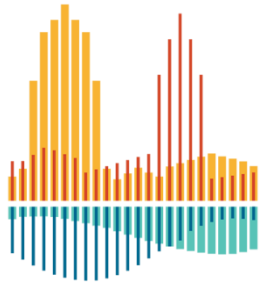


# PRECISION MEDICINE RESEARCH

## WHAT YOU NEED TO KNOW



INSTITUTE FOR  
COMPUTATIONAL  
BIOLOGY

May 11, 2018

Dana C. Crawford, PhD

Associate Professor

Population and Quantitative Health Sciences

Institute for Computational Biology

# Precision Medicine

## A New Initiative

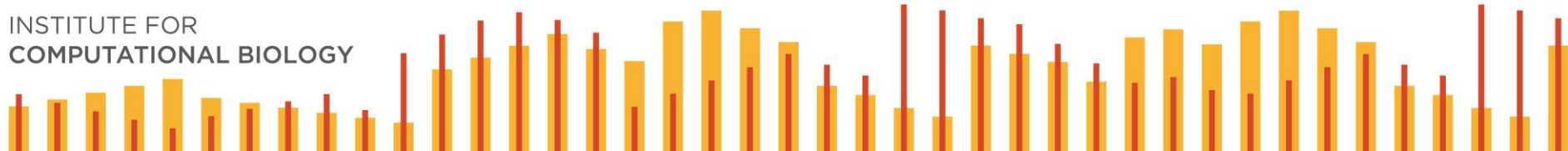


*Lancet* 385:2448-2449 (2015)

- Launched January 20, 2015
- 2016 budget of \$215M



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



# Precision Medicine

## The Concept



RELAX  
AND BE  
PATIENT

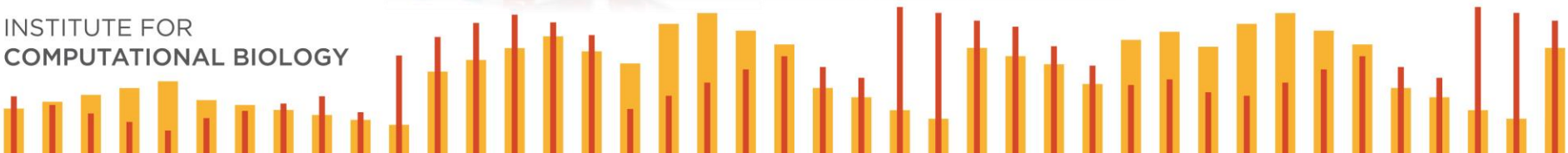
Right Patient

Right Drug



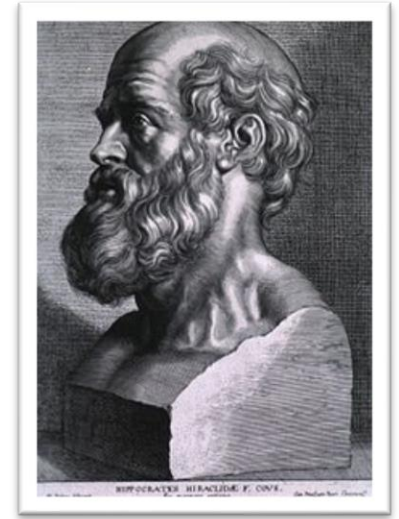
Right Dose

First Time



# Precision Medicine

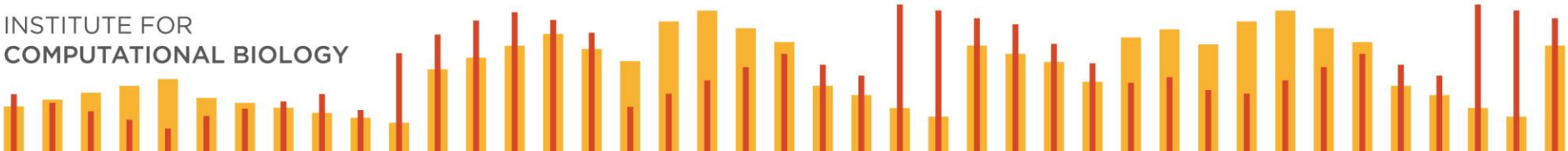
## The Concept



<https://en.wikipedia.org/wiki/Hippocrates>

- AKA Personalized Medicine
- Not a new concept

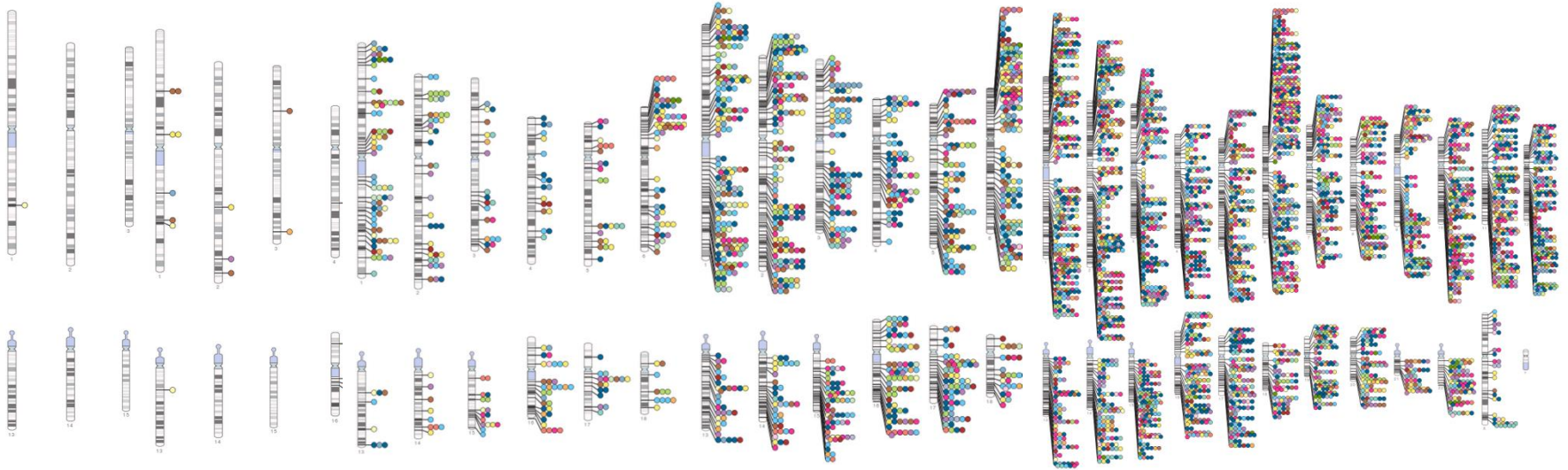
It's far more important to know what person the disease has than what disease the person has. – Hippocrates (c. 460 – c. 370 BC)



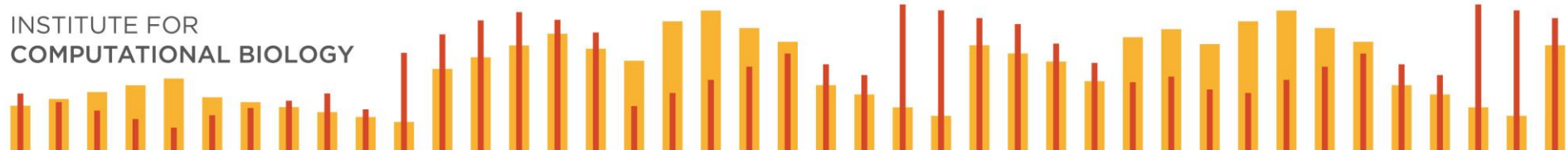
# Precision Medicine

## The Concept

Accelerated Genomic Discovery



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



<https://www.ebi.ac.uk/gwas/>

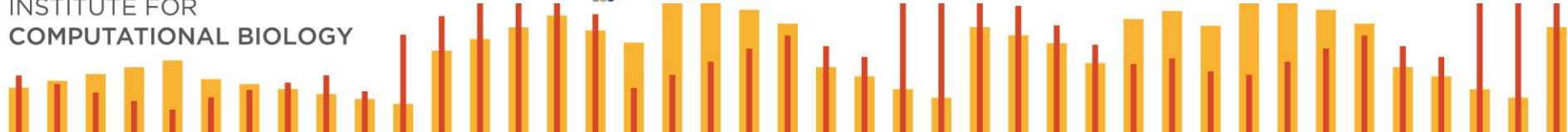
## GWAS as of 5/03:

3,361 publications

61,173 associated  
SNPs ( $10^{-5}$ )



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



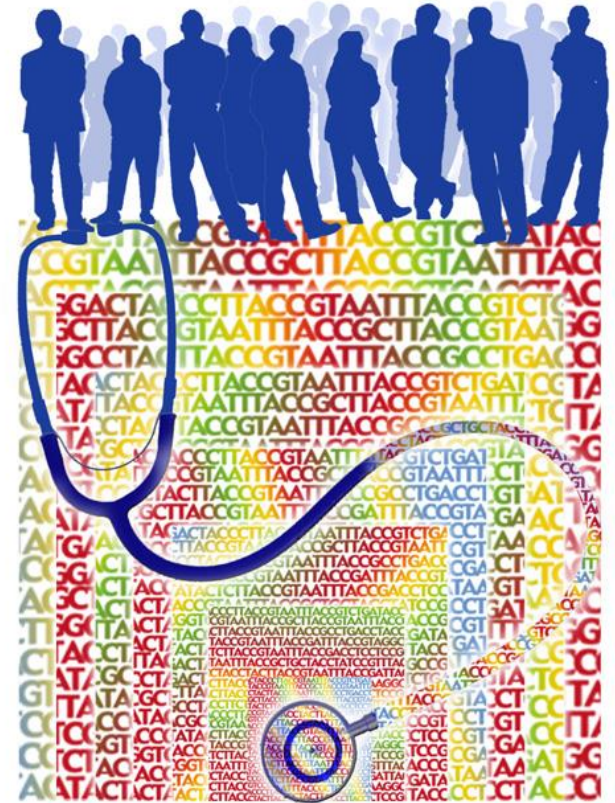
# Precision Medicine

## The Concept

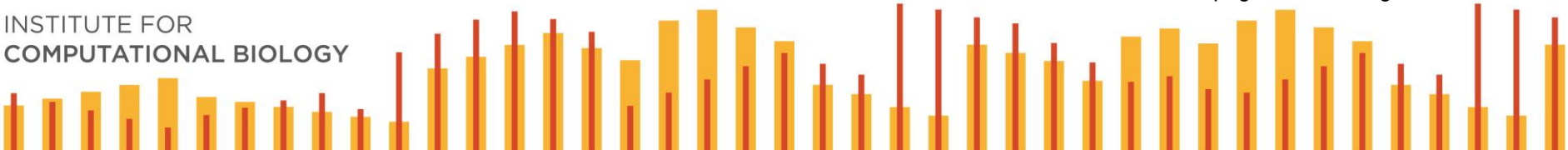
### Cost-effective Clinical Sequencing

- Whole exome  
~\$5,000 - \$7,000\*
- Whole genome  
~\$10,000

(\$300-500 at Geisinger)



