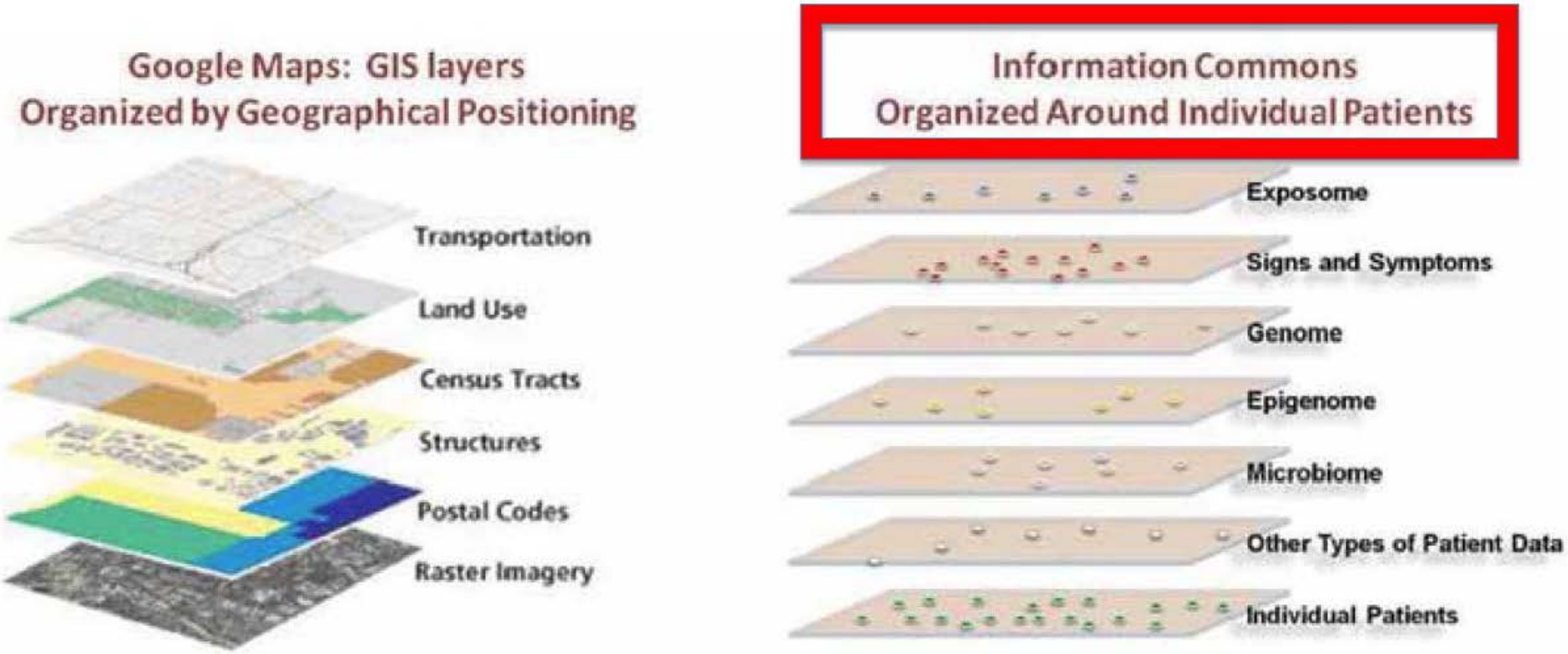
Avoiding Fake Diseases and Biomarkup In Pediatrics

Isaac S. Kohane, MD, PhD

HARVARD MEDICAL SCHOOL
DEPARTMENT OF Biomedical Informatics



Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease

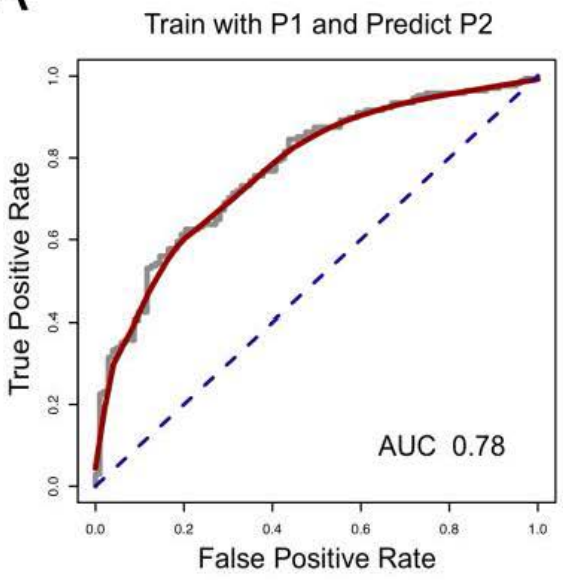
Report from National academy of science, USA, 2011

Some of the points I will make

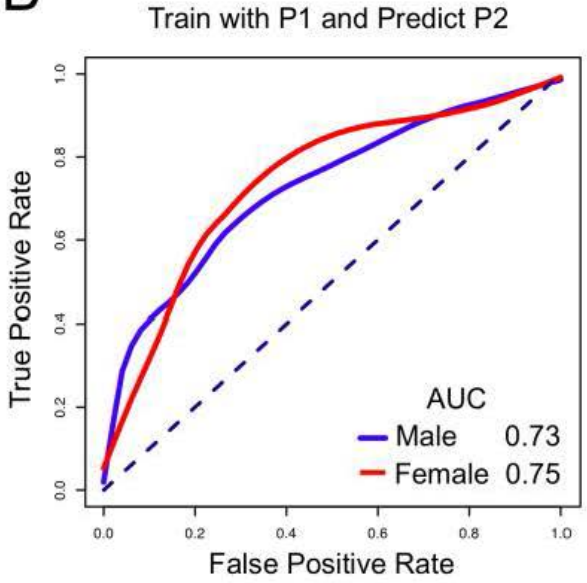
- Most population genetics and clinical medicine addicted to categorical diagnosis
- Fake/Wrongly taxonomized diseases makes social & commercial construction of disease likely.
- Genomics-first/Genomics only unnecessarily limiting
- Clinical predictive accuracy does not imply shared physiology
- Disease overlaps demonstrates a fundamental problem
- Clinical data can be used to lesson the confusion.
- But multi-modal approaches are the most robust.

