with cardiovascular disease
- Data from a significant majority of clinical practices

EHR Data Mining

- Comprehensive EHR data from two academic centers
- Expanding to key integrated health systems

HCP & Individual Engagement

- Multichannel tools to engage health systems and individual HCPs
- Tools for clinicians and individuals with FH

Machine Learning

- Software that learns by example.

- We show the model examples of FH and Non-FH patients.

- Patients are described to the model using features (inputs):

  - Lab Results

  - Patient Age

  - ICD9 codes

  - Etc…

- The model learns correlation between certain features and FH rate.

- Model can classify FH in new patients using just their features.

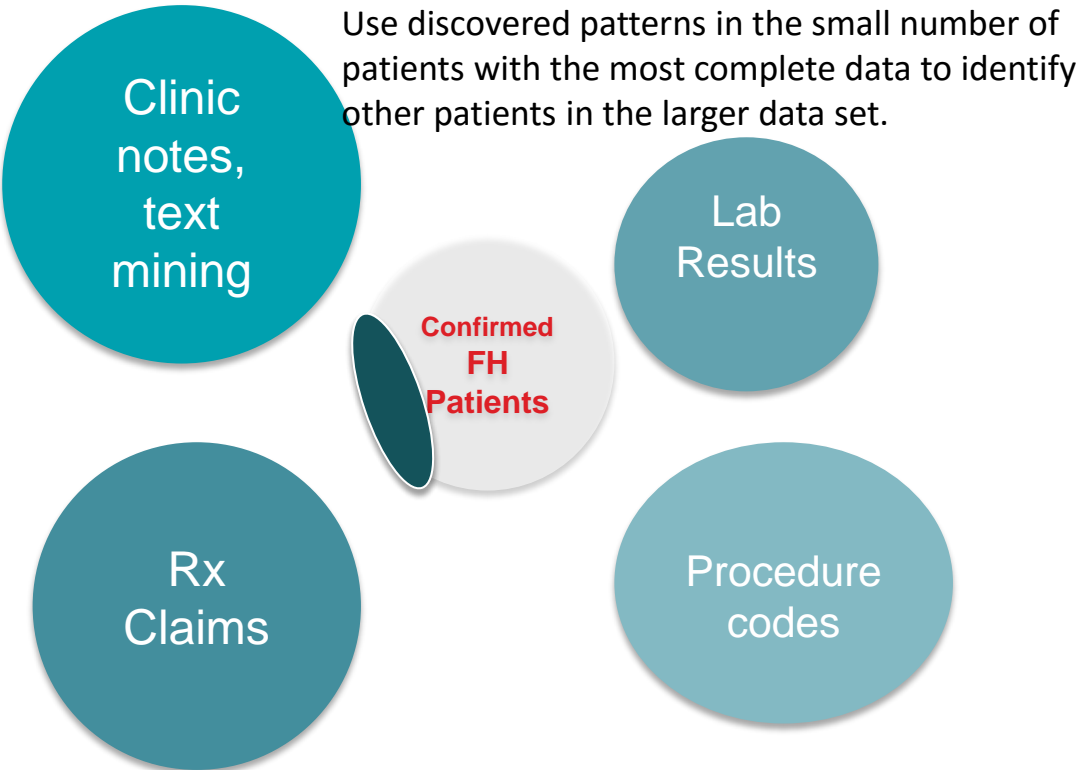# Machine Learning Example
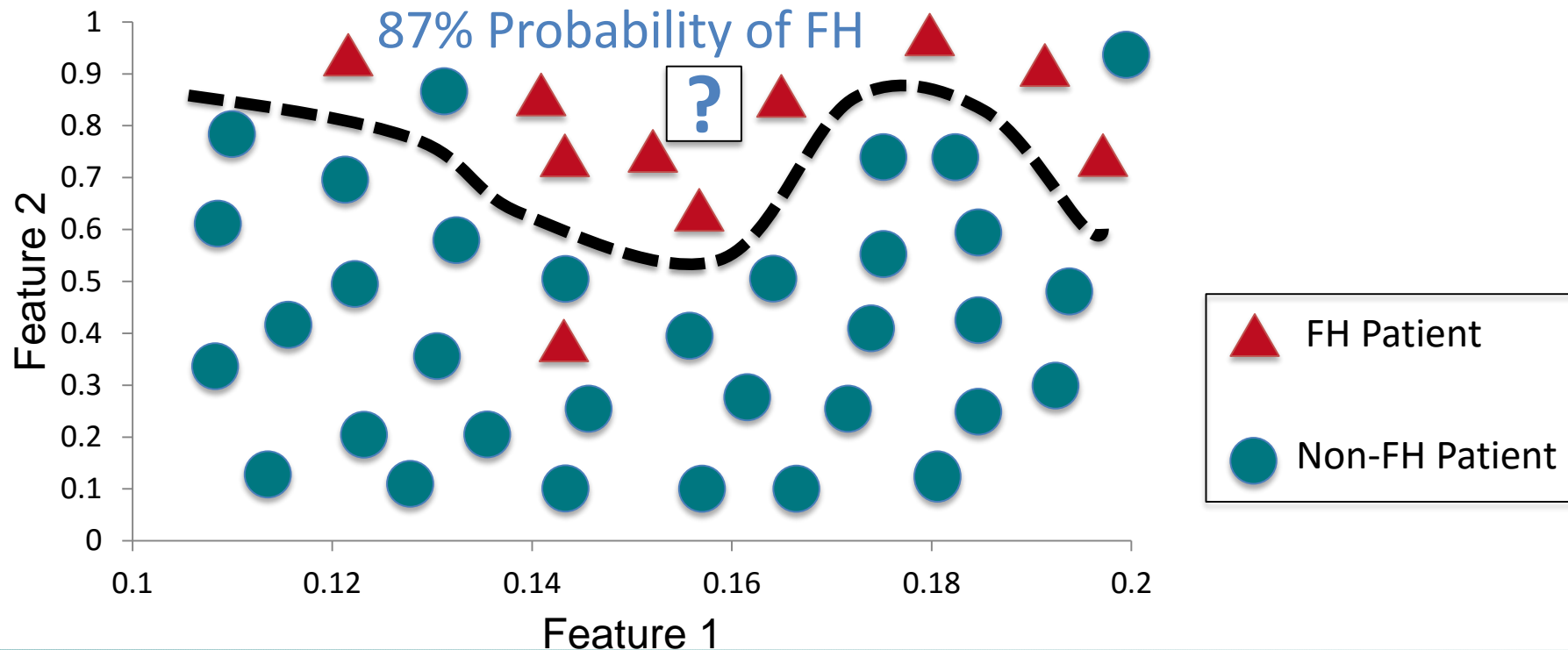
The Sources of Features (Weber et al)