[www.phgfoundation.org](http://www.phgfoundation.org)



# Precision Medicine

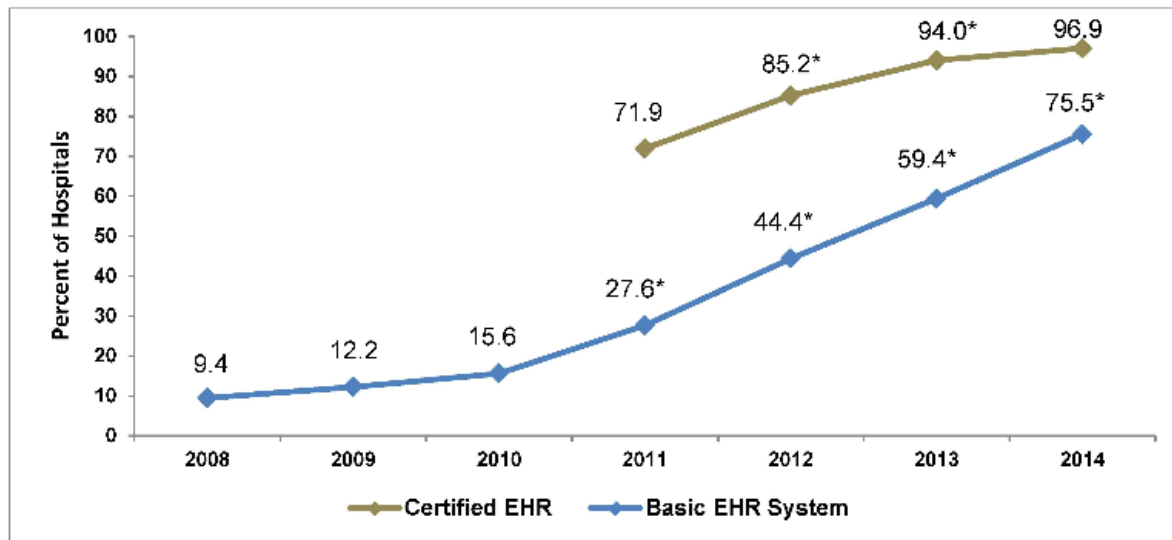
## The Concept



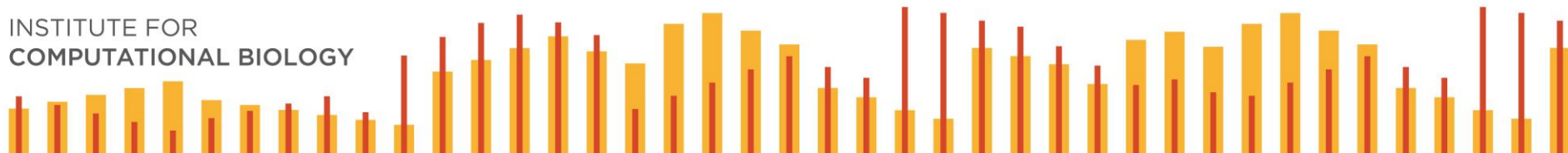
The rapid rise of  
EHRs:

As of 2014, 3 out of 4  
US hospitals have at  
least a basic EHR

Figure 1: Percent of non-Federal acute care hospitals with adoption of at least a Basic EHR with notes system and possession of a certified EHR: 2008-2014

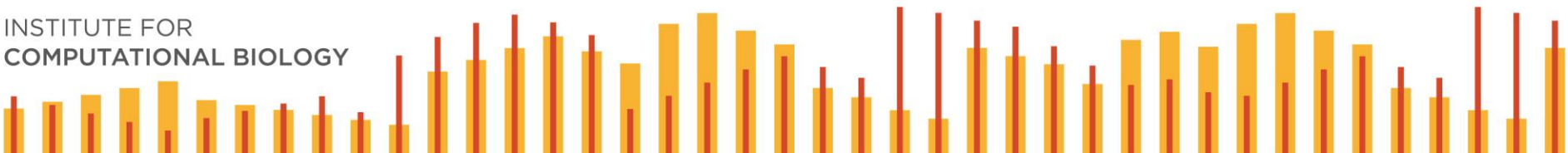


From <https://www.healthit.gov/sites/default/files/data-brief/2014HospitalAdoptionDataBrief.pdf>



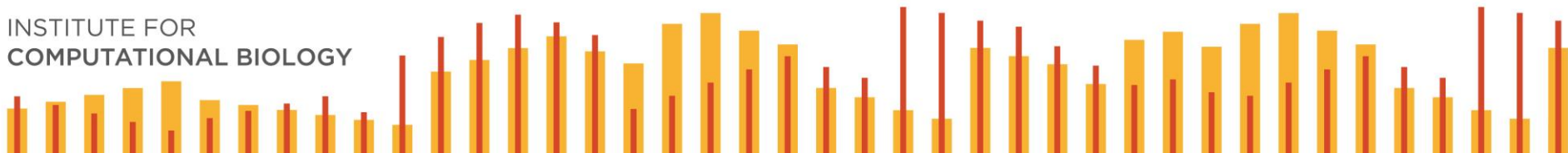
# WHAT'S AN ELECTRONIC MEDICAL RECORD (EMR)?

- A digital version of a patient's paper chart



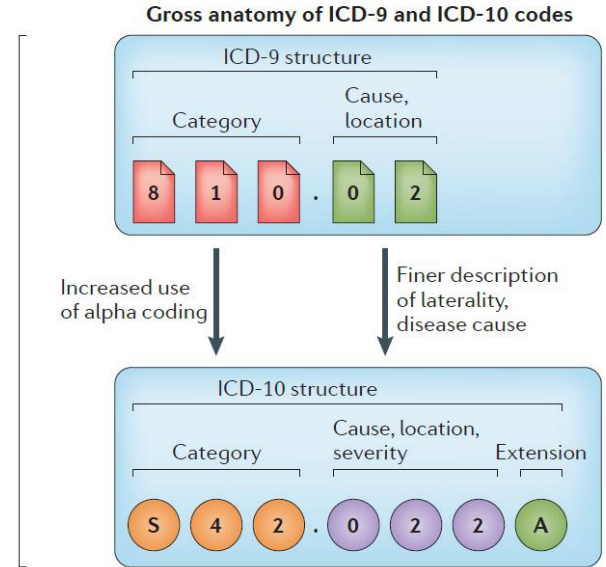
# WHAT'S IN AN EMR?

- Demographics
  - Structured and unstructured text
- Vitals
  - Structured
- Medical History
  - Structured data and unstructured text
- Medical encounter
  - Structured data and unstructured text
- Orders and prescriptions
  - Structured data
- Test results
  - Structured data



# WHAT'S IN AN EMR?

- International Classification of Diseases (ICD) codes
  - Aka billing codes
  - Coding system used to code signs, symptoms, injuries, diseases, and conditions
    - ICD-9-CM (Ninth Revision, Clinical Modification)
    - ICD-10-CMS/PCS
- Current Procedural Terminology (CPT) codes



Bush, Oetjens, Crawford *Nat Rev Genet* 17(3):129-45 (2016)

# WHAT'S IN AN ELECTRONIC MEDICAL RECORD (EMR)?

## Clinical narrative

- Social history
- Behavior/lifestyle
- Family history
- Unstructured data

Subject Status: ☐ Included ☐ Excluded ☐ Undetermined ☒ Not Reviewed

Select tabs to display

**Documents** × Medications × Labs × ICD9 Codes × Comments ×

Open Filters Open Highlight

**High Value Documents** Other Documents Problem Lists

- Synthroid 0.15 mg qd

ALLERGIES:

- NKDA

FAMILY MEDICAL HISTORY:

- Mother(70): Deceased, CHF, hypertension

- Father(40): Deceased, stomach ulcers, colitis

- Siblings(x10): Four deceased, diabetes, CAD, MI, hypertension

- Daughter(39): Unknown blood disorder

SOCIAL HISTORY:

- Marital status: Single

- Education: Completed 8th grade

- Unemployed: Disability/SSI

- Tobacco: None

- Alcohol: No history

- Illicit drug use: No history

EXAMINATION

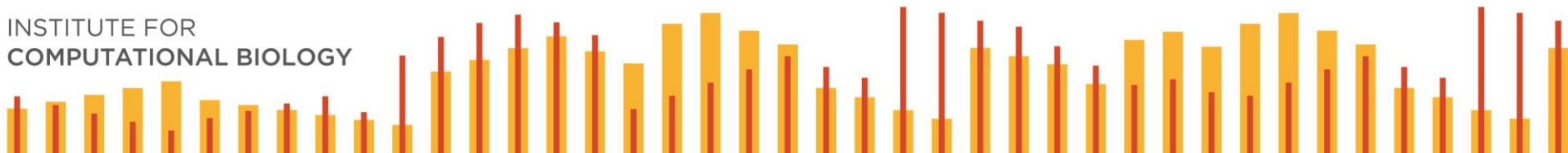
VITAL S: P: 64 BPM BP: 170/50 mm Hg Wt: 165 lb Ht: 63.0



