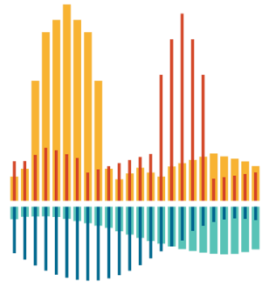


PRECISION MEDICINE RESEARCH

WHERE GENOMICS MEETS BIG DATA BIOMEDICAL INFORMATICS



INSTITUTE FOR
COMPUTATIONAL
BIOLOGY

November 3, 2017

Dana C. Crawford, PhD

Associate Professor

Population and Quantitative Health Sciences

Institute for Computational Biology

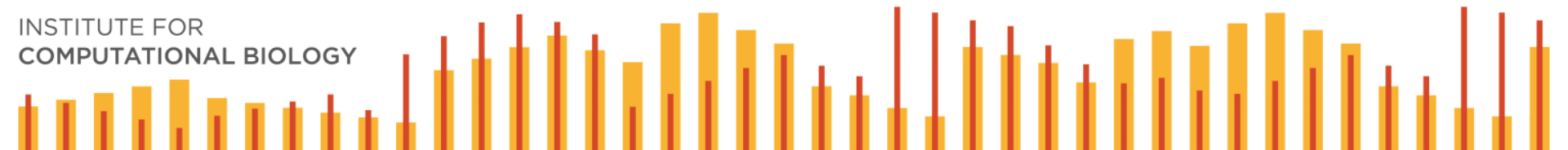
Precision Medicine

A New Initiative



Lancet 385:2448-2449 (2015)