Weber GM et al. *JAMA*. 2014;311:2479-80.
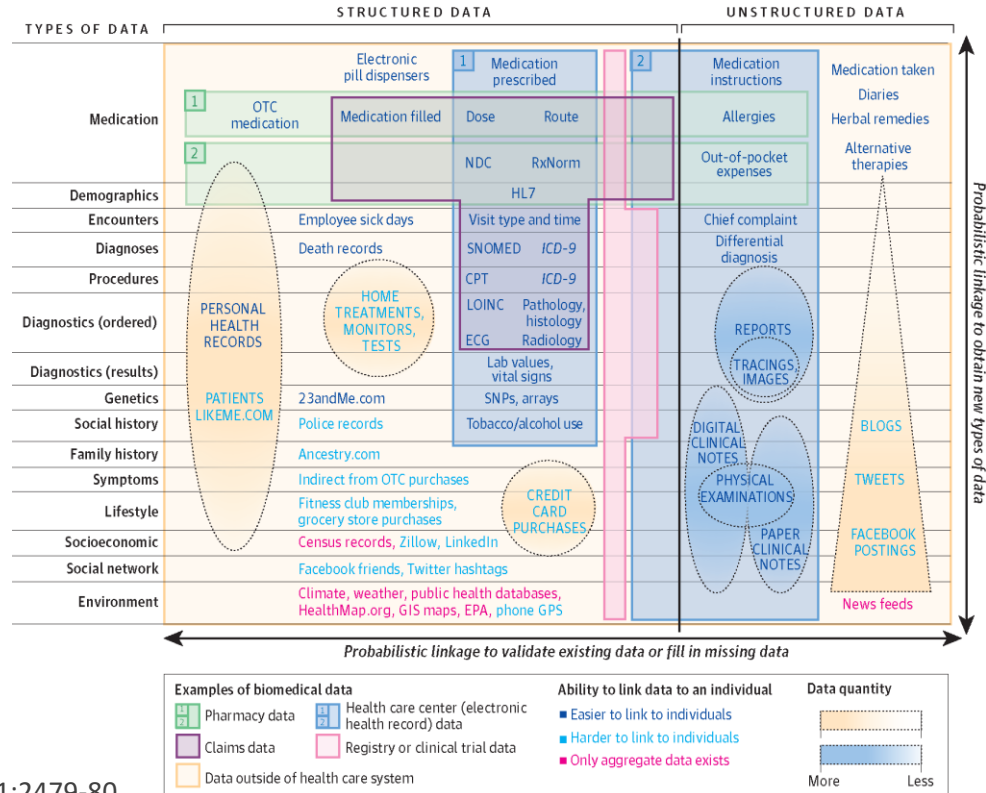
Identifying FH Patient Characteristics Using Orthogonal Data

Use discovered patterns in the small number of patients with the most complete data to identify other patients in the larger data set.