Blood Gene Expression Detection

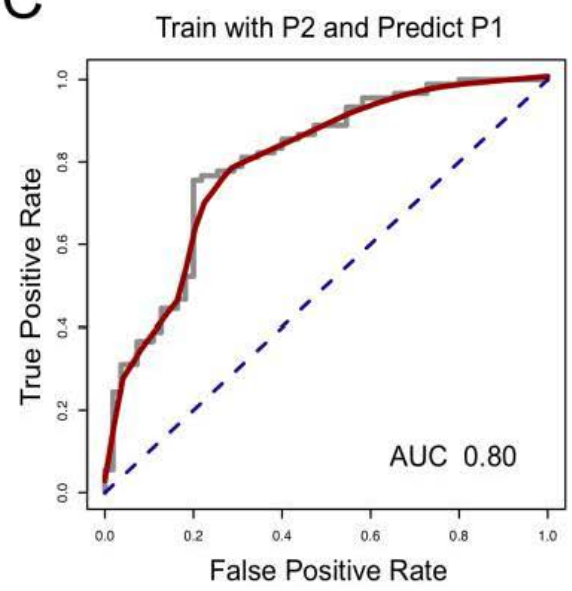
A



B

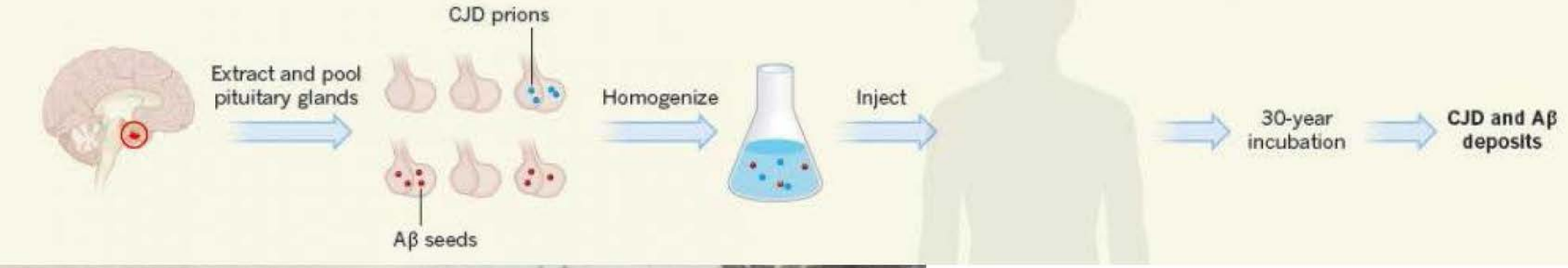


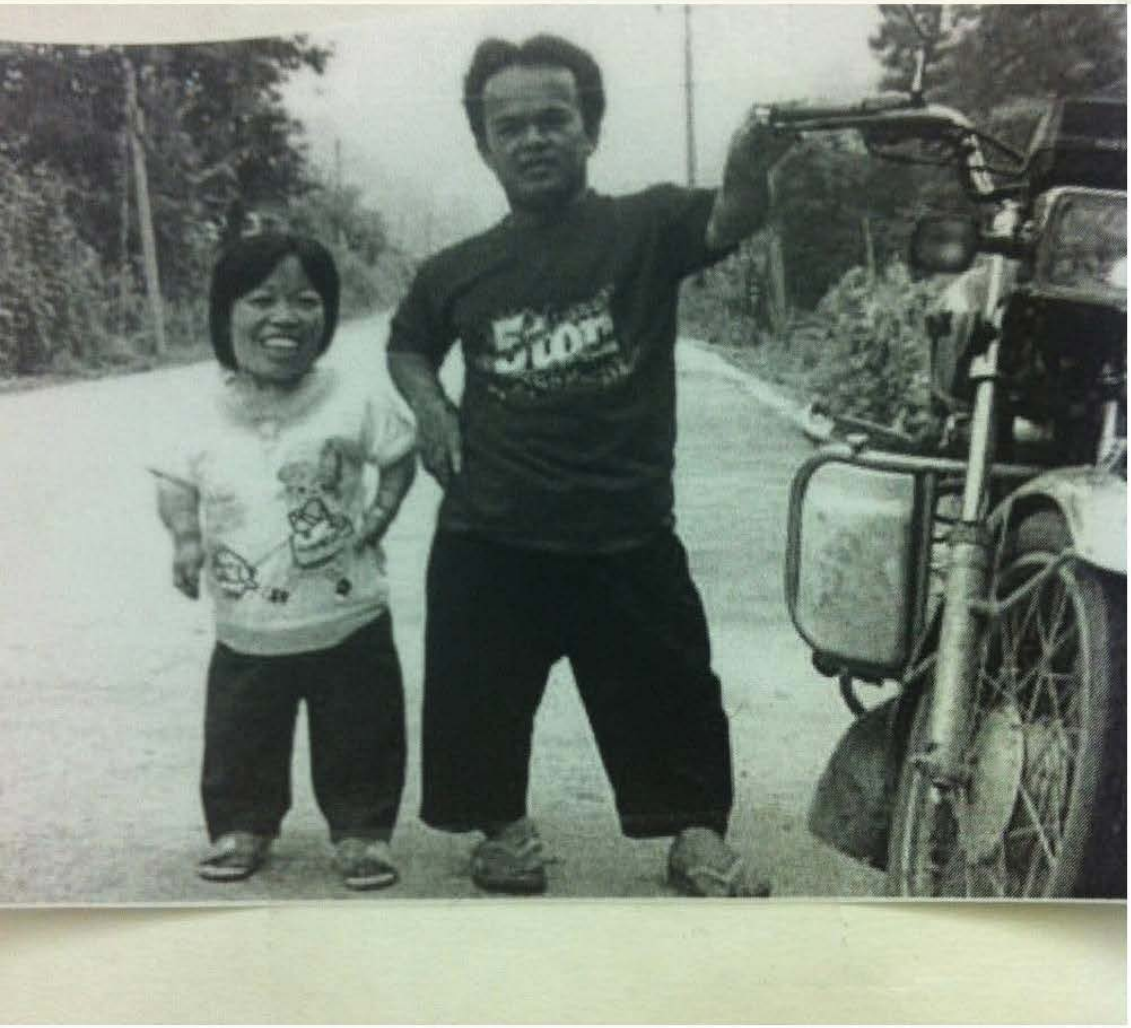
C

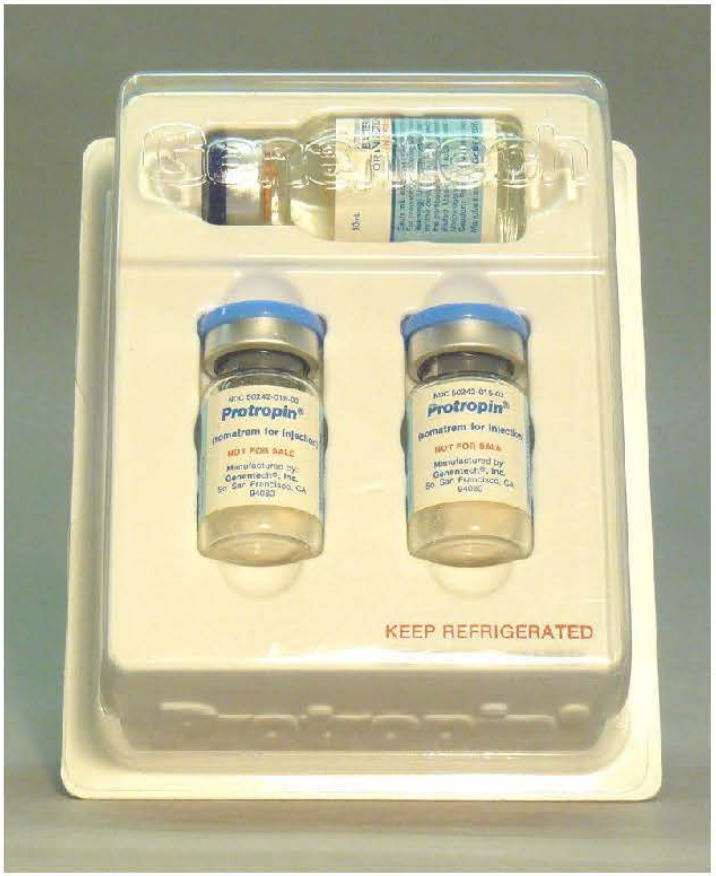


Invited to HLS Meeting

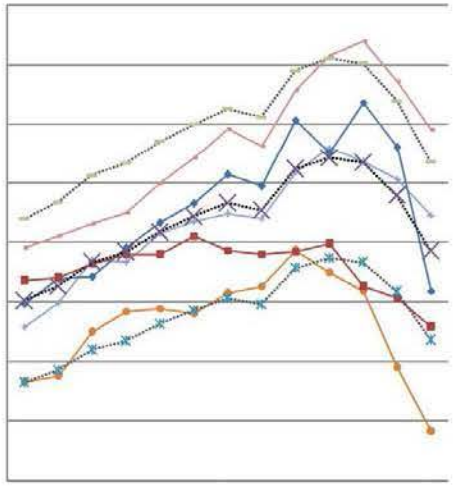
"I think when Ari [Ne'eman] talks about autism and I talk about autism, we're talking about people with different clusters of autism. I know he doesn't like the word 'cure.' If my daughter could function the way Ari could, I would consider her cured," says Singer. "I have to believe my daughter doesn't want to be spending time peeling skin off her arm."







HARVARD MEDICAL SCHOOL | DEPARTMENT OF Biomedical Informatics



GIANT study

- 100's of genes
- Not like diseases at tails
- So what do you call short stature and is that a diagnosis?

Criteria for Treatment

- "Growth hormone deficiency (GHD)"
- "Idiopathic short stature (ISS), defined by height standard deviation score ≤-2.25"
  associated with growth rates unlikely to result in normal adult height, in whom other causes of short stature have been excluded

and a little story from
25 years ago

But about once a month, an ambitious parent with cash to burn asks Desrosiers, a pediatric endrocrinologist in Florida, if he would be willing to give growth hormones to a short but otherwise healthy child.

Desrosiers turns these patients away, but he says the requests still come.

null

Just last week, the father of a young baseball player -- a 14-year-old who was already 5 feet 6 inches tall -- expected Desrosiers to prescribe recombinant growth hormone (rGH) to add height to his budding athlete.

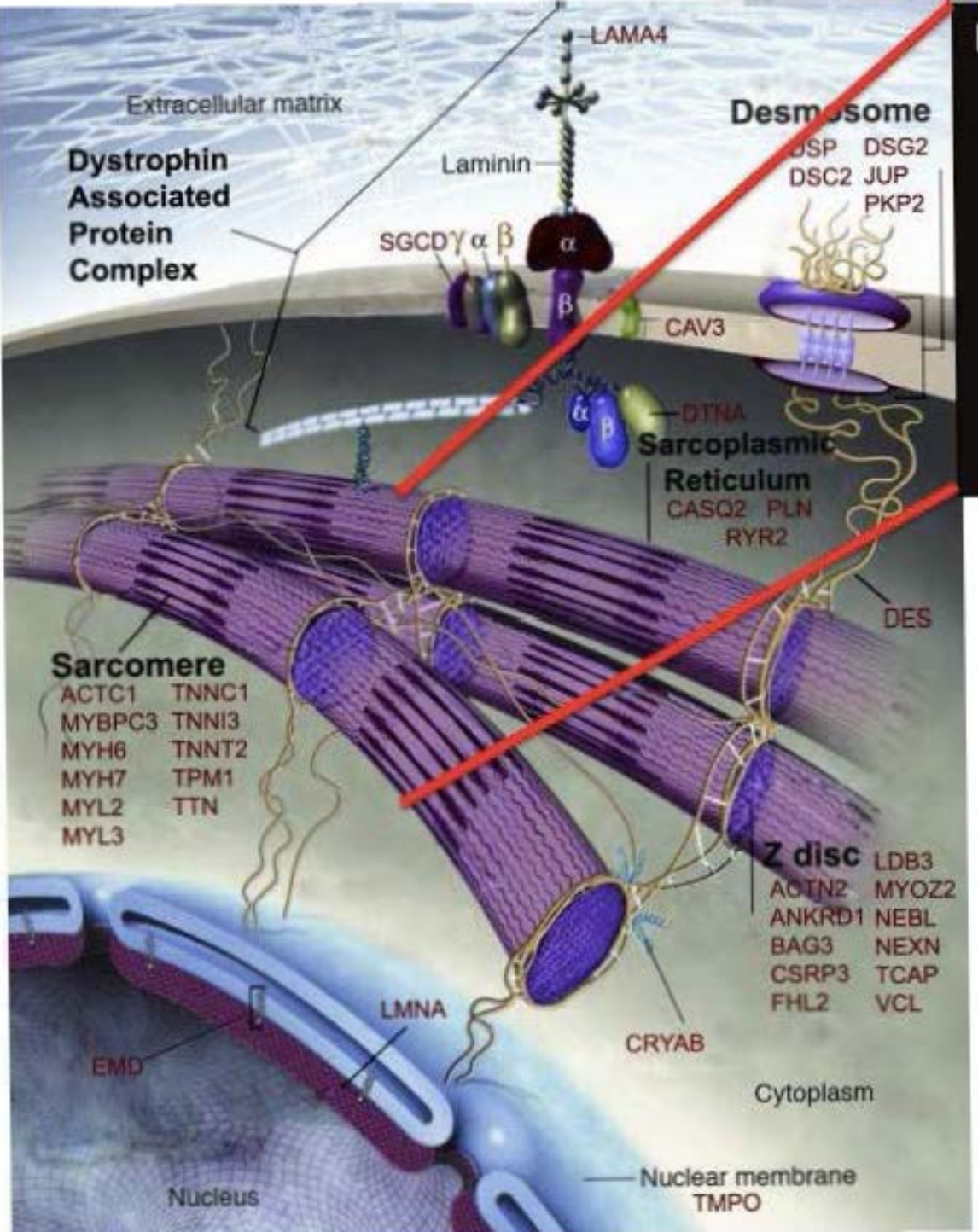
"He wanted to make his kid big, and he thinks he's going to walk out with the shots," said Desrosiers, director of pediatric endocrinology at Arnold Palmer Children's Hospital in Orlando. "He was willing to pay more than $45,000 a year, and didn't even bat an eyelash."