# WHAT'S AN ELECTRONIC HEALTH RECORD (EHR)?



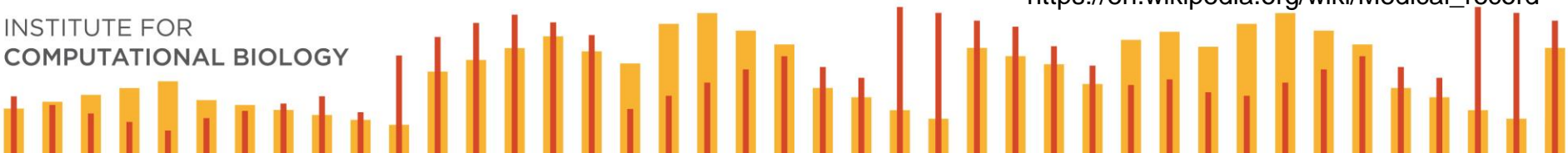
- A digital version of a patient's paper chart
- EHRs go beyond the data collected in the provider's office and include a more comprehensive patient history.
  - Designed to contain and share information from all providers involved in a patient's care.
  - Data can be created, managed, and consulted by authorized providers and staff from across more than one health care organization.
  - Also allow a patient's health record to move with them—to other health care providers, specialists, hospitals, nursing homes, and even across states.



# EHRs make population-scale research possible



[https://en.wikipedia.org/wiki/Medical\\_record](https://en.wikipedia.org/wiki/Medical_record)



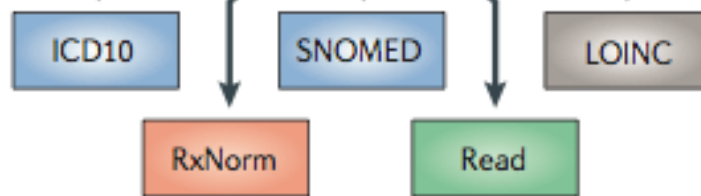
Mrs Jones is an 84-year-old African-American woman admitted from the emergency department with complaint of **crushing substernal** pain ... past medical history is significant for a 20 year history of **type 2 diabetes mellitus** controlled with **oral hypoglycaemics**, **2 ppd history smoking** ...

Family history: Sister died from **myocardial infarction** at 74 years ...

Mrs Jones was discharged on a 1,500 ml fluid restriction, **nitroglycerin** 0.4 mg/spray 1-2 spray po. **Aciphex** 20 mg (20 mg tablet DR take 1) PO

Discharge diagnosis: **acute MI**, **diabetes mellitus** ...

### Unified Medical Language System



UMLS maps a data type from one vocabulary to another

Standardized vocabularies in which all EHR-data items are represented

Diagnoses

Environment

Medications

Family history

Symptoms

Billing codes

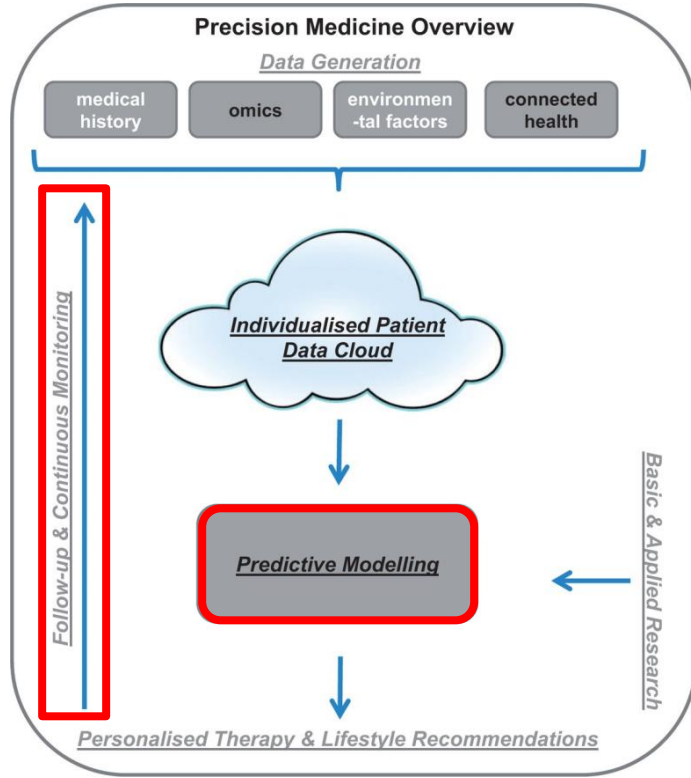
Dispensed medications

Laboratory measurements

EHR-derived database composed of natively codified and NLP-codified data types

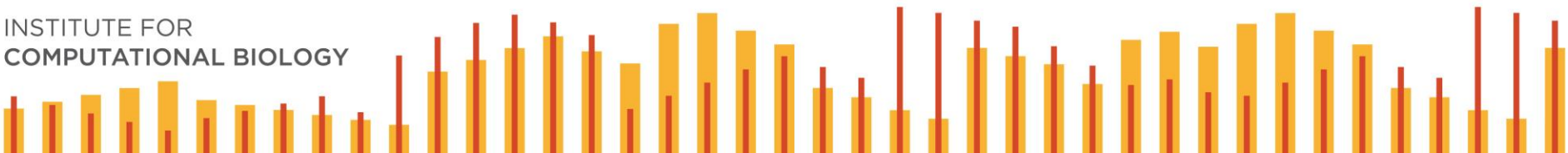
# Precision Medicine

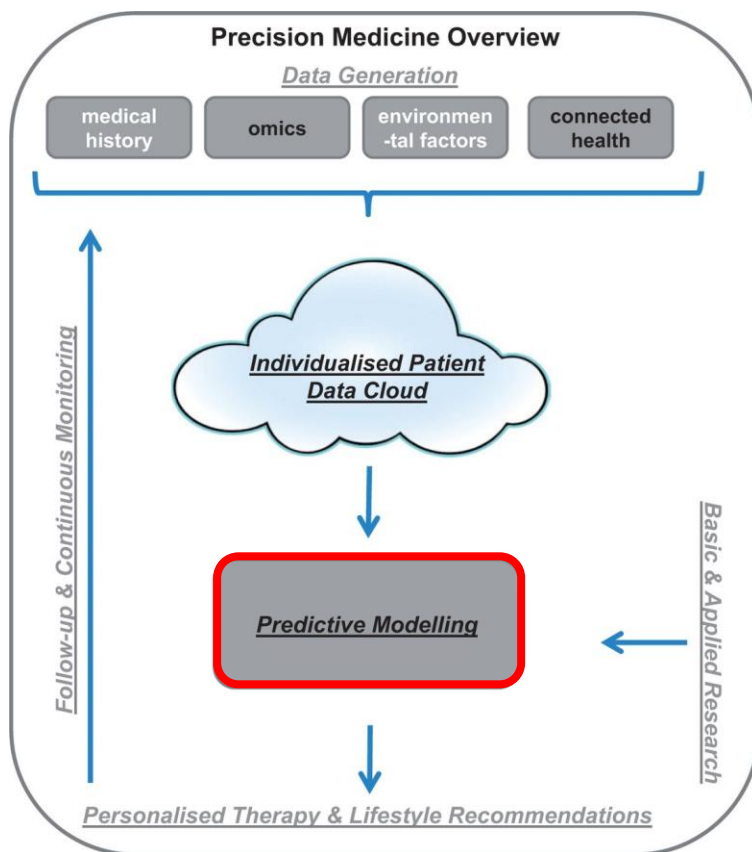
## The Concept



Duffy DJ *Brief Bioinform* 17(3):494-504 (2016)

- Do we have the data for this?
- What will this require?

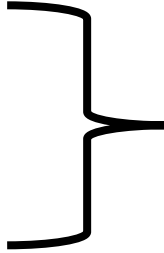




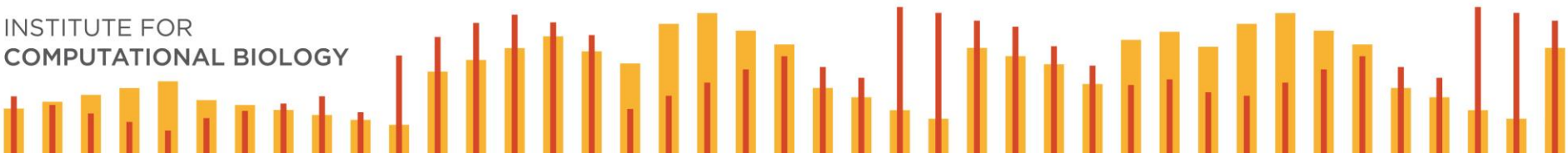
Duffy DJ *Brief Bioinform* 17(3):494-504 (2016)

## Need More Data?

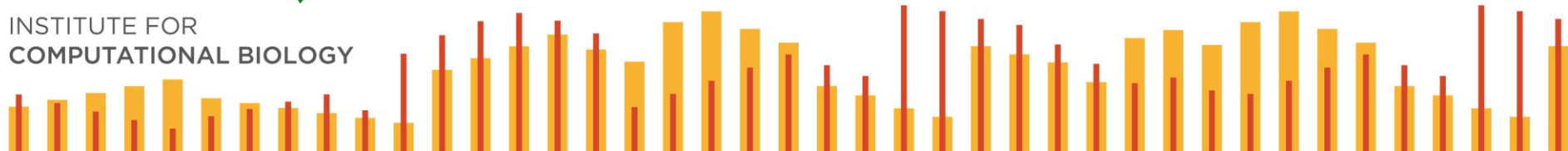
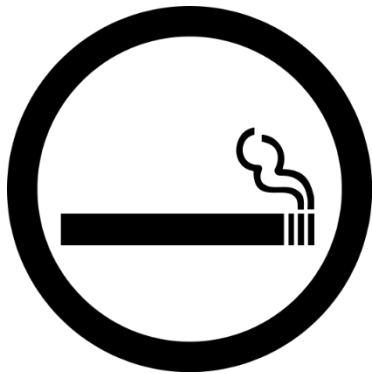
- More genetic data
- Other 'omic data
- Co-morbidities