Clinic notes, text mining

Lab Results

Confirmed **FH** Patients

Rx Claims

Procedure codes

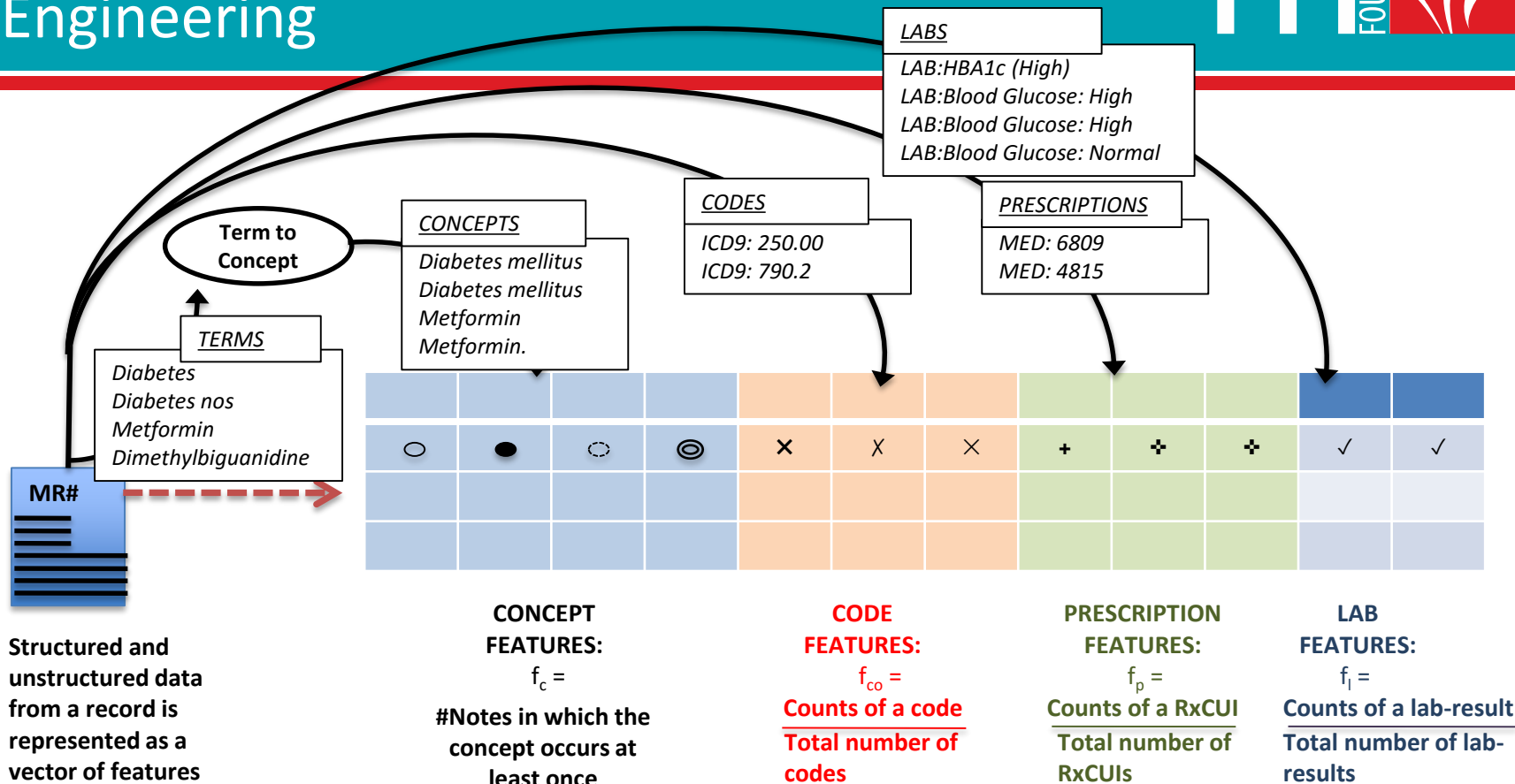|	
Unstructured data	Clinic notes, dictations for key words, phrases • Personal medical Hx: age of cardiac event, procedure • Disease names: FH, hyperlipidemia • Family history: premature coronary disease • Signs: xanthoma, xanthelasma, arcus
Structured data	Labs: LDL-C, Total-C Procedure codes: cardiac cath, PCI, CABG, stress test Drug lists: statin and non-statin agents

Electronic Phenotyping

- **Phenotype**
  - A physical attribute: height, weight, blood type, etc
  - A condition: obesity, rheumatoid arthritis, FH …
  - A pattern of events: "Patients with known poor left ventricular function with an ejection fraction of 45%, having community acquired pneumonia with a serum hemoglobin less than …"

- **Phenotyping:** A method for assigning a phenotype label to a patient record

- ***We can view phenotyping as a supervised learning problem.***
  - ***So we need a set of true positive, and a set of true negative cases.***
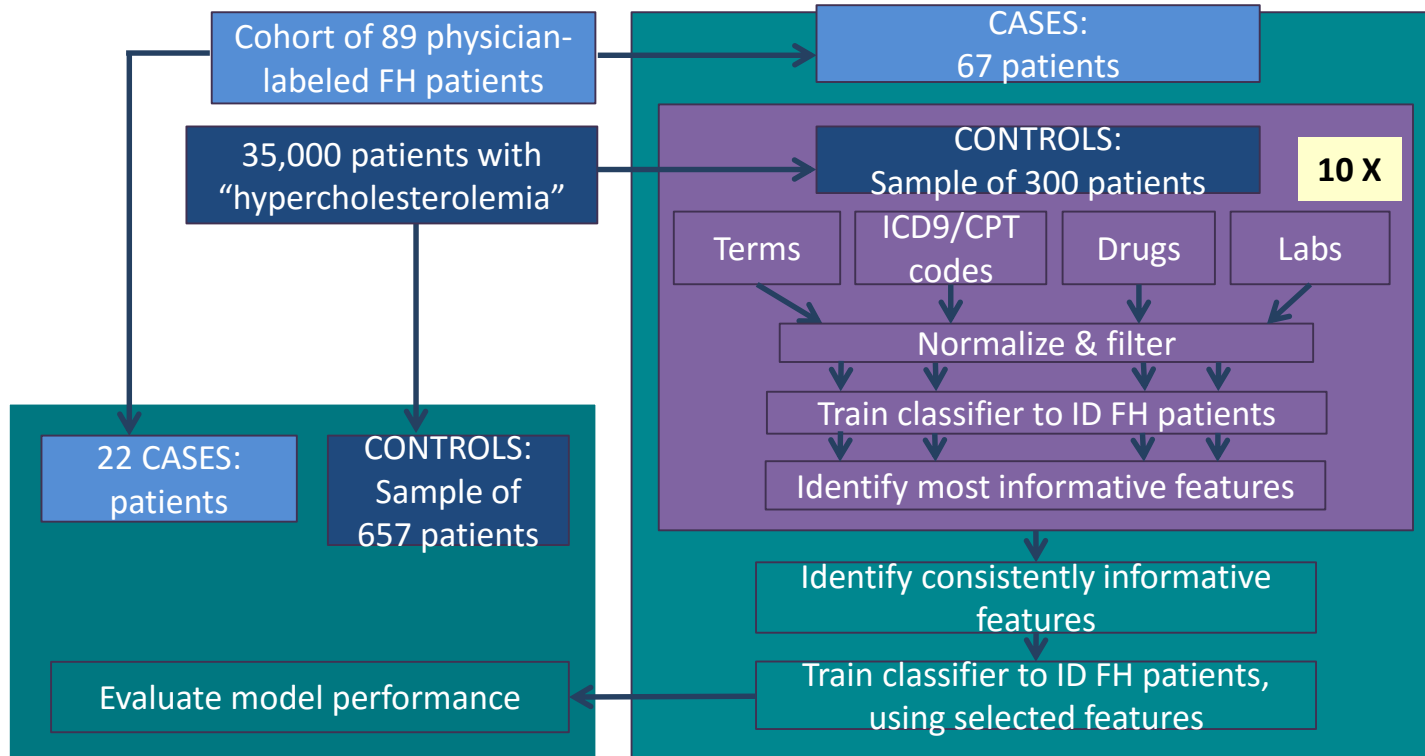
Examples of the Data We Encounter