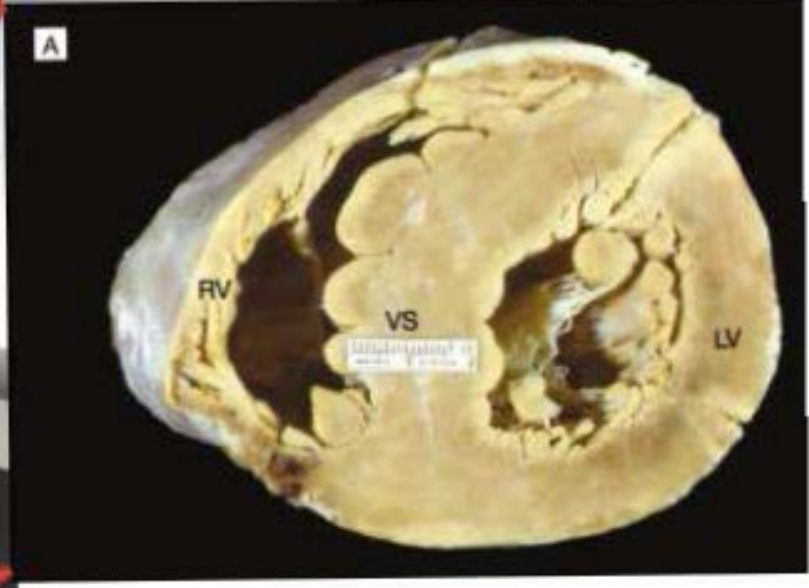
Desrosiers said he even gets requests for growth hormone from "Jolly Green Giant" families, where children are likely to be tall.

In 2003, the U.S. Food and Drug Administration approved the use of rGH for children with idiopathic -- or unexplained -- short stature, without a diagnosed metabolic hormone deficiency.

Pathological Interaction Between Clinical Annotation and Genetics.

HARVARD MEDICAL SCHOOL
DEPARTMENT OF Biomedical Informatics





Hypertrophic Cardiomyopathy (HCM)

- Heart failure
- Arrhythmias
- Obstructed blood flow
- Infective endocarditis
- Sudden cardiac death

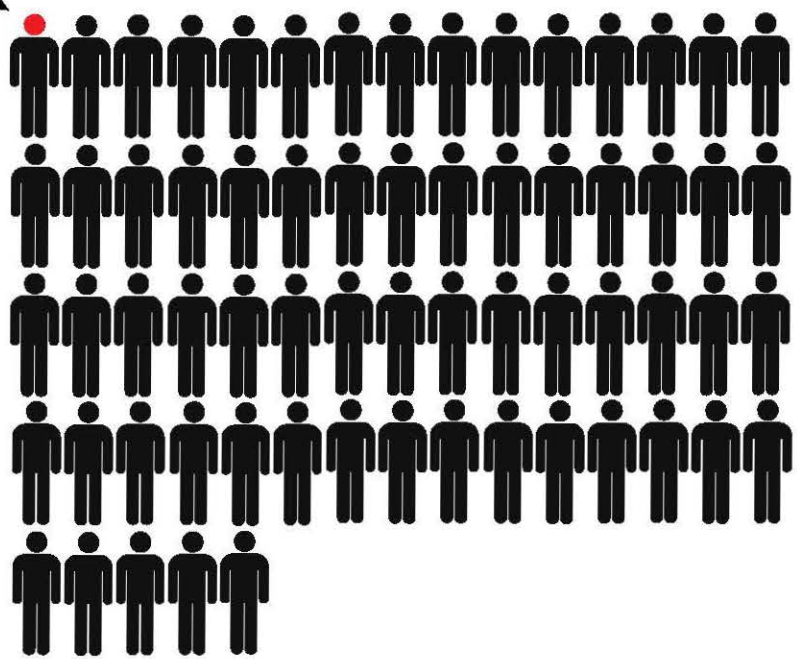
Prevalence 1:500

Autosomal Dominant

HARVARD MEDICAL SCHOOL
DEPARTMENT OF Biomedical Informatics

6503 individuals

NHLBI ESP

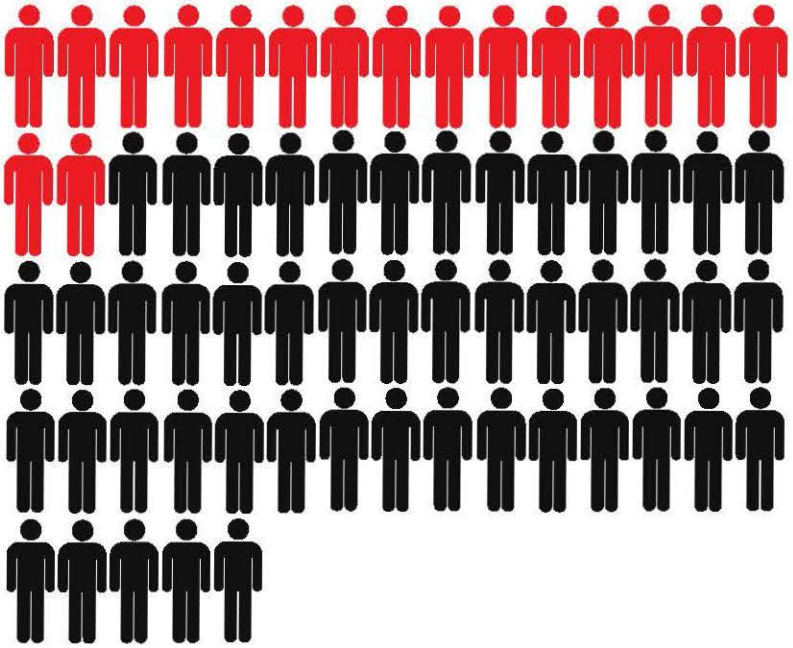


expected

100 individuals

6503 individuals

NHLBI ESP

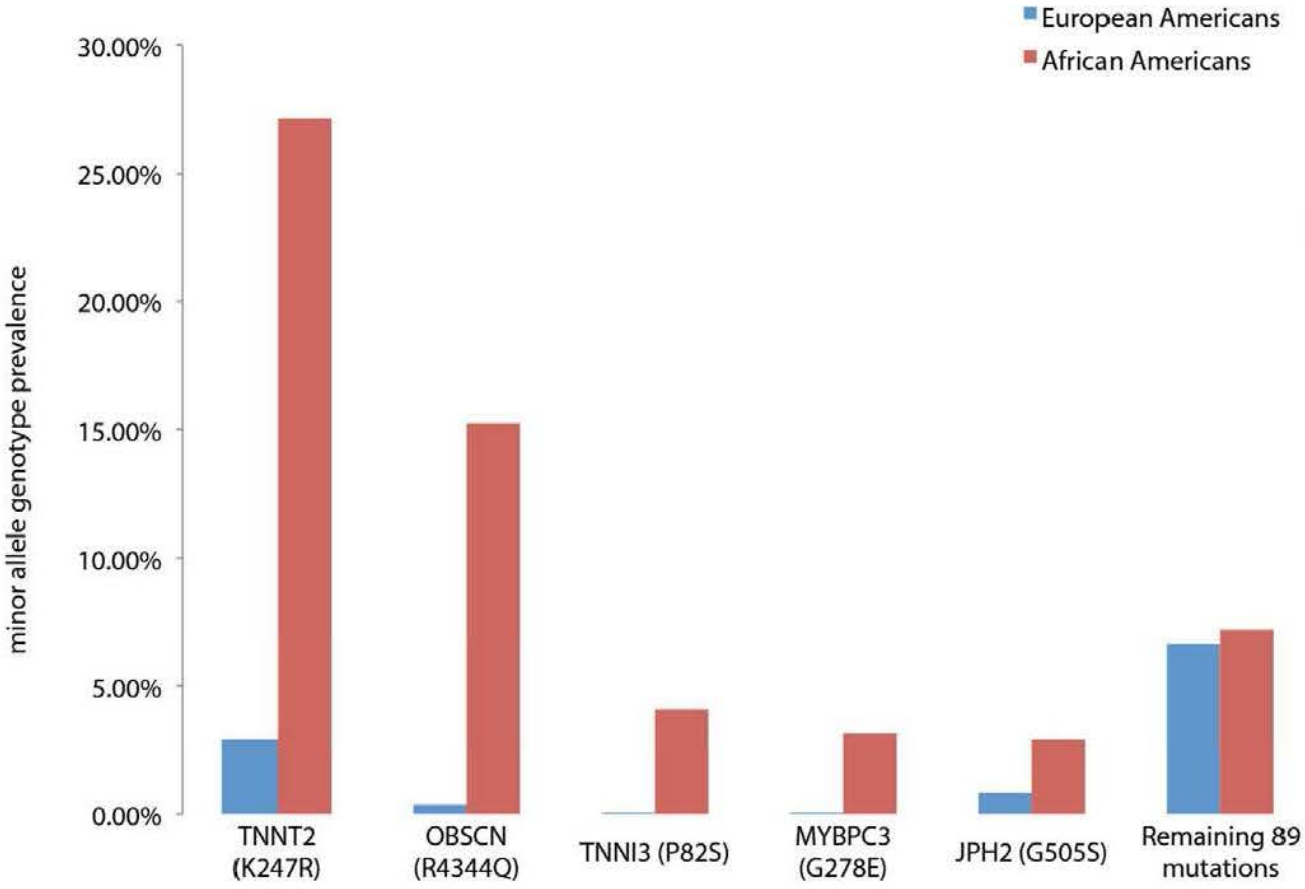


observed

HCM Prevalence = 1:500

HCM Inheritance = Autosomal Dominant

Genetics-induced Health Disparities



	TNNT2 (K247R)	OBSCN (R4344Q)	TNNI3 (P82S)	MYBPC3 (G278E)	JPH2 (G505S)	Remaining 89 mutations
European Americans	2.88%	0.33%	0.03%	0.02%	0.80%	6.67%
African Americans	27.14%	15.27%	4.07%	3.15%	2.92%	7.18%