- Longitudinal data
- Environmental exposures

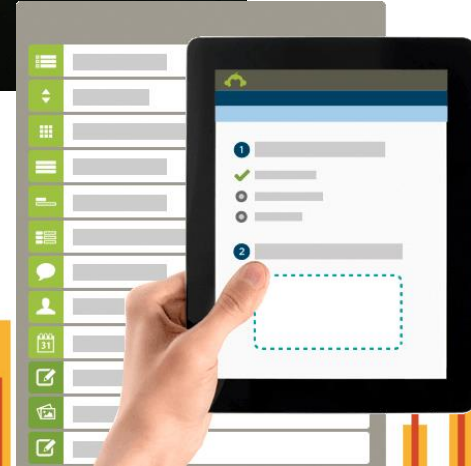


# EHRs and Exposure Data Consistently Inconsistent



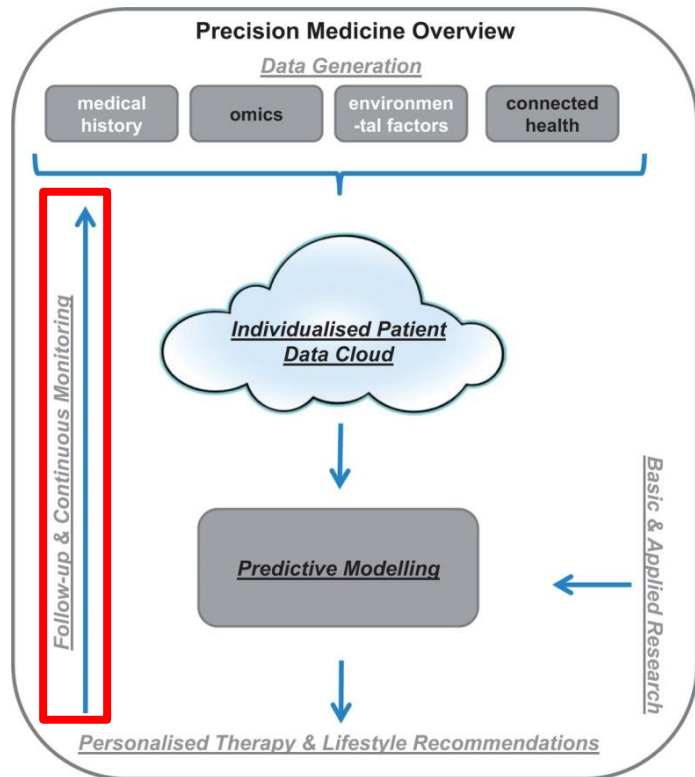
# Emerging Tools for Patient-Provided Data

## mHealth, Wearables, and Patient Portals



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY





Duffy DJ *Brief Bioinform* 17(3):494-504 (2016)

# EHRs and Follow-Up

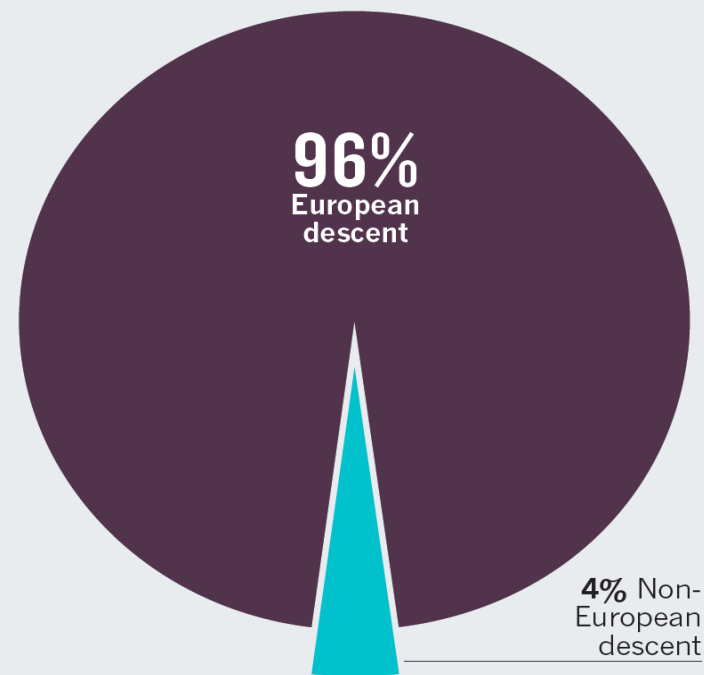
- Patients enter and leave systems
- Some clinic visits/services outside of system
- Valuable longitudinal data lost

# Genomic data based mostly on studies of European-descent

Bustamante, Burchard, de la Vega (2011)  
*Nature* 475:163-165

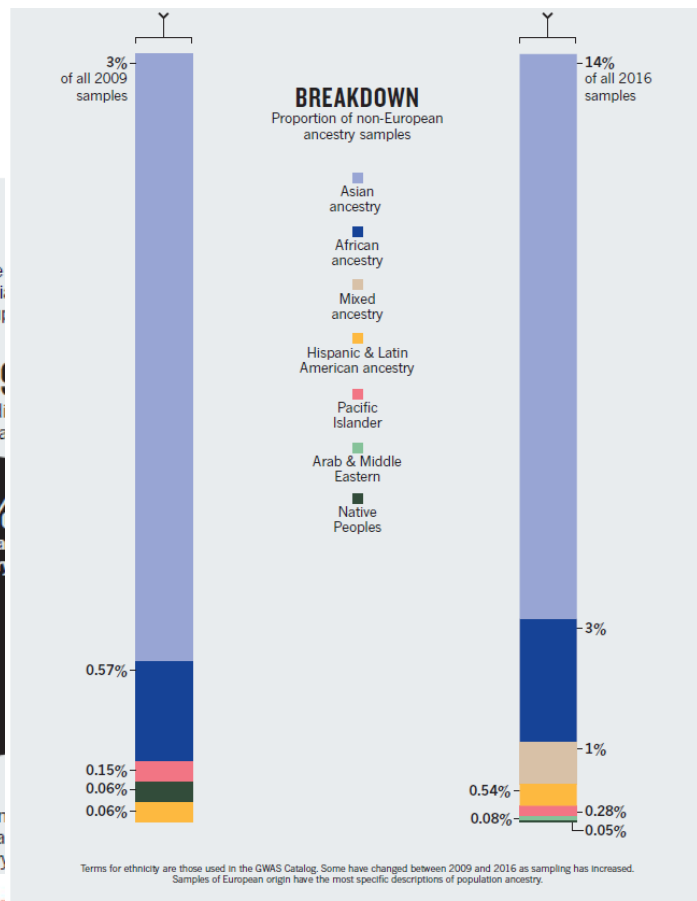
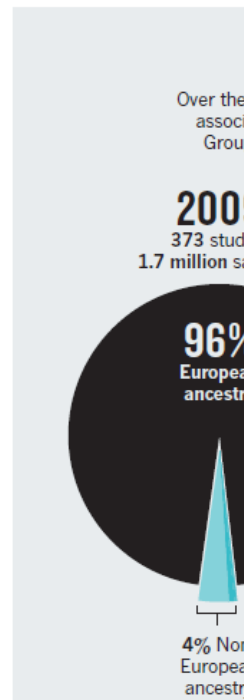
## SAMPLING BIAS

Most genome-wide association studies have been of people of European descent.



# Getting Better?

Popejoy and Fullerton (2016)  
*Nature* 538: 161-164



INSTITUTE FOR  
 COMPUTATIONAL BIOLOGY

# Diversity in Research Matters

## SPECIAL ARTICLE

### Genetic Misdiagnoses and the Potential for Health Disparities

Arjun K. Manrai, Ph.D., Birgit H. Funke, Ph.D., Heidi L. Rehm, Ph.D., Morten S. Olesen, Ph.D., Bradley A. Maron, M.D., Peter Szolovits, Ph.D., David M. Margulies, M.D., Joseph Loscalzo, M.D., Ph.D., and Isaac S. Kohane, M.D., Ph.D.

N Engl J Med 2016; 375:655-665 | [August 18, 2016](#) | DOI: 10.1056/NEJMsa1507092

## REPORT

### Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans

Giulio Genovese<sup>1,2,\*</sup>, David J. Friedman<sup>1,3,\*</sup>, Michael D. Ross<sup>4</sup>, Laurence Lecordier<sup>5</sup>, Pierrick Uzureau<sup>5</sup>, Barry I. Freedman<sup>6</sup>, Donald W. Bowden<sup>7,8</sup>, Carl D. Langefeld<sup>8,9</sup>, Taras K. Oleksyk<sup>10</sup>, Andrea L. Uscinski Knob<sup>4</sup>, Andrea J. Bernhardt<sup>1</sup>, Pamela J. Hicks<sup>7,8</sup>, George W. Nelson<sup>11</sup>, Benoit Vanhollebeke<sup>5</sup>, Cheryl A. Winkler<sup>12</sup>, Jeffrey B. Kopp<sup>11</sup>, Etienne Pays<sup>5,†</sup>, Martin R. Pollak<sup>1,13,†</sup>

+ Author Affiliations

†To whom correspondence should be addressed. E-mail: [mpollak@bidmc.harvard.edu](mailto:mpollak@bidmc.harvard.edu) (M.R.P.); [epays@ulb.ac.be](mailto:epays@ulb.ac.be) (E.P.)

\* These authors contributed equally to this work.