Patient A

Patient B
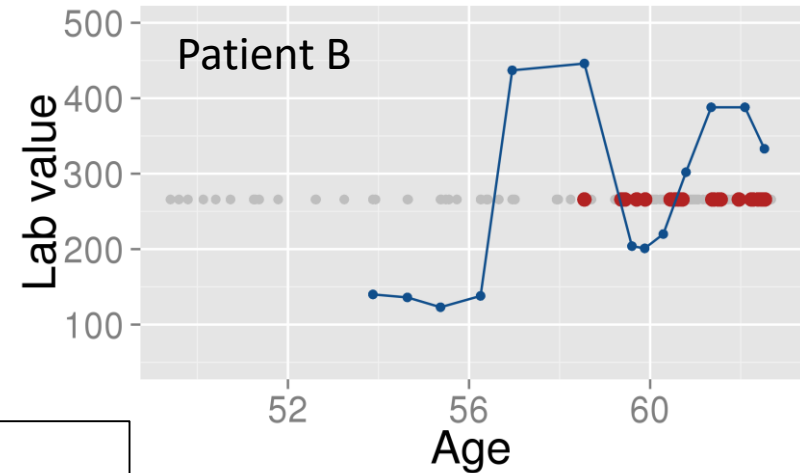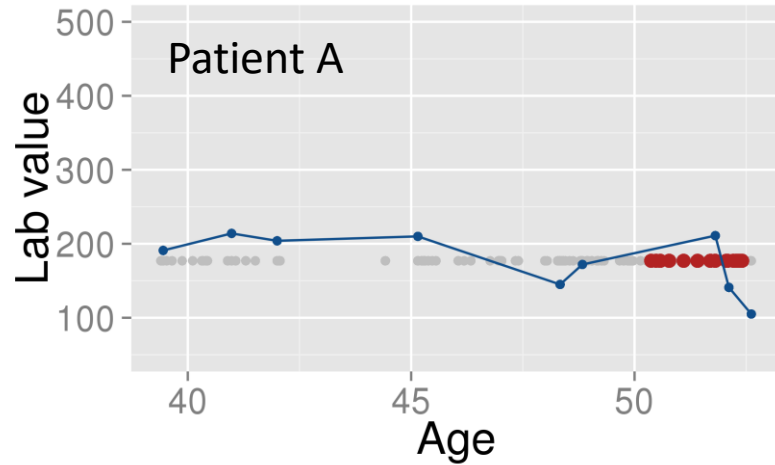
LDL cholesterol levels

Visits during which the common statins are mentioned

Any visit

Feature Engineering

LABS
LAB:HBA1c (High)
LAB:Blood Glucose: High
LAB:Blood Glucose: High
LAB:Blood Glucose: Normal

CODES
ICD9: 250.00
ICD9: 790.2

PRESCRIPTIONS
MED: 6809
MED: 4815

Term to Concept

CONCEPTS
Diabetes mellitus
Diabetes mellitus
Metformin
Metformin.

TERMS
Diabetes
Diabetes nos
Metformin
Dimethylbiguanidine

MR#

Structured and unstructured data from a record is represented as a vector of features

CONCEPT FEATURES:
$f_c =$
#Notes in which the concept occurs at least once

CODE FEATURES:
$f_{co} =$
Counts of a code / Total number of codes

PRESCRIPTION FEATURES:
$f_p =$
Counts of a RxCUI / Total number of RxCUIs

LAB FEATURES:
$f_l =$
Counts of a lab-result / Total number of lab-results

19

Classifier Building

Cohort of 89 physician-labeled FH patients

35,000 patients with "hypercholesterolemia"

CASES:
67 patients

CONTROLS:
Sample of 300 patients

10 X

Terms | ICD9/CPT codes | Drugs | Labs

Normalize & filter

Train classifier to ID FH patients

Identify most informative features

22 CASES:
patients

CONTROLS:
Sample of
657 patients

Identify consistently informative features

Train classifier to ID FH patients, using selected features

Evaluate model performance

Performance Characteristics

N (FH cases)	N (controls)	AUC	PPV	Sensitivity	Specificity	F1 score
89	300 (high cholesterol)	0.91	0.76	0.78	0.92	0.76
89	300 (one comorbidity)	0.92	0.75	0.83	0.93	0.79

QUESTION: DO WE KNOW WHAT WOULD BE THE PERFORMANCE
CHARACTERISTICS just choosing something dumb like LDL levels?

DO we have any metrics on how much the performance is boosted using
orthogonal data?

Universe of Patients to Make Predictions on:

	Stride 6 patients
Criteria 1 - High cholesterol	401,978
Criteria 2 - One comorbidity	379,224
Criteria 3 – Had > 1 visit (total), had a visit w/in 2 years, have "one comorbidity"	249,983

Run classifier

Ranked list of patients

Sanity check by chart review

- What are the clinical characteristics of the patients flagged by the algorithm?
- How well do the flagged patients fit existing diagnostic criteria?
- "Precision at a certain rank"
  - Top 50?
  - Top 100?
  - Top 150?