Potential

JGUST 18, 2016

Age	Ethnicity	Report Year	Originally Reported Status	Current Status	Indication for Test	
46	Unavailable	2005	P	B	Clinical Diagnosis of HCM	Pro82Ser
75	Unavailable	2005	P	B	Family History and Clinical Symptoms of HCM	
32	Black or African American	2005	P	B	Clinical Diagnosis of HCM	
34	Black or African American	2005	U	B	Clinical Diagnosis and Family History of HCM	
12	Black or African American	2006	U	B	Family History of HCM	
40	Black or African American	2007	U	B	Clinical Diagnosis of HCM	
45	Black or African American	2007	U	B	Clinical Features of HCM	
16	Asian	2008	U	B	Clinical Diagnosis and Family History of HCM	
59	Black or African American	2006	P	B	Clinical Features of HCM	Gly278Glu
15	Black or African American	2007	P	B	Clinical Diagnosis of HCM	
16	Black or African American	2007	P	B	Clinical Diagnosis of HCM	
22	Black or African American	2007	P	B	Clinical Diagnosis and Family History of HCM	
48	Black or African American	2008	U	B	Clinical Diagnosis of HCM	

P = Pathogenic and Presumed Pathogenic
U = Pathogenicity Debated and Unknown Significance

Valsartan for Attenuating Disease Evolution In Early Sarcomeric HCM (VANISH)

This study is currently recruiting participants. (see Contacts and Locations)

Verified July 2013 by New England Research Institutes

Sponsor:
New England Research Institutes

Collaborator:
National Heart, Lung, and Blood Institute (NHLBI)

Information provided by (Responsible Party):
New England Research Institutes

ClinicalTrials.gov Identifier:
NCT01912534

First received: June 5, 2013
Last updated: October 5, 2015
Last verified: July 2013
History of Changes

Full Text View | Tabular View | No Study Results Posted | Disclaimer | How to Read a Study Record

Group 1 (Overt HCM Cohort)

1. LV wall thickness ≥12 mm and ≤25 mm or z score ≥3 and ≤18 as determined by rapid assessment by the echocardiographic core laboratory
2. NYHA functional class I or II; no perceived or only slight limitations in physical activities
3. No resting or provokable LV obstruction (peak gradient ≤ 30 mmHg) on clinically-obtained Exercise Tolerance Test (ETT)-echo within the past 24 months or transthoracic echo with Valsalva maneuver within the past 12 months
4. Age 8-45 years
5. Able to attend follow-up appointments, complete all study assessments, and provide written informed consent

Group 2 (Preclinical HCM Cohort (G+/LVH-))

1. LV Wall Thickness <12 mm and z score <3 , as determined by rapid assessment by the echocardiographic core laboratory
2. Age 10-25 years
3. E' z score ≤ -1.5 OR ECG abnormalities other than NSSTW changes (Q waves, T wave inversion, repolarization changes) OR LV wall thickness z-score 1.5-2.9 combined with LV thickness to dimension ratio ≥0.19 (as determined by rapid assessment by the echocardiographic core laboratory)
4. Able to attend follow-up appointments, complete all study assessments, and provide written informed consent

Is Prediction the Acid Test of Diagnosis?

Survival 3 Years After a WBC Test

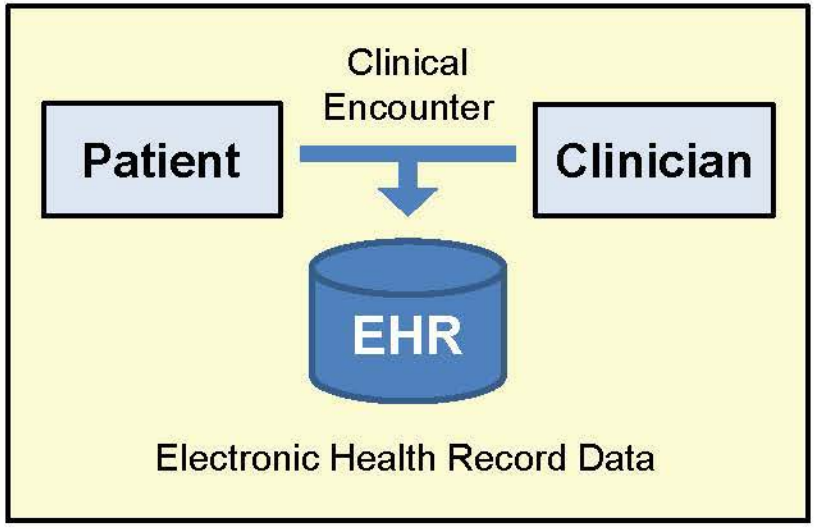
(White, Male, 50-69 Years; Using Last WBC Between 7/28/05 and 7/27/06)

Repeat Interval	Result Time	WBC Value				Patients
		Low	Normal	High	Any	
< 1 Day	12a-8a	43.33%	84.68%	63.24%	76.39%	1830
	8a-4p	54.55%	86.61%	79.40%	83.15%	1442
	4p-12a		77.30%	67.53%	72.49%	229
< 1 Year	12a-8a	47.83%	79.58%	66.67%	74.39%	1644
	8a-4p	76.96%	90.73%	80.80%	88.53%	8812
	4p-12a	81.65%	92.99%	86.01%	91.69%	2769
> 1 Year	12a-8a		95.65%	96.97%	96.00%	175
	8a-4p	97.30%	98.13%	91.98%	97.83%	4280
	4p-12a	92.68%	97.35%	96.67%	97.20%	1932
Any		73.17%	91.79%	78.11%	88.95%	23113
Patients		1286	18775	3052	23113	

Healthcare System Dynamics

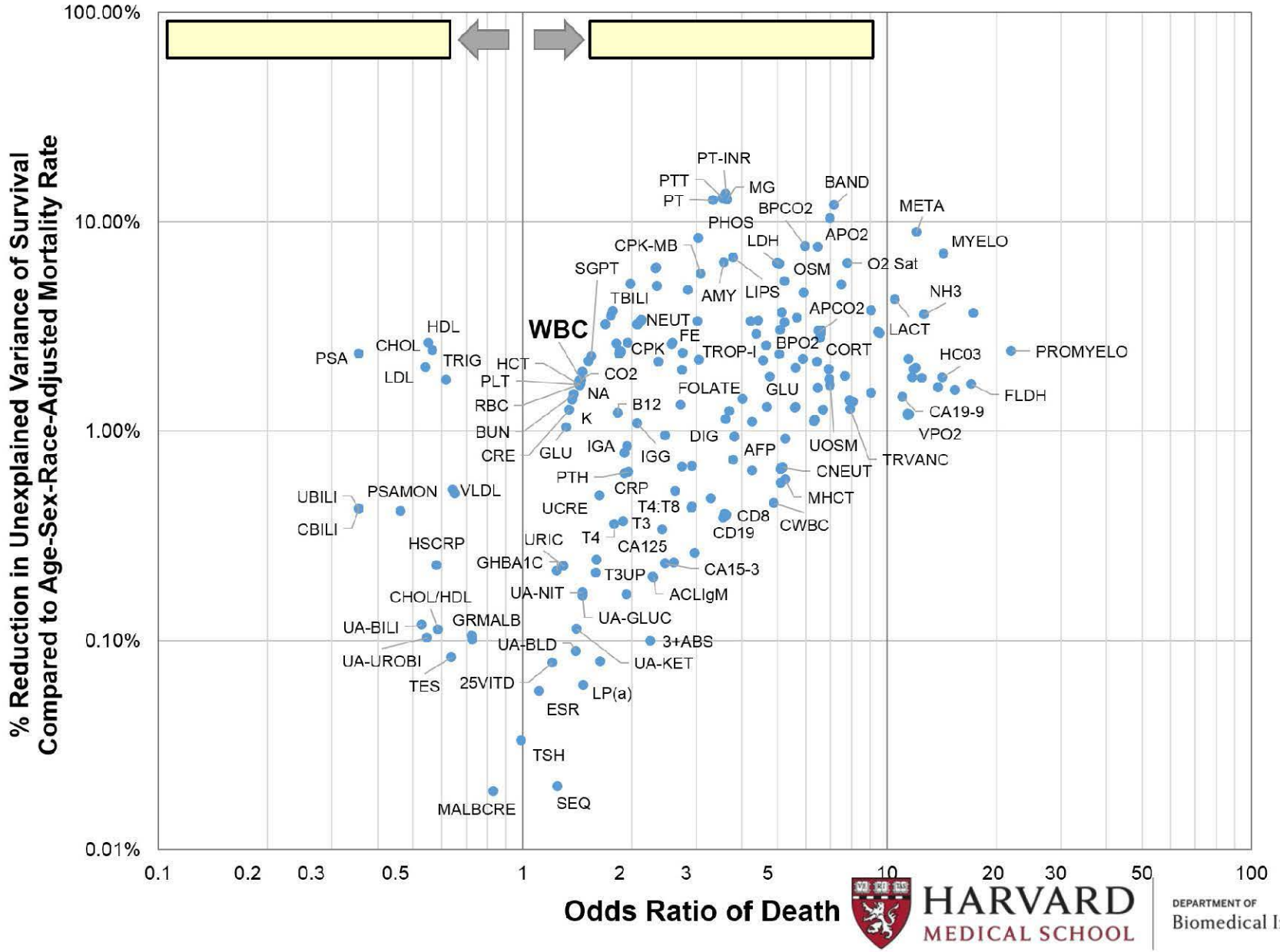
Clinical data reflect both patients' health AND their interactions with the healthcare system.