Science 13 Aug 2010:  
Vol. 329, Issue 5993, pp. 841-845  
DOI: 10.1126/science.1193032

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



# National Precision Medicine

## A New Initiative



### 1 Million Cohort

Large

Linked to EHR and payer databases

Oversample subgroups

Recontact for return of results

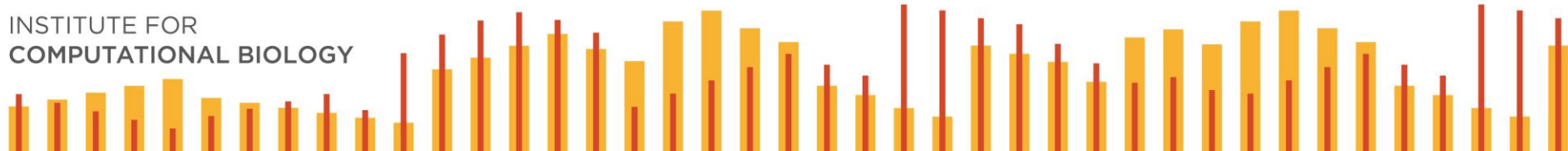
Wide age range

Sophisticated lifestyle data collection

Rare and common disease

Extensive follow-up

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



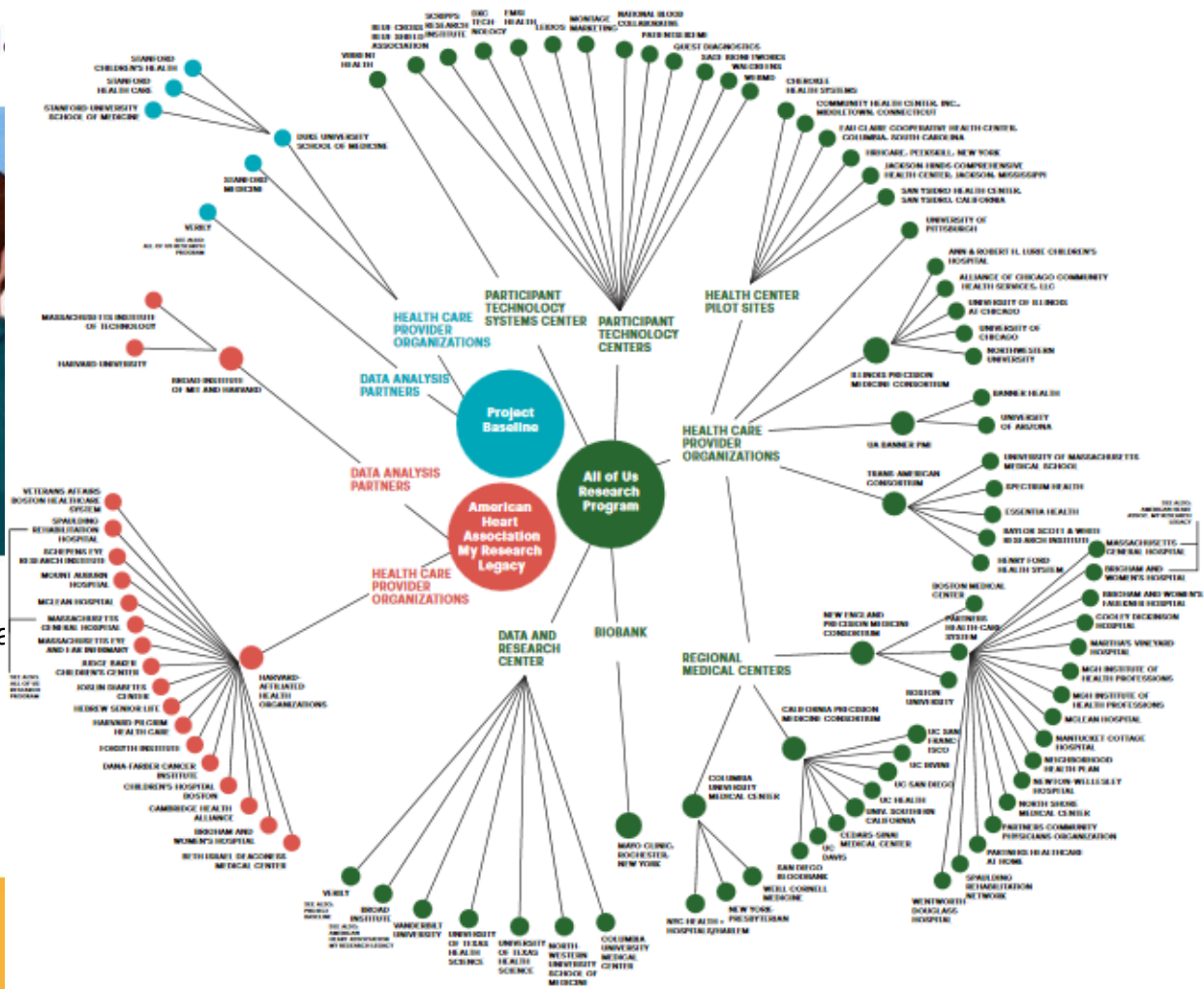
# Precision Medicine Initiative Cohort Program

(Photo credit: Dr. Janina Jeff at the 2017 New  
Balance Bronx 10 Mile)



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

# Health Care Provider Org



<https://allofus.nih.gov/about/program-components/health-care-provider-organiza>

<https://datasociety.net/output/precision-medicine-national-actor-map/>

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

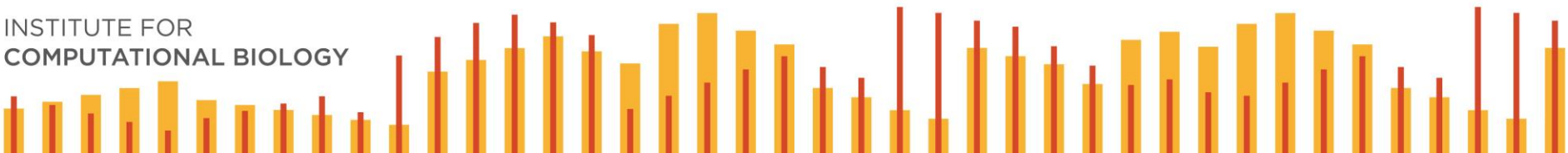


# Local Precision Medicine

## CLE is All In!



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

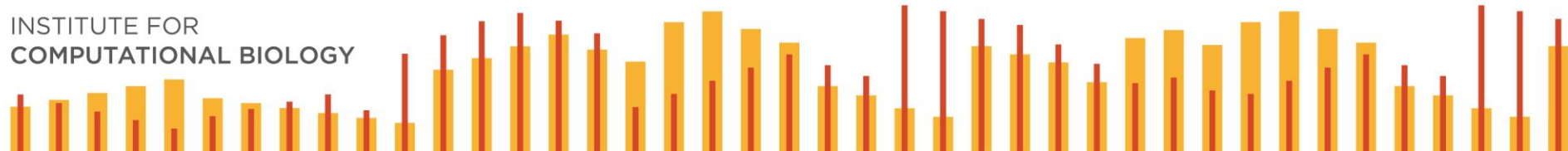


# The CLE is All In! Diversity

	Cleveland, OH	Cuyahoga County, OH	United States
Population	396,815	1,280,122	308,745,538
Persons under 18 years	24.6	22.7	24.0
Persons 65 years and over	12.0	15.5	13.0
White alone	37.3	63.6	72.4
Black or African American alone	53.3	29.7	12.6
High school graduate or higher	77.4	87.8	86.3
Bachelor's degree or higher	15.2	30.3	29.3
With a disability, < 65	15.3	10.2	8.5
Without health insurance, < 65	18	10	12.0
Median household income	\$26,179	\$44,203	\$53,482
Persons in poverty	35.9	19.6	14.8
Mean travel time to work (min.)	24.2	24.1	25.7