Iterate

Test in other EHR
- Geisinger, Mayo, Penn
- Assess performance in Gene + vs gene -

Scaling to National Level
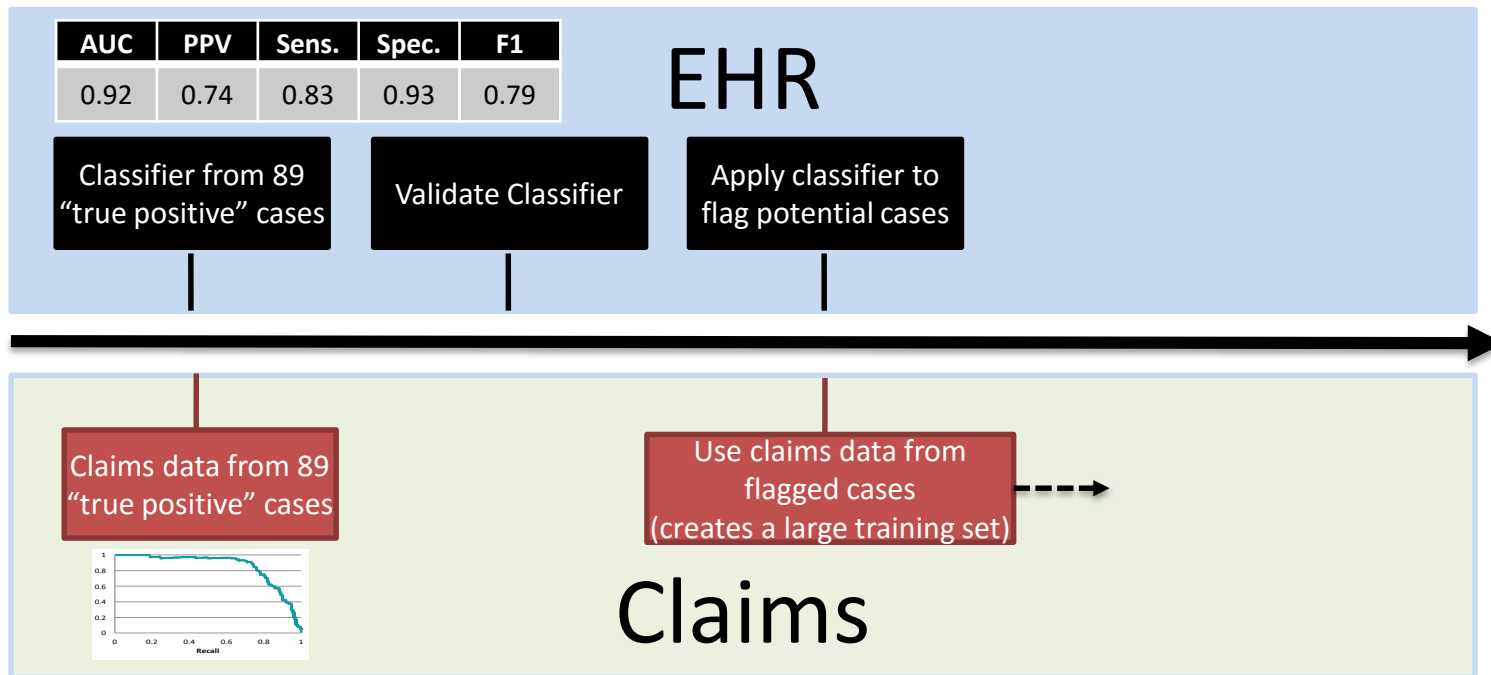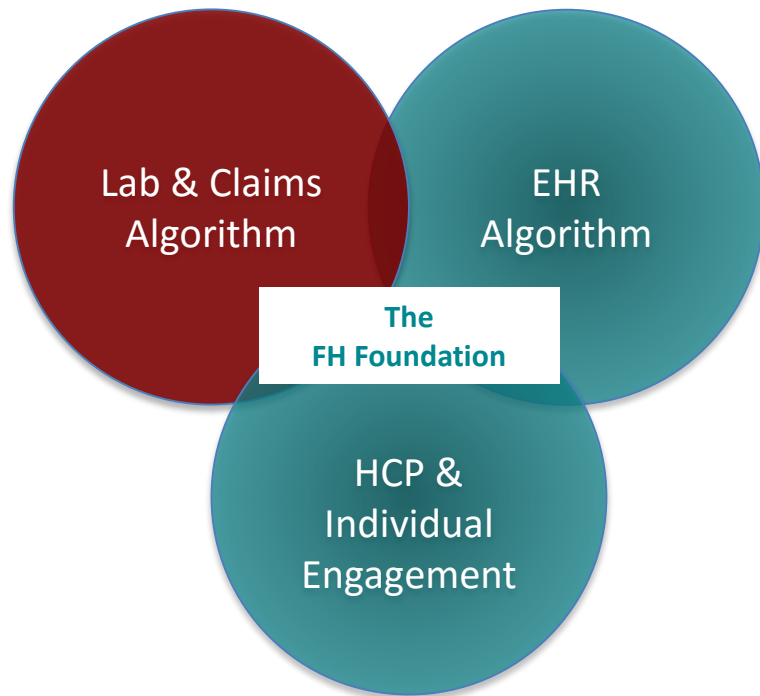
AUC	PPV	Sens.	Spec.	F1
0.92	0.74	0.83	0.93	0.79

EHR

Classifier from 89 "true positive" cases

Validate Classifier

Apply classifier to flag potential cases

Claims data from 89 "true positive" cases

Use claims data from flagged cases (creates a large training set)

Claims

FIND FH®

A multiyear screening and engagement initiative to identify and encourage the diagnosis and treatment of FH

Lab & Claims Algorithm

EHR Algorithm

The FH Foundation

HCP & Individual Engagement

Lab & Claims Data Mining

- Healthcare encounter data on 89 million Americans with cardiovascular disease
- Data from a significant majority of clinical practices

EHR Data Mining

- Comprehensive EHR data from two academic centers
- Expanding to key integrated health systems

HCP & Individual Engagement

- Multichannel tools to engage health systems and individual HCPs
- Tools for clinicians and individuals with FH