Patient Pathophysiology	Healthcare System Dynamics	Data Quality
Patient Demographics	Number of Observations	Data Entry Errors
Diagnoses	Time of Day of Observations	Dictation Mistakes
Laboratory Test Results	Time Between Observations	Data Compression Loss
Vital Signs	Cost of a Test or Treatment	Unstructured Data
Genetic Markers	Clinical Setting / Clinician Type	Missing Data



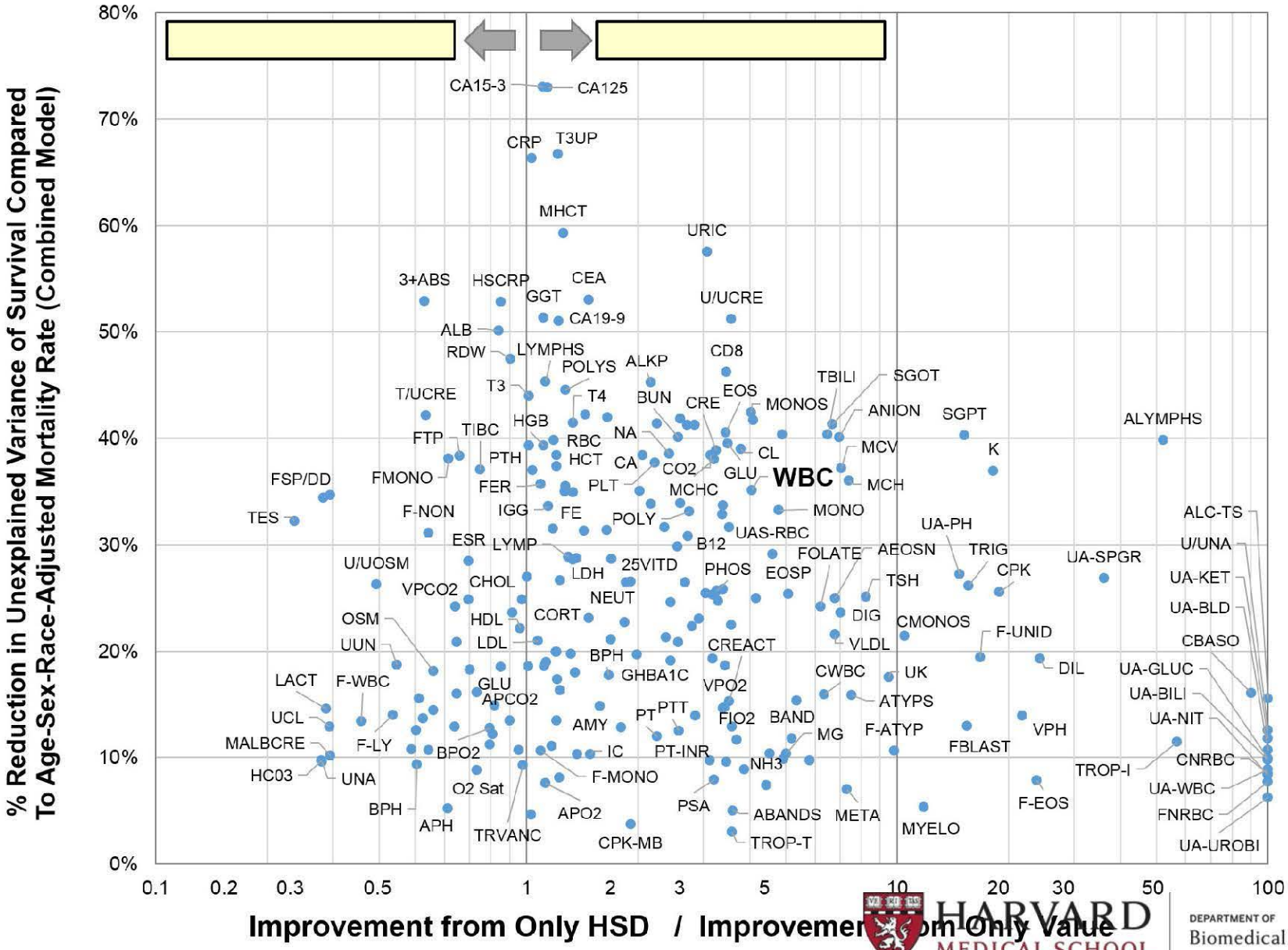
	Patient Pathophysiology	
Healthcare System Dynamics	**Normal**	**Abnormal**
Normal	Best Outcomes	Moderate Outcomes
Abnormal	Moderate Outcomes	Worst Outcomes

Predicting Survival from **Ordering** a Lab Test



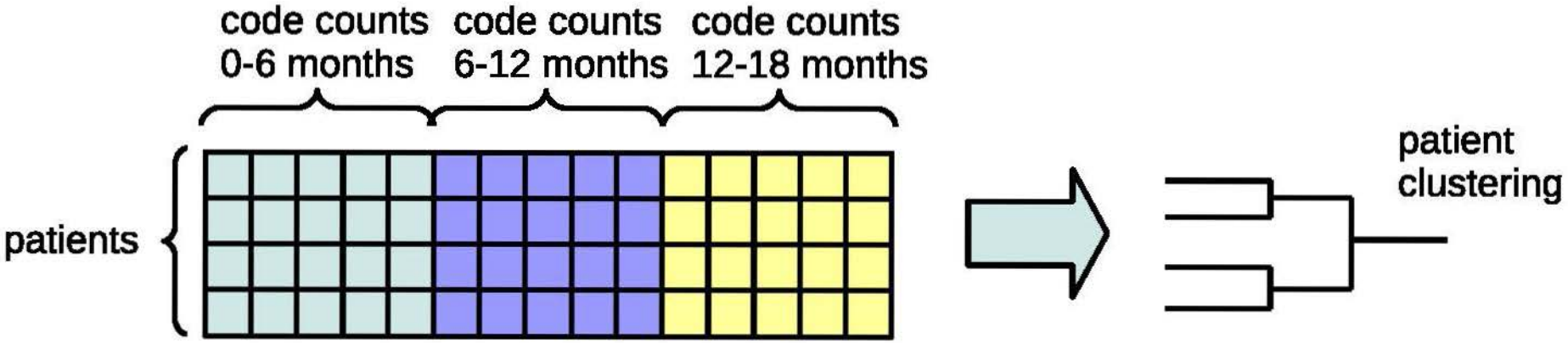
HARVARD MEDICAL SCHOOL
DEPARTMENT OF Biomedical Informatics

Predicting Survival Using **Lab Value** & **HSD**

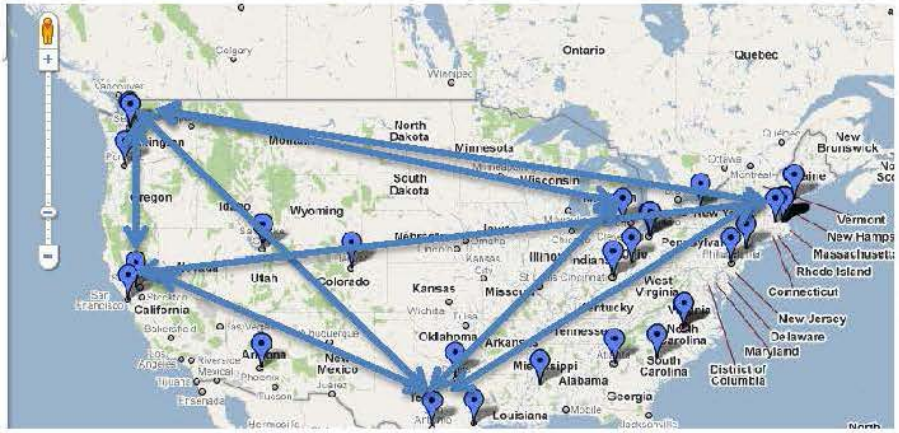


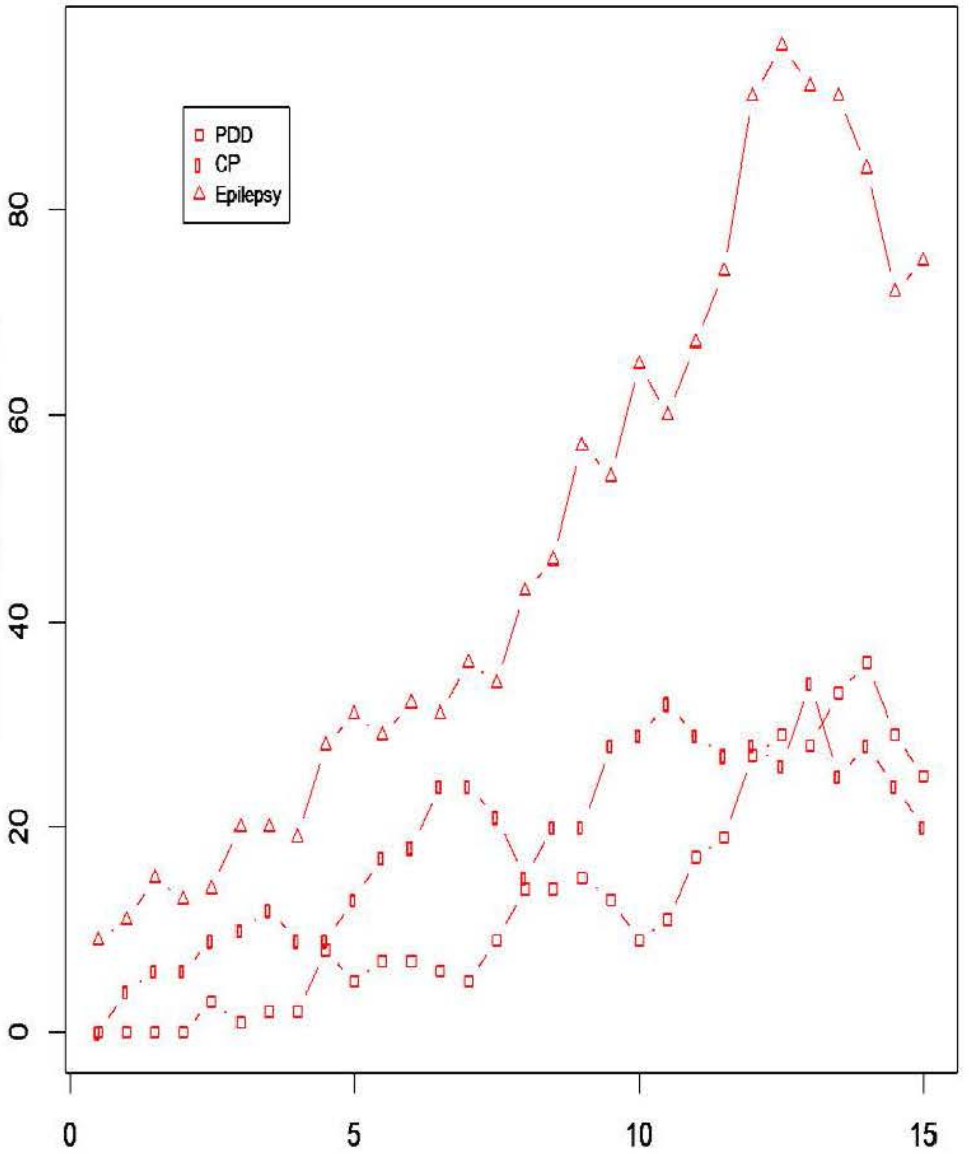
Clinical Data To Clarify Diagnostic Boundaries?