2010 US Census

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

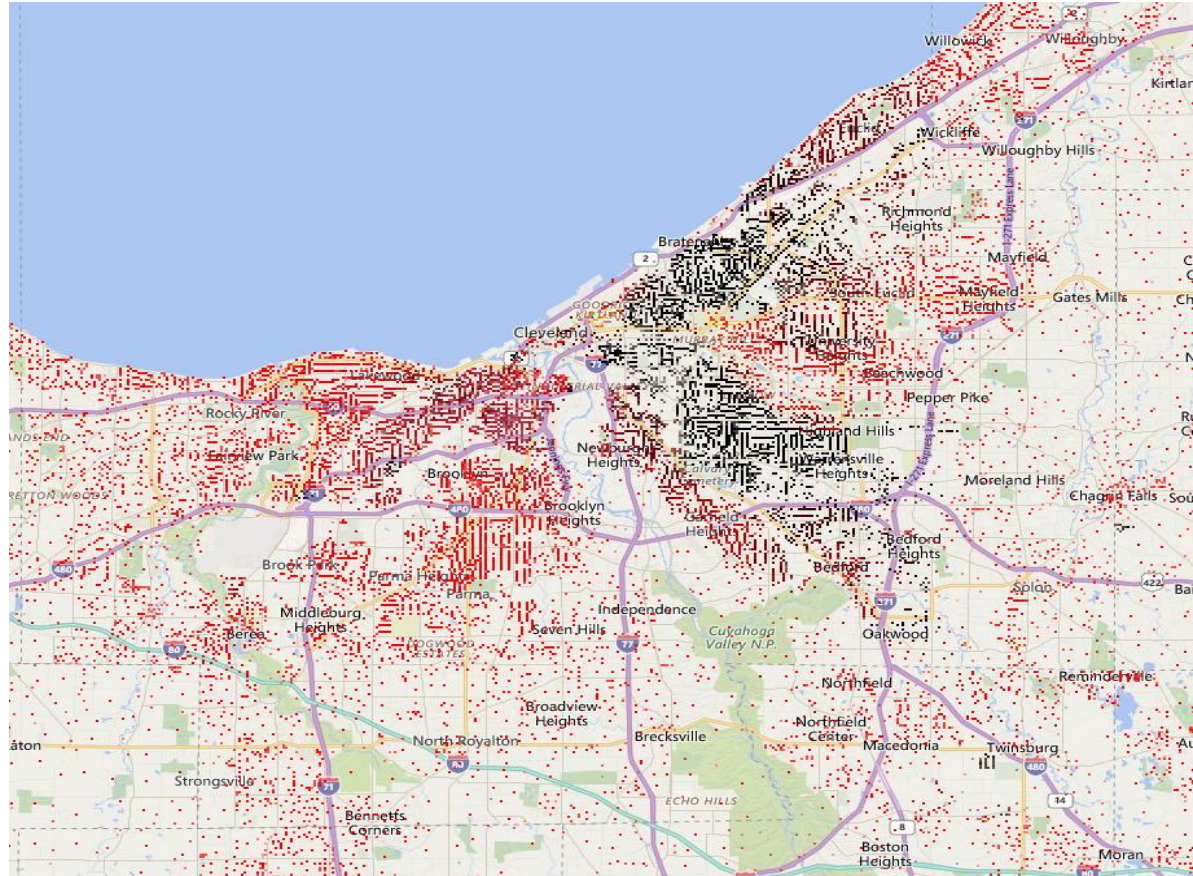




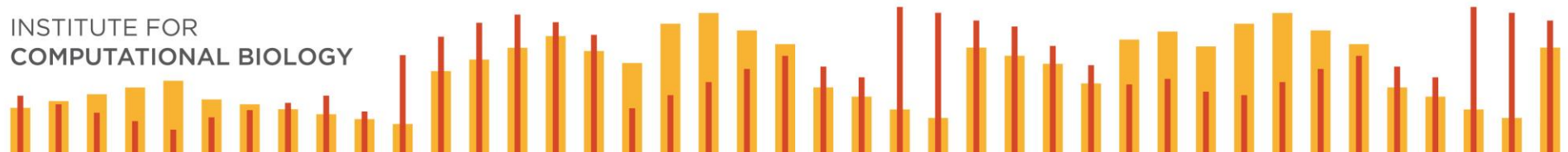
African  
American



European  
American



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



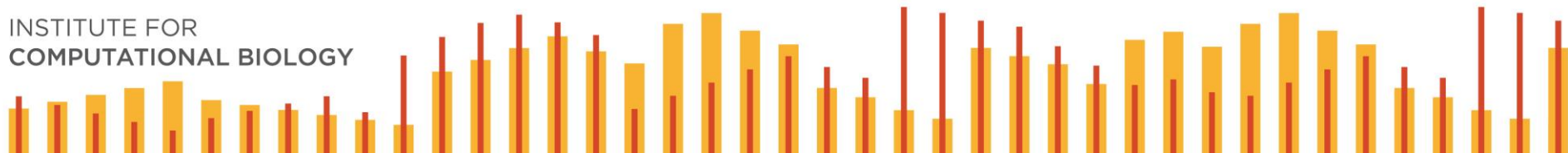
# The CLE is All In!

## Local Precision Medicine



2.9 million lives  
(>75% of all healthcare in Cleveland)

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



 **MetroHealth** Epic installed in 1999; stage 7 (as of 2014)



Allscripts installed in 2011; stage 6 (as of 2014)

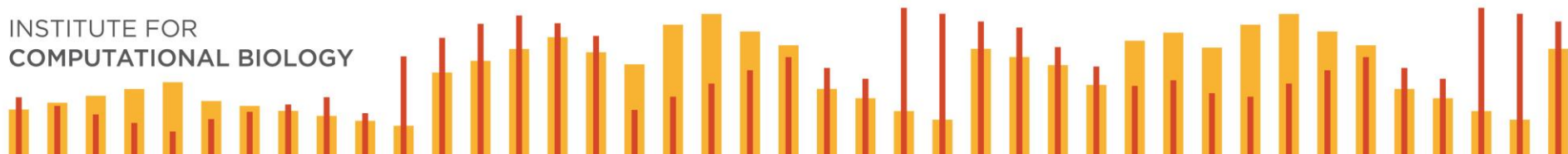


Epic installed in 2002; stage 7 (as of 2014)



VistA\*\*\* installed in late 1970s

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



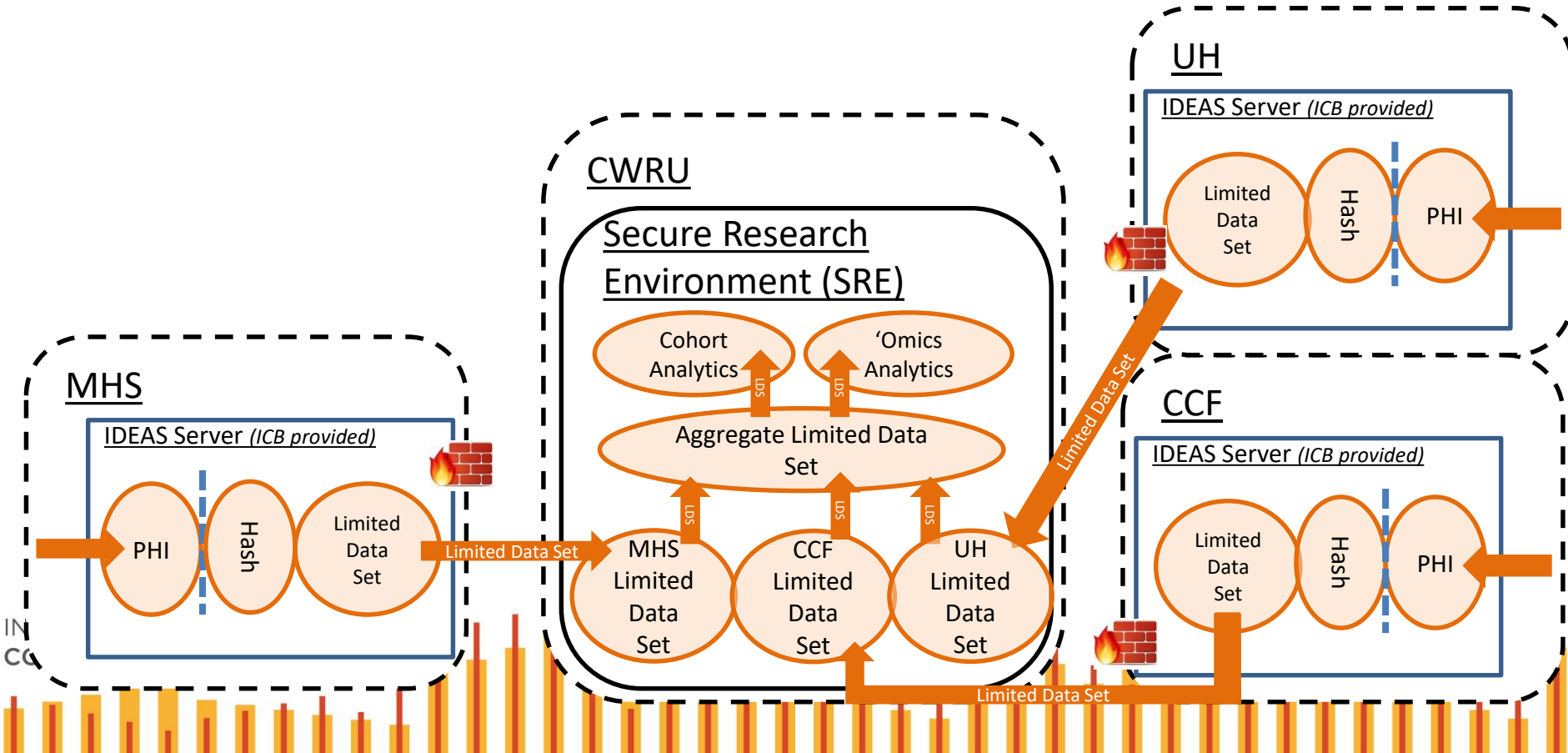
# CLEARPATH:

CLEVELAND AREA RESearch Platform for Advancing Translational Healthcare

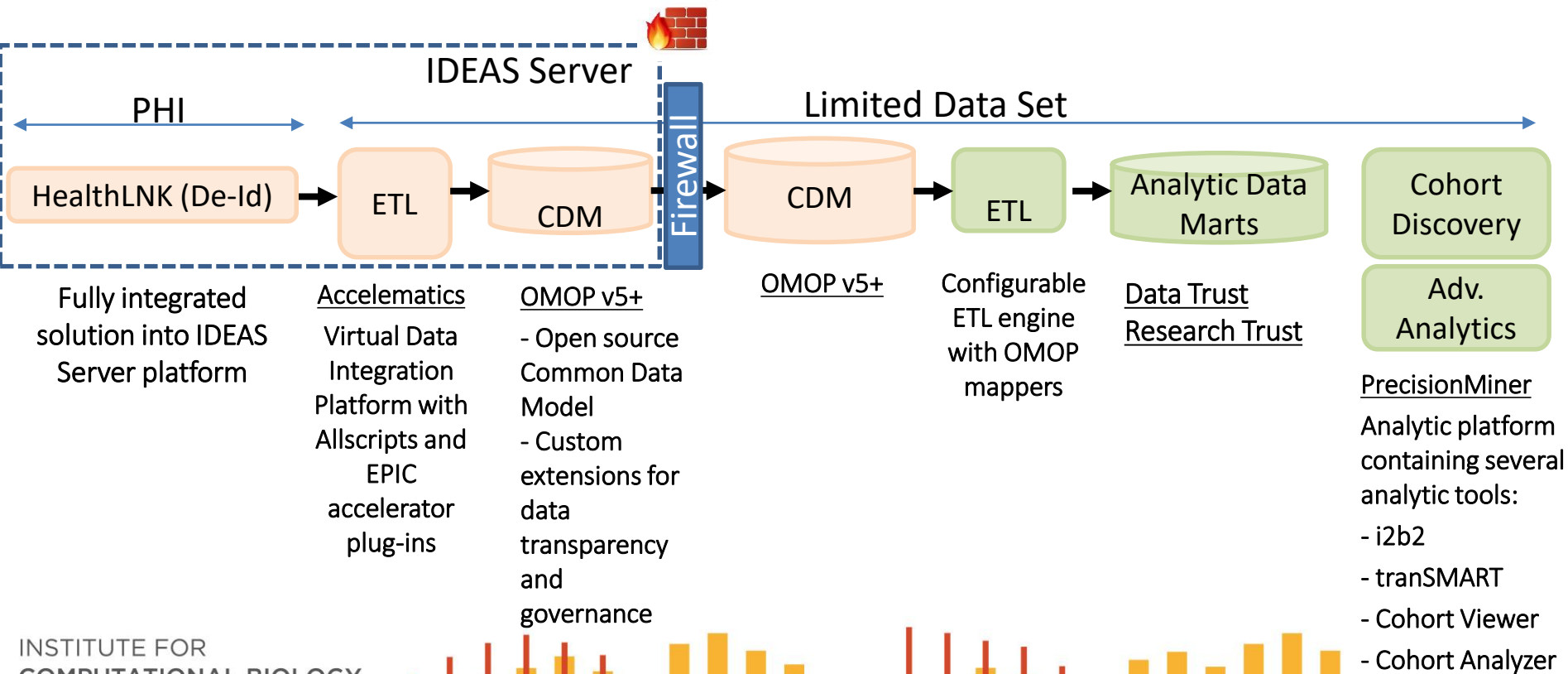
- Resource offered by the CWRU ICB
- Biomedical Big Data Warehouse
  - Provide a platform to capture and integrate multiple streams of data
  - Provide access and resources to
    - Query the data
    - Develop datasets for research
    - Analyze the data and interpret the results