FH Foundation
Lab & Claims Database

19,149,553
Unique Patients in Lx Data

19,095,699
Unique Patients in Lx/Rx Overlap

12,861,217
Unique Patients in Lx/Dx/Px Overlap

12,834,404
Unique Patients in Lx/Rx/Dx/Px Overlap

40,328,108
Unique Patients in Rx/Dx/Px Overlap

40,422,524
Unique Patients in Dx/Px Data

89,112,339
Unique Patients in Rx Data

Ensemble Model

- Multiple models working together
- We build a Random Forest classifier for each class of features
- Output of each model becomes a feature into the ensemble classifier
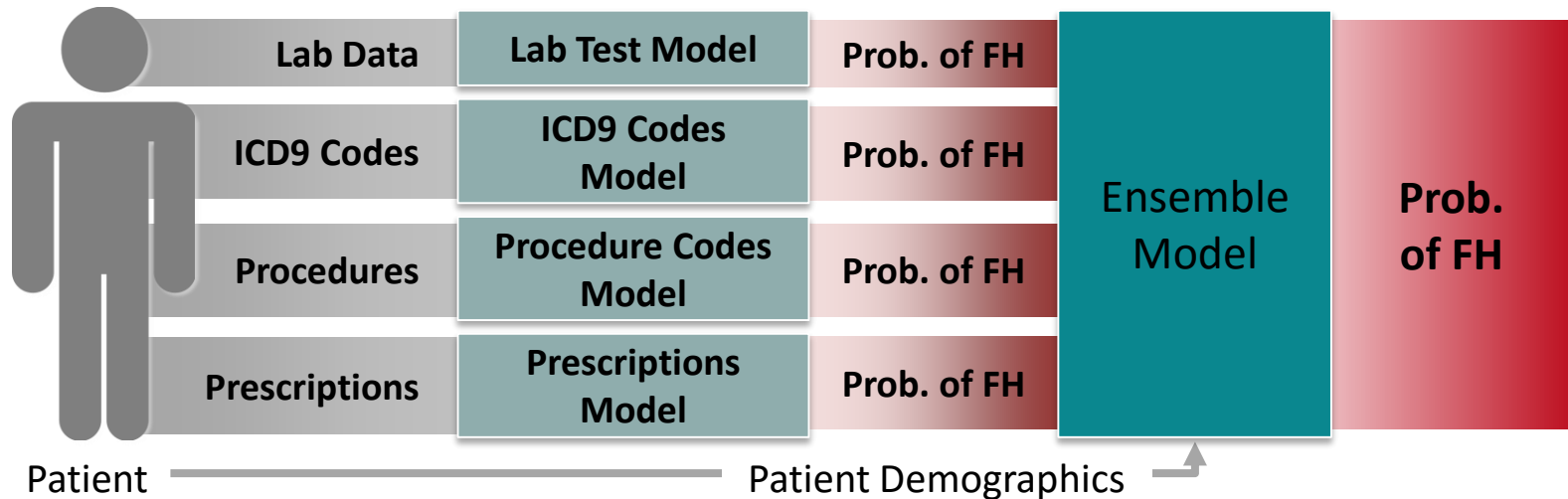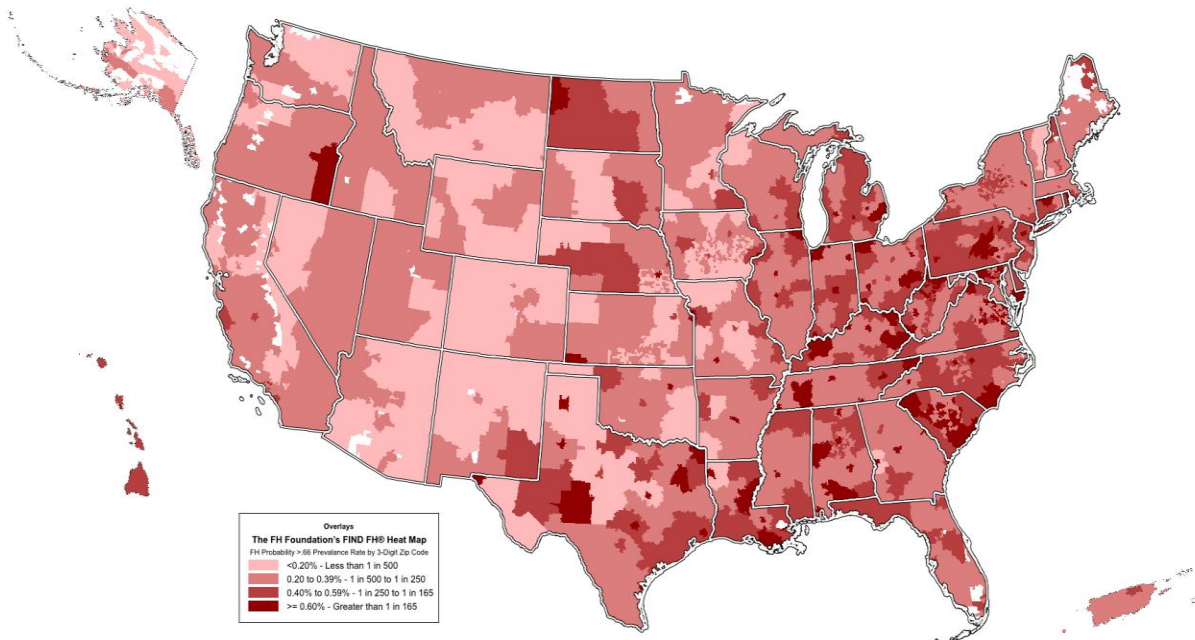
The FH Foundation's FIND FH® Heat Map
FH Probability > .86 Prevalence Rate by 3-Digit Zip Code

<0.20% - Less than 1 in 500
0.20 to 0.39% - 1 in 500 to 1 in 250
0.40% to 0.59% - 1 in 250 to 1 in 165
>= 0.60% - Greater than 1 in 165

FIND FH® Lab & Claims Algorithm Developed by The FH Foundation
Claims Data Source: IMS Health Real World Data: LRx longitudinal prescriptions and Dx medical claims