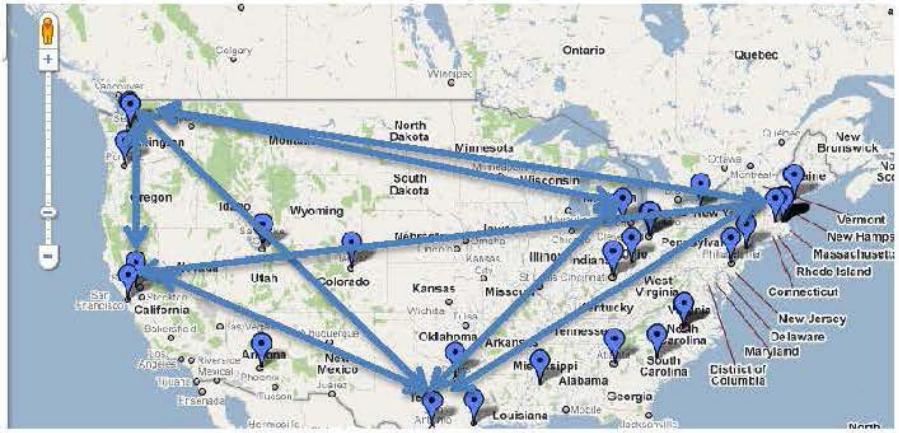
Mothers told me about bowel problems but pediatricians told me…

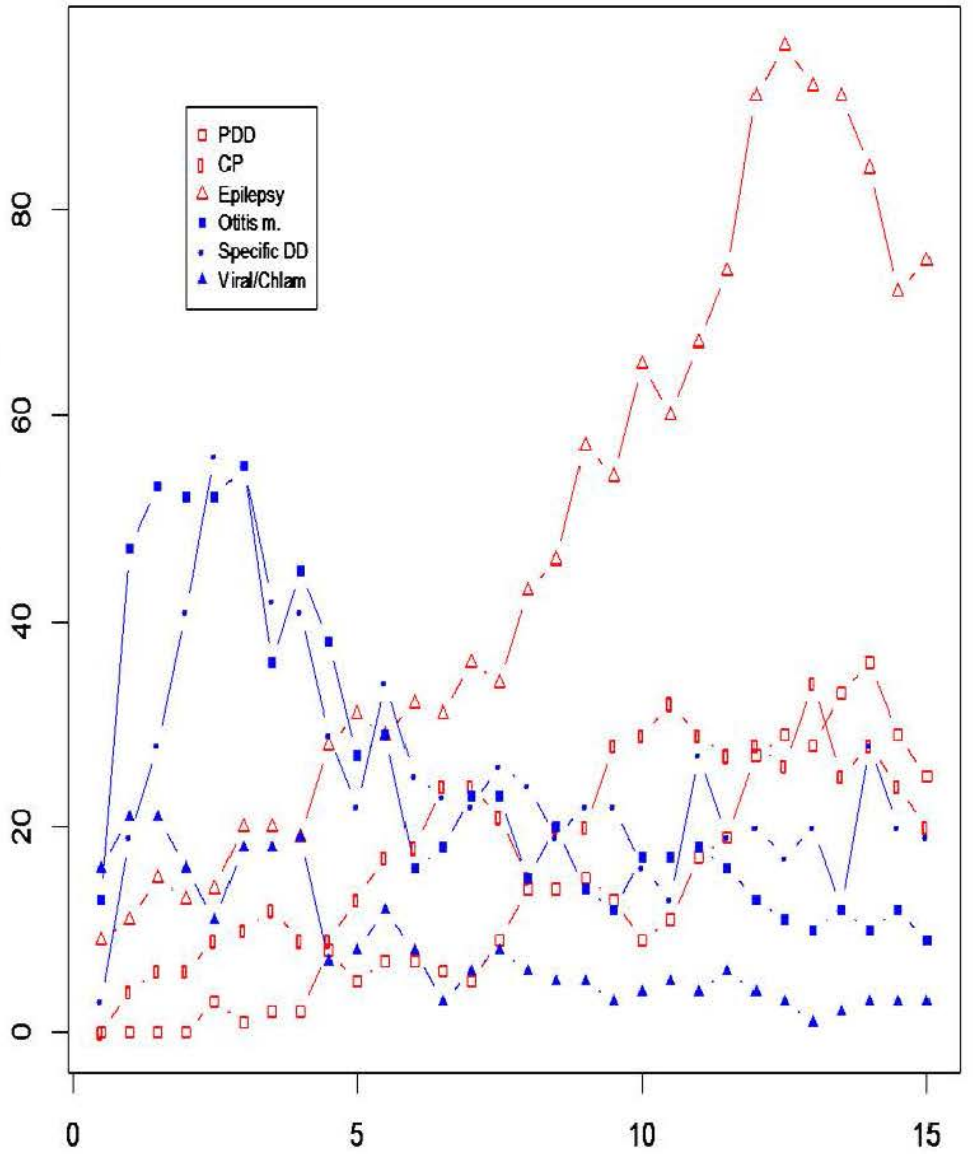


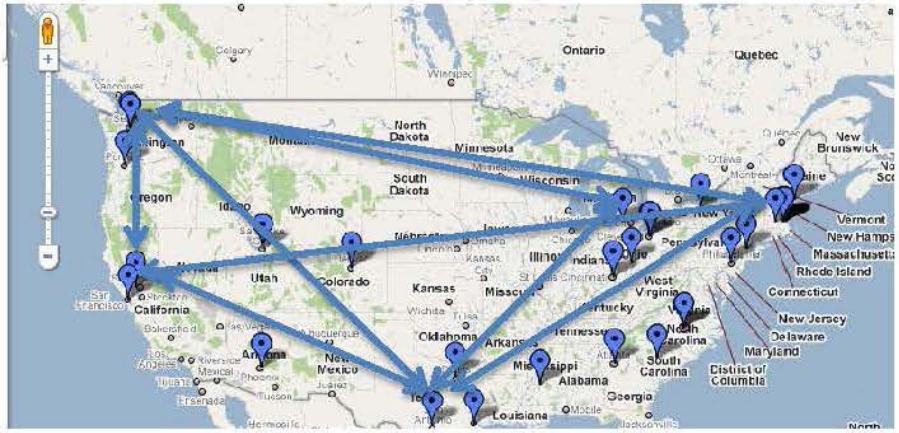
Autism or Autisms?

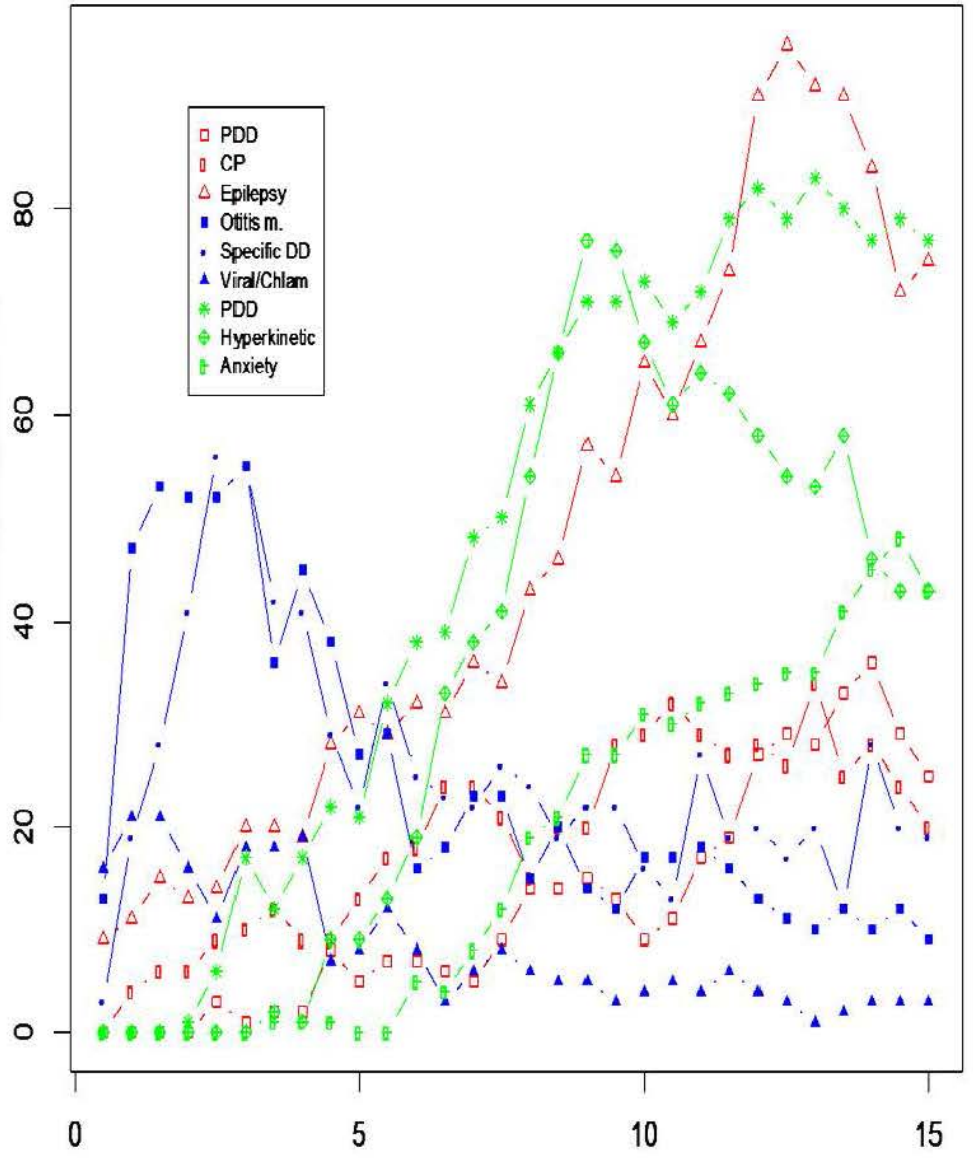





Ontario
Quebec
North Dakota
Minnesota
South Dakota
Wisconsin
Oregon
Wyoming
Utah
Colorado
Kansas
California
Oklahoma
Arkansas
Louisiana
Alabama
Georgia
South Carolina
West Virginia
Kentucky
Vermont
New Hampshire
Massachusetts
Rhode Island
Connecticut
New Jersey
Delaware
Maryland
District of Columbia
New Mexico