# CLEARPATH: CONCEPTUAL MODEL



# CLEARPATH Architecture



# MetroHealth/ICB Pilot Study (MIPs)



David Kaelber, MD  
Chief Medical  
Informatics Officer



John Sedor, MD



John O'Toole, MD



Will Bush, PhD, MS

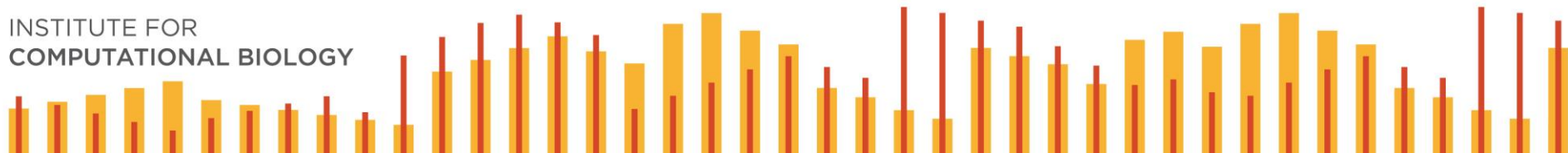


Jessica Cooke  
Bailey, PhD



**MetroHealth**

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



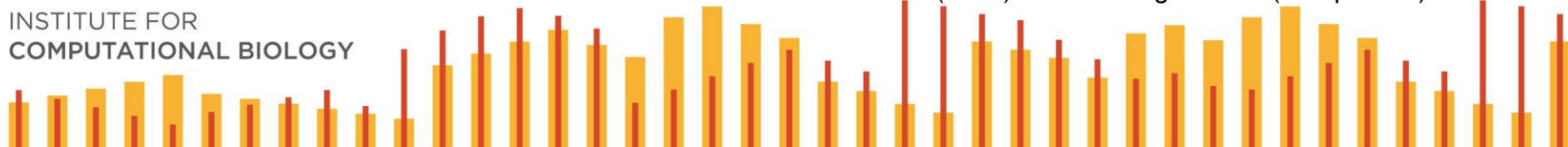
# MetroHealth/ICB Pilot Study (MIPs)

- Cleveland's first academic medical center (1837)
  - 731-bed hospital
  - 17 outpatient sites
- Annually, the system has approximately
  - 900,000 outpatient visits
  - 100,000 ED visits
  - 30,000 inpatient admissions.

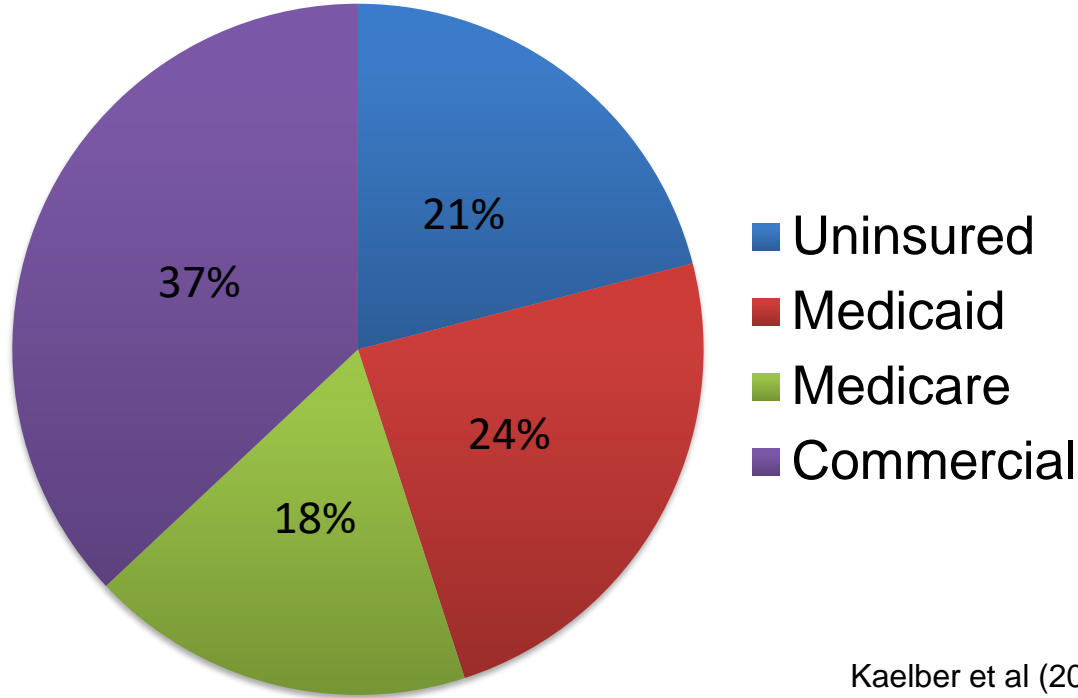


Kaelber et al (2013) *Am J Manag Care* 19(10 Spec No):SP337-43

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

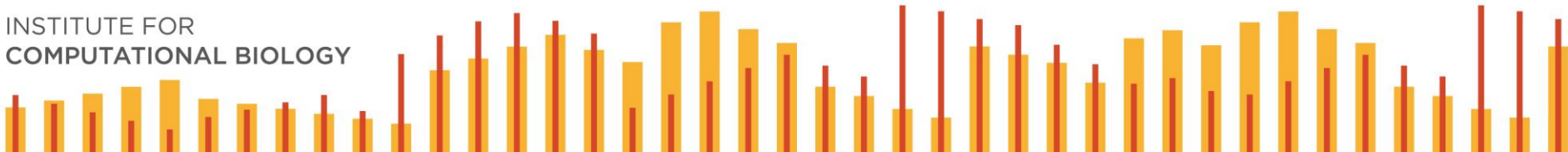


# MetroHealth/ICB Pilot Study (MIPs)



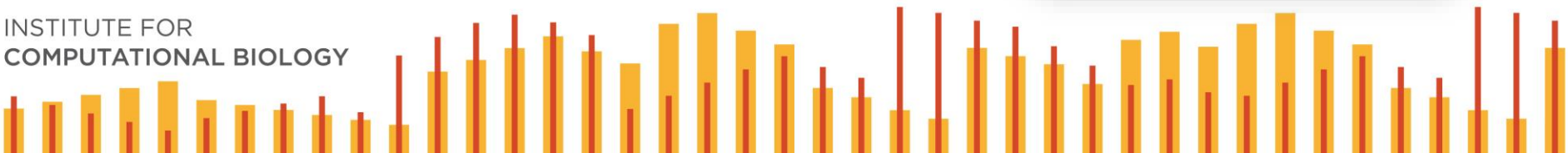
Kaelber et al (2013) *Am J Manag Care* 19(10 Spec No):SP337-43

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



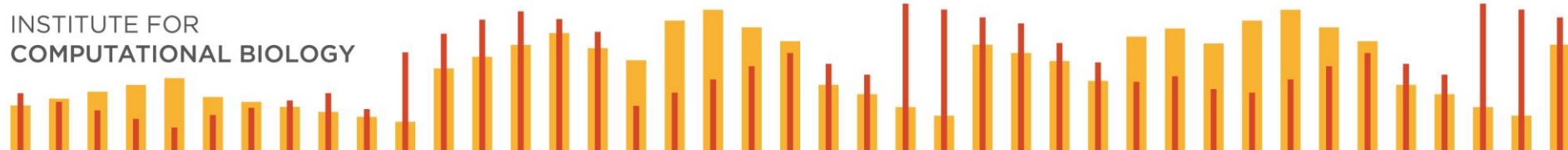
# MetroHealth/ICB Pilot Study (MIPs)

- Survey participants about attitudes towards PMI-like efforts
- Ascertain participants for biospecimen collection and 'omics