Patients w/Dx
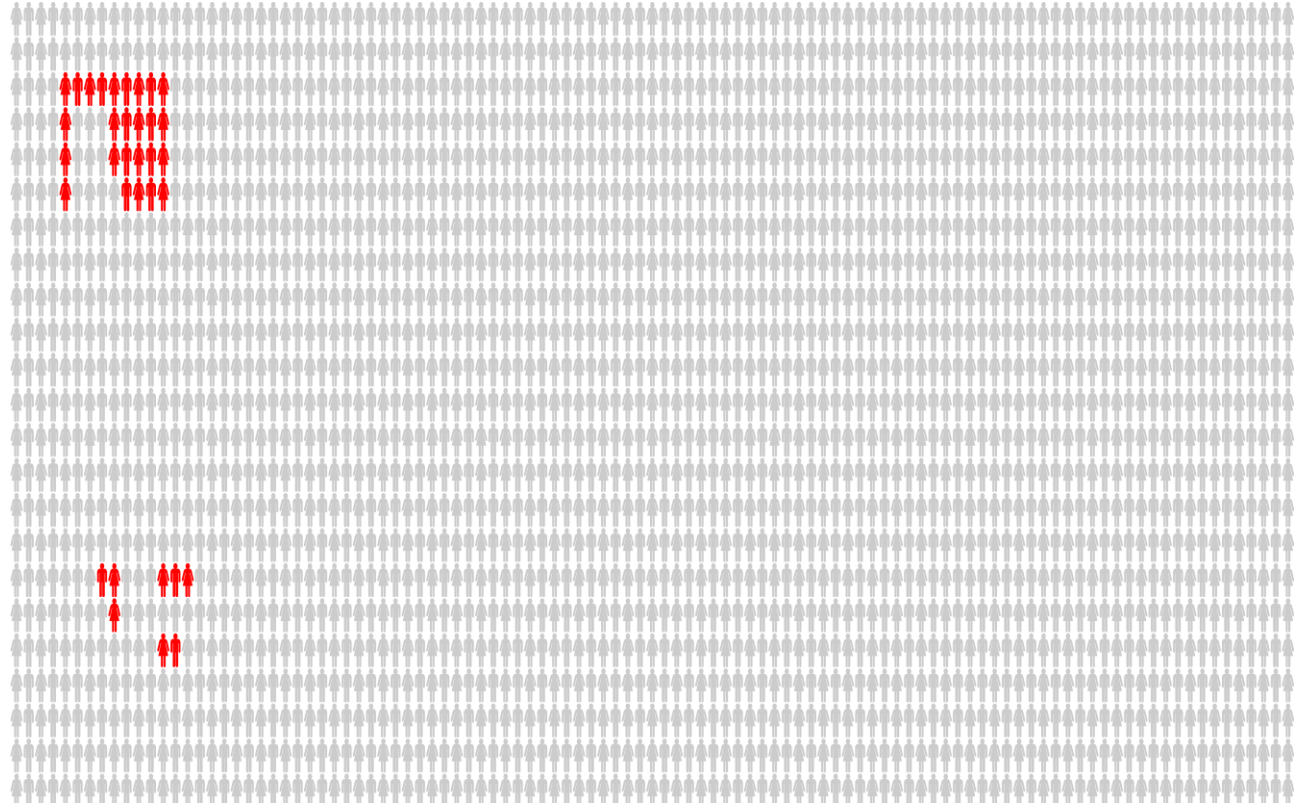
Patients w/out Dx

Striking a Balance Between Precision & Recall
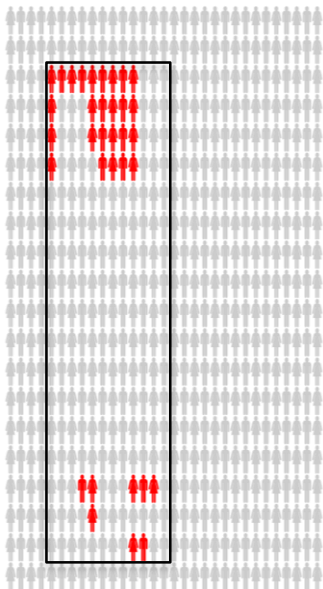
Perfect Recall; Low Precision
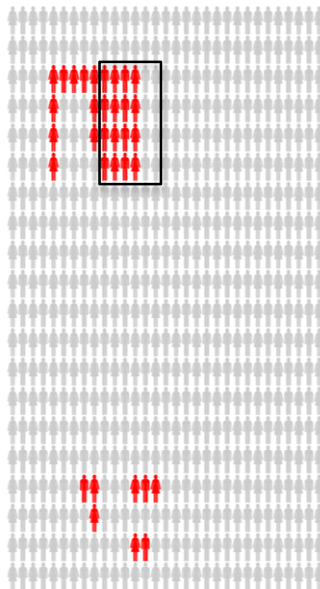
Balanced Recall and Precision