PDD
CP
Epilepsy
Otitis m.
Specific DD
Viral/Chlam
0
20
40
60
80
0
5
10
15
Years



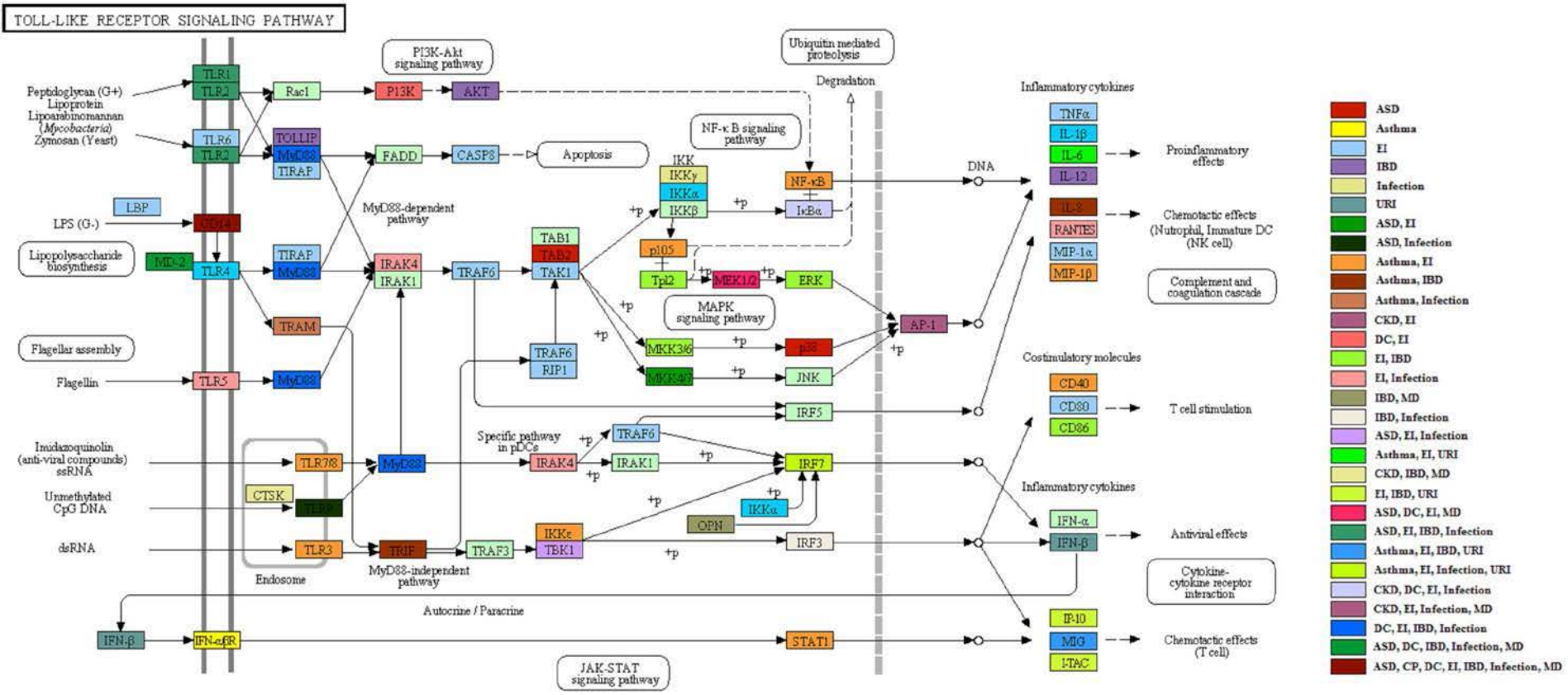


HARVARD MEDICAL SCHOOL
DEPARTMENT OF Biomedical Informatics

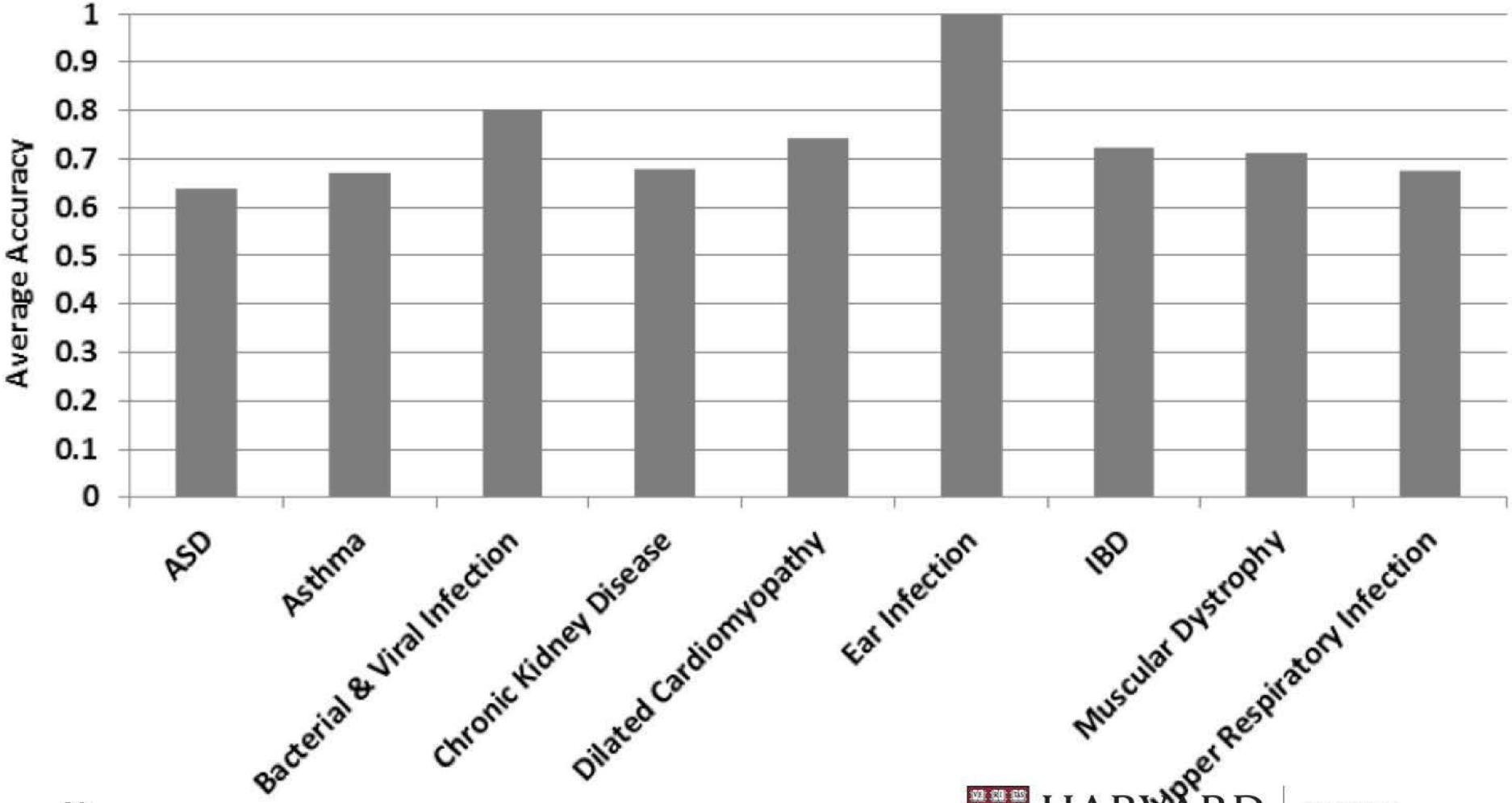
What about increasing overlaps across diagnoses?

What genes are shared across co-morbid conditions?