	Demographic Group	Weighted N	% who said the study definitely or probably should be done	Beta	S.E.	p-value	% definitely or probably willing to participate in hypothetical biobank	Beta	S.E.	p-value
<i>Total</i>		2,601	79%				54%			
<i>Gender</i>	Men	1,251	77%	-0.160	0.098	0.10	54%	0.007	0.080	0.93
	Women	1,350	80%	ref			54%	ref		
<i>Race and Ethnic Group</i>	White, non-Hispanic	1,721	79%	ref			53%	ref		
	Black, non-Hispanic	296	77%	-0.032	0.156	0.84	55%	0.100	0.131	0.44
	Hispanic (all races)	385	78%	0.136	0.149	0.36	59%	0.291	0.123	0.02
	Other non-Hispanic	200	81%	0.555	0.541	0.31	56%	0.058	0.167	0.73
<i>Survey language, among Hispanics</i>	Spanish	186	80%	0.159	0.252	0.53	61%	0.383	0.205	0.06
	English	199	80%	ref			56%	ref		
<i>Age</i>	21–29	447	81%	-0.047	0.048	0.32	60%	-0.156	0.039	<0.0001
	30–44	694	80%				58%			
	45–59	735	79%				53%			
	60+	724	77%				47%			

Kaufman et al *PLoS One* 11(8):e0160461 (2016)



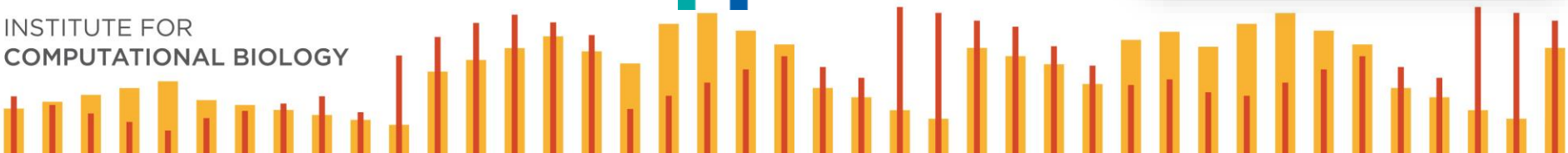
# MetroHealth/ICB Pilot Study (MIPs)

103 patients surveyed at Nephrology Clinic

- 50% African American
- 54% female
- Mean age 61.45 years  
(range: 18 – 91 years)



INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



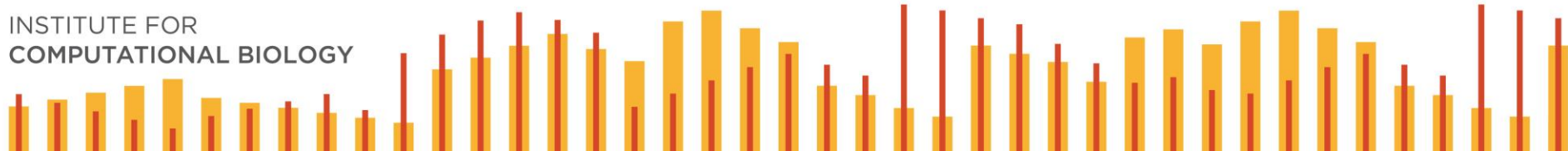
# MetroHealth/ICB Pilot Study (MIPs)

Would you be willing to allow your health records and genetic information to be stored in a national biorepository coordinated by the government as a part of their “Precision Medicine Initiative”? If so this information may be available to researchers nationally and internationally with the understanding that your privacy would be protected.



71% YES

69% YES



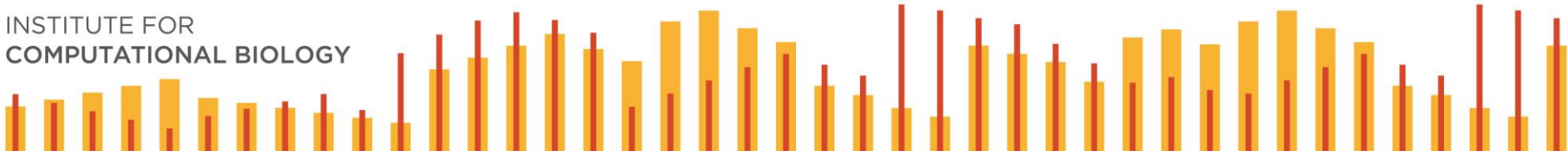
# MetroHealth/ICB Pilot Study (MIPs)

Would you be willing to install a free phone app that would be able to track your physical activity, measures of your health and location with the understanding that your privacy would be protected?



38% YES

46% YES



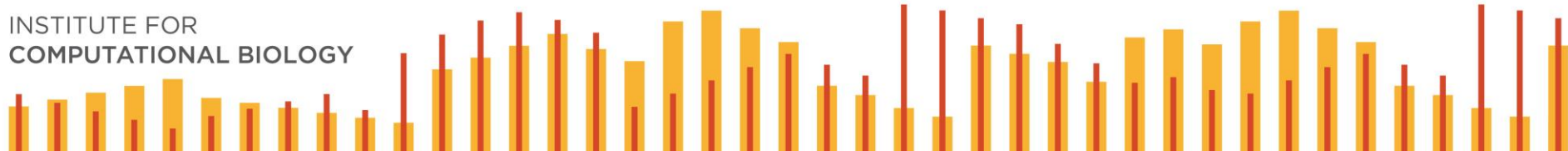
# MetroHealth/ICB Pilot Study (MIPs)

Would you be willing to allow the information collected using the phone app (question 3) to be sent to the national coordinating center where it may be shared with researchers nationally and internationally with the understanding that your privacy would be protected?



37% YES

54% YES



# MetroHealth/ICB Pilot Study (MIPs)

If you participated in a study that collected your genetic and health information, how important is it to you that you receive results from the study? Circle the number that is closest to how you feel.



1

Not at all important

~12%

- 2

Somewhat important

~20%

- 3

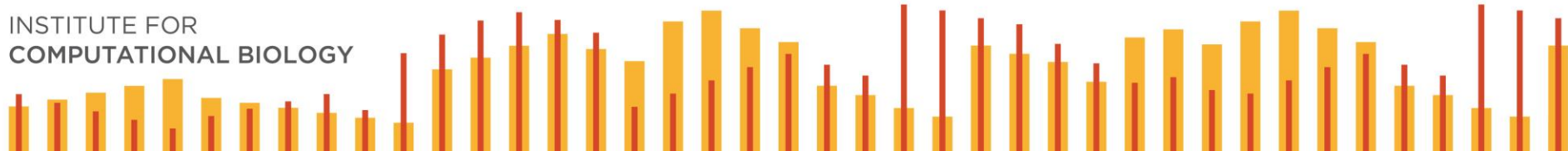
- 4

- 5

Very important

~62%

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



# MetroHealth/ICB Pilot Study (MIPs)

What type of results would you like to receive, please check all that apply

A – Information about average results – no personal results

B – Information about yourself your doc may have already told you

C – Information about your genes that may influence your doc's approach to care

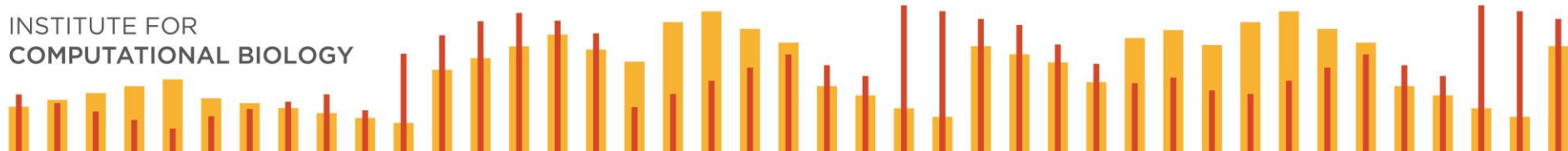
D – Information about your genes that has uncertain significance and won't change treatment

E – I do not want to receive any results



~77% GENETICS (C)

~10% NOTHING



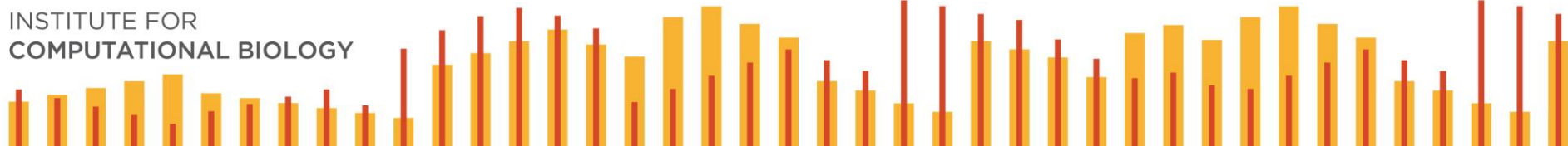
# MetroHealth/ICB Pilot Study (MIPs)

## Survey Summary and Comparisons:

- Higher willingness to participate  
70% v ~54%\*
- Same willingness to send data via phone app  
48% v ~43% (to share social media data)\*
- Same interest in genetic data  
77% (*at least*) v 74%\*

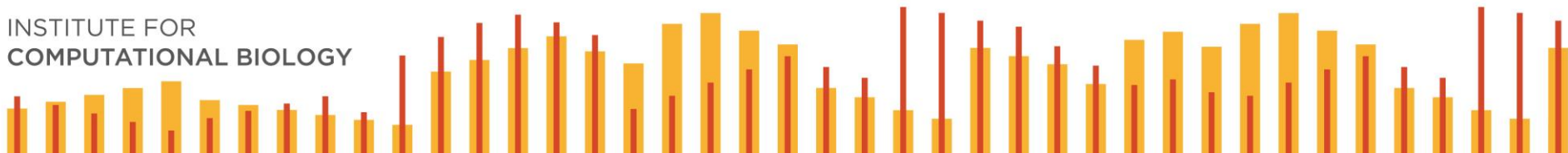


\*Kaufman et al *PLoS One* 11(8):e0160461 (2016)



# Summary

- Everyone seems excited about precision medicine research
- Ambitious efforts underway locally and nationally
- Outstanding questions remain
  - Will underrepresented groups participate?
  - Will the EHR data be useful?
  - Will wearable data be useful?
  - Will this really move precision medicine forward?



# Questions?