Low Recall; Perfect Precision
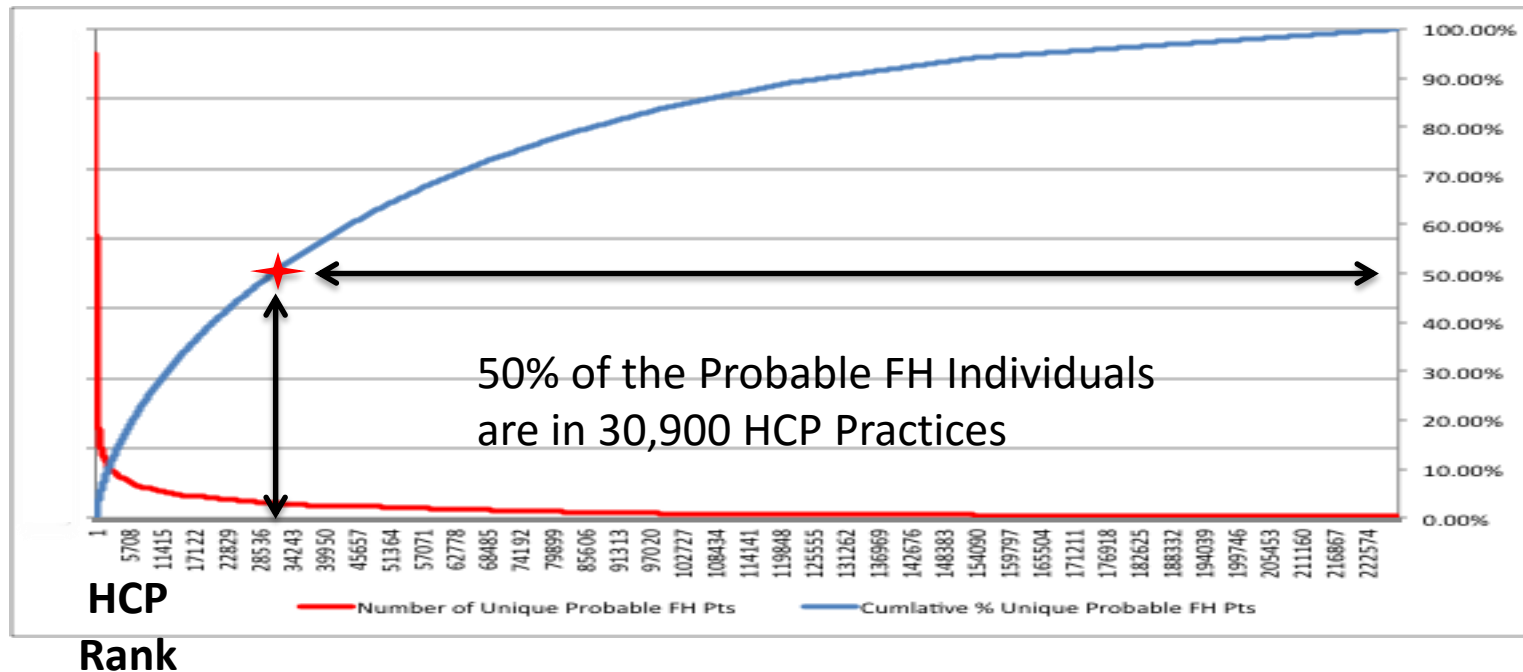
Patients w/Dx   Patients w/out Dx

75% of the Top HCPs Are PCPs

Hcpid	Specialty	City	State	Zip3	Unique Tot CVD Pts	Unique Probable FH > 0.9
1984181	FAMILY MEDICINE	EVANSVILLE	KY	424	246	124
400657	INTERNAL MEDICINE	FAYETTEVILLE	NC	283	810	107
486555	GENERAL PRACTICE	VAN NUYS	CA	914	822	101
2058521	INTERNAL MEDICINE	NEWARK	NJ	070	930	88
957011	GASTROENTEROLOGY	PITTSBURGH	PA	152	160	86
1954569	FAMILY MEDICINE	FT WORTH	TX	760	976	82
1962562	FAMILY MEDICINE	SOUTH FLORIDA	FL	330	762	81
321530	FAMILY MEDICINE	PALATINE	IL	600	168	81
2131272	CARDIOVASCULAR DISEASES	HOUSTON	TX	770	1,401	76
1761176	NURSE PRACTITIONER	HACKENSACK	NJ	076	845	75
1492128	FAMILY MEDICINE	EUREKA	CA	955	172	68
1880622	INTERNAL MEDICINE	TUSCALOOSA	AL	354	541	68
436410	PHYSICIAN ASSISTANT	BUFFALO	NY	142	505	67
1390586	FAMILY MEDICINE	HUNTSVILLE	AL	356	577	66
1886494	FAMILY MEDICINE	SHREVEPORT	LA	711	1,016	64
2033271	FAMILY MEDICINE	SALISBURY	MD	218	166	63
24358	INTERNAL MEDICINE	BIRMINGHAM	AL	352	1,115	62
351493	ENDOCRINOLOGY & METAB..	BOWLING GREEN	KY	421	517	62
866261	CARDIOVASCULAR DISEASES	INGLEWOOD	CA	902	765	61
1123362	INTERNAL MEDICINE	HONOLULU	HI	967	414	59
1349873	FAMILY MEDICINE	HONOLULU	HI	967	578	59
1062287	GENERAL PRACTICE	SN BERNARDINO	CA	925	767	58
232373	FAMILY MEDICINE	CHARLOTTE	NC	282	1,075	57
207773	FAMILY MEDICINE	HACKENSACK	NJ	076	446	55
691682	INTERNAL MEDICINE	KNOXVILLE	TN	377	493	55
960935	GASTROENTEROLOGY	KNOXVILLE	TN	378	336	55

Harnessing the Power Law

50% of the Probable FH Individuals are in 30,900 HCP Practices

Number of Unique Probable FH Individuals per HCP

HCP Rank

Number of Unique Probable FH Pts — Cumlative % Unique Probable FH Pts

FIND FH 2016

- Two Additional Clinical Partnerships
  - Geisinger Health System (GHS)
  - Ohio State University Wexner Medical Center (OSU)
- FIND FH Algorithm Clinical Validation
- KOL Network Engagement Campaign
- Manuscript
  - Submit at Least 1 Manuscript

FiveThirtyEight

Politics   Sports   **Science & Health**   Economics   Culture

PUBLIC HEALTH | 4:01 PM | JAN 13, 2016

An Algorithm Could Know You Have Genetic Disease Before You Do

BY ANNA MARIA BARRY-JESTER

FIND FH® Lab & Claims Algorithm Developed by The FH Foundation
Claims Data Source: IMS Health Real World Data: LRx longitudinal prescriptions and Dx medical claims

CMHC
WEST
CARDIOMETABOLIC HEALTH CONGRESS

March 4-5, 2016 • San Francisco, CA