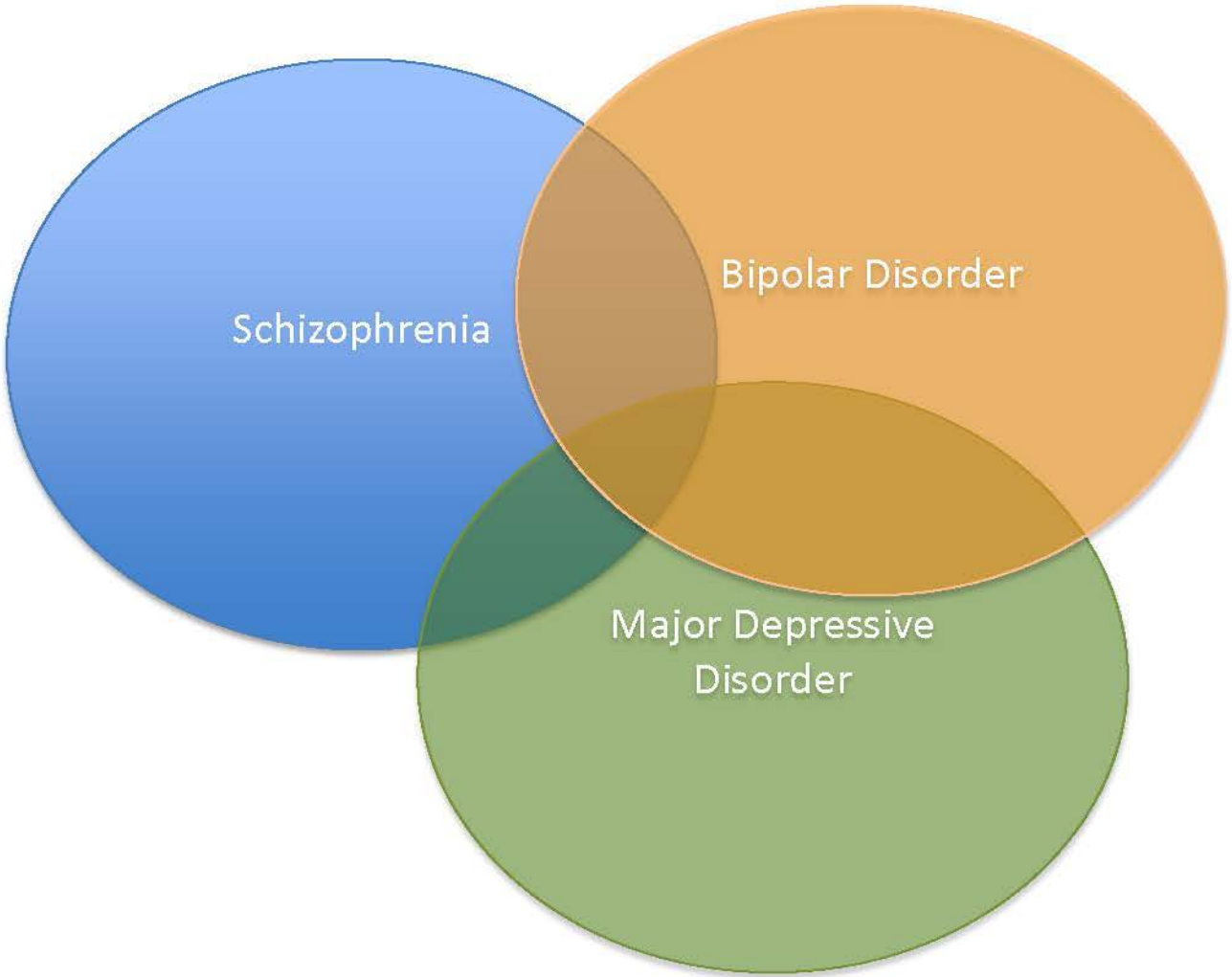
Nazeen et al. *Genome Biology* 17 (1): 228, 2016.



Classification accuracy reveals shared biology



What is the disease?



Shared Genes?
Shared Symptoms?

Autism(s)

- Implications for study and treatment
- Currently not obvious from a genetics-first/only approach
- Why clinicians might miss it
- Unsupervised vs self-referential & circular supervised
- Multi-modal-first exploration.

What about when there is NO diagnosis?

UDN Data Process Overview

Undiagnosed Diseases Network

If accepted

Data Mgt. Step #2: referral and data transfer to a UDN clinical center

Data Mgt. Step #1: Patient Application & initial response

UDN CC

UDN Gateway

Public summaries

UDN Data Warehouse

Metrics outcomes

UDN Record Archive and Electronic Trial Mgt.

STANFORD HOSPITAL & CLINICS | UCLA Health System | BAYLOR Health Care System | Vanderbilt University Medical Center | U | Harvard | National Institutes of Health

Data Mgt. Step #4: post visit reporting and review

Digital files, workflow, admin reports, phenotype, genotype

STANFORD HOSPITAL & CLINICS | UCLA Health System | BAYLOR Health Care System | Vanderbilt University Medical Center | U | Harvard | National Institutes of Health

Data Mgt. Step #3: in[illegible]ert[illegible]ation

HARVARD MEDICAL SCHOOL | DEPARTMENT OF Biomedical Informatics

Seven clinical sites

1 Baylor College of Medicine and Texas Children's Hospital
2 Duke Medicine with Columbia University Medical Center
3 Harvard Teaching Hospitals (BCH, BWH, MGH)
4 National Institutes of Health
5 Stanford Medicine
6 UCLA School of Medicine
7 Vanderbilt University Medical Center

Six additional research sites

Central Biorepository
8 Vanderbilt University Medical Center

Coordinating Center
9 Harvard Medical School

DNA Sequencing Core Facilities
10 Baylor College of Medicine
11 a HudsonAlpha Institute for Biotechnology *with*
b Illumina

Metabolomics Core Facility
12 a Pacific Northwest National Laboratories *with*
b Oregon Health & Science University

Model Organisms Screening Center
13 a Baylor College of Medicine *with*
b University of Oregon

Applications Received	Applications Under Review	Participants Accepted
822	311	298

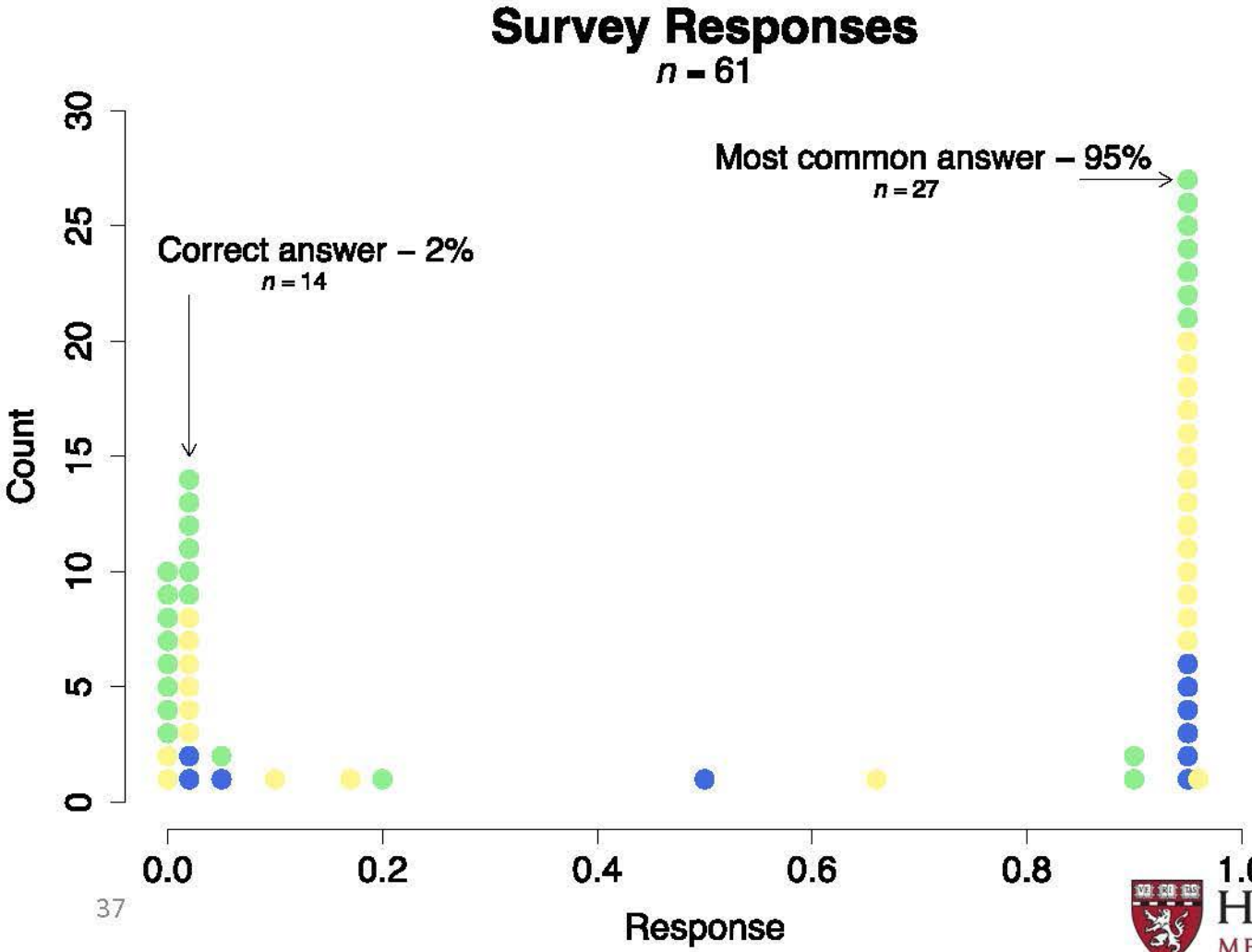
Model for the hard cases?

When there is no diagnosis

- If there is a high probability causal variant:
  - Diagnostic label links clinical findings in that patient to that variant
- Do all individuals with that variant have disease?
  - How does genetic background/environment contribute in the general population
  - $p(D|V) \neq p(V|D)$ {cf. HFE & Hemochromatosis}

Why is medicine so dependent on categorical diagnoses?

Medicine's Uncomfortable Relationship With Math: Calculating Positive Predictive Value



Summary

- Addiction to categorical diagnoses is cognitively useful
  - But our patients have beaten us to the Google-reflex
- Categorical diagnosis can result in spurious biological & clinical inference
  - Often manipulated for $$$, 2ndry agenda
  - Single measurement modality (incl. genomics) easier to manipulate
- Prediction of diagnostic class not necessarily evidence of biological etiology
- Diagnoses are more **robust** and **useful** when
  - Data-driven
  - Formally model health systems dynamics
  - Statistically-informed
  - ***Multi-modal***
  - Unsupervised or lightly supervised.

Acknowledgments

Nathan Palmer

Susanne Churchill

Arjun "Raj" Manrai

Andrew Beam

Denis Agniel

Matt Might

Griffin Weber

Rachel Ramoni

Sumaiya Nazeen

Bonnie Berger

Alexa McCray

HARVARD MEDICAL SCHOOL

DEPARTMENT OF Biomedical Informatics