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



Precision Medicine

The Concept



RELAX
AND BE
PATIENT

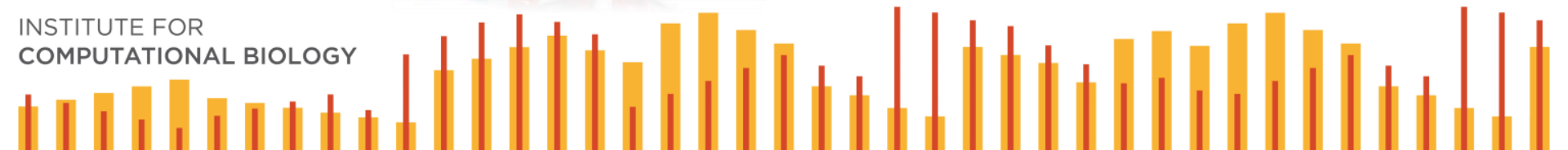
Right Patient

Right Drug



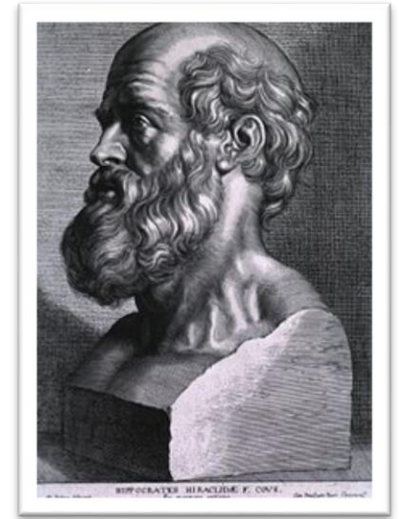
Right Dose

First Time



Precision Medicine

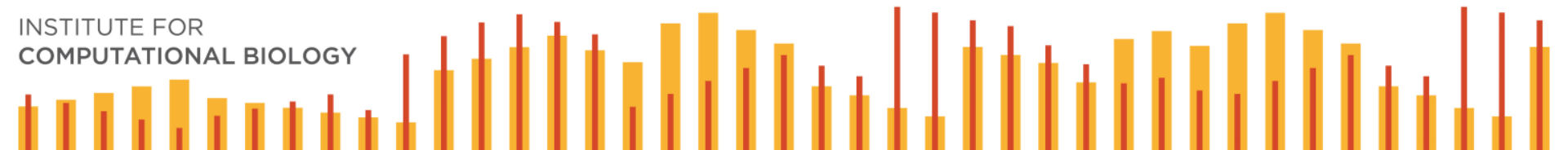
The Concept



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

- AKA Personalized Medicine
- Not a new concept

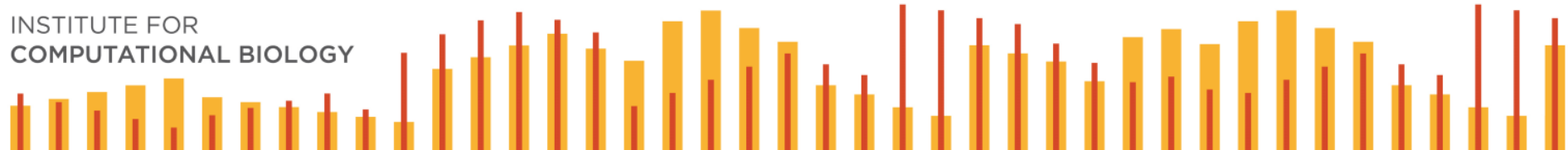
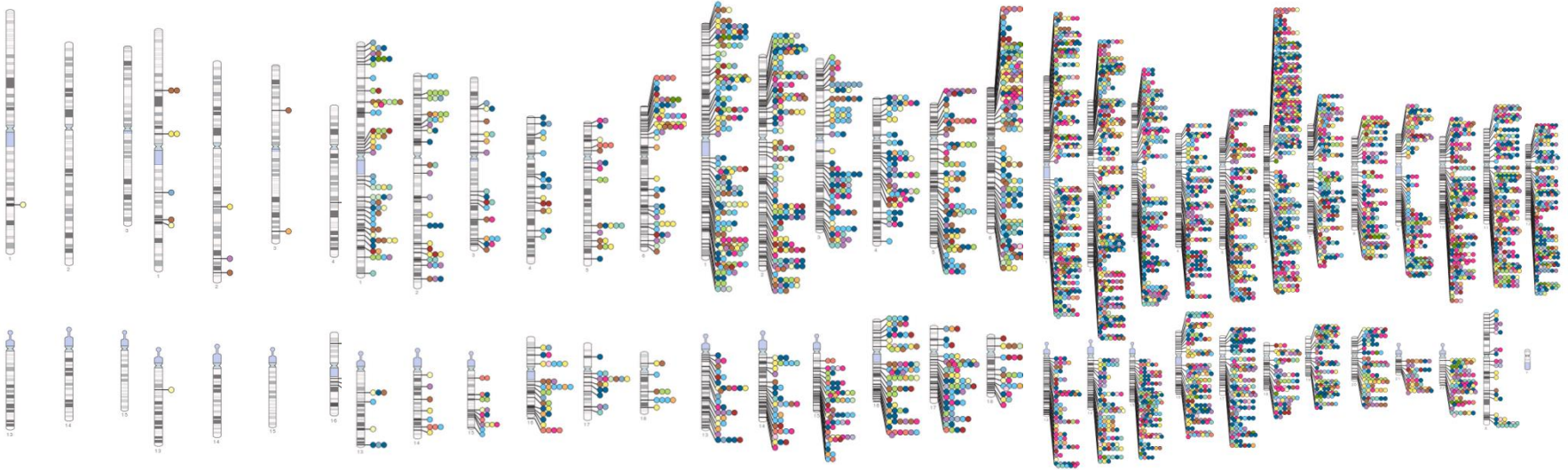
It's far more important to know what person the disease has than what disease the person has. – Hippocrates (c. 460 – c. 370 BC)



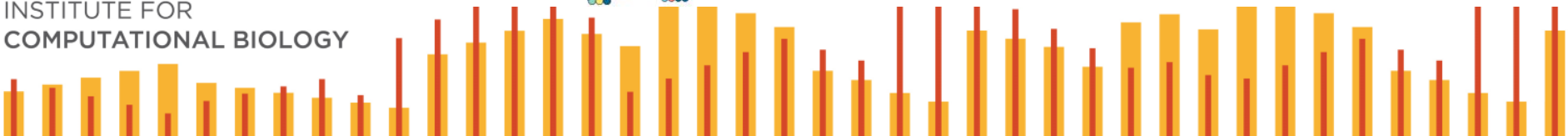
Precision Medicine

The Concept

Accelerated Genomic Discovery



GWAS as of 5/29:
2,940 publications
36,066 associated
SNPs (10^{-5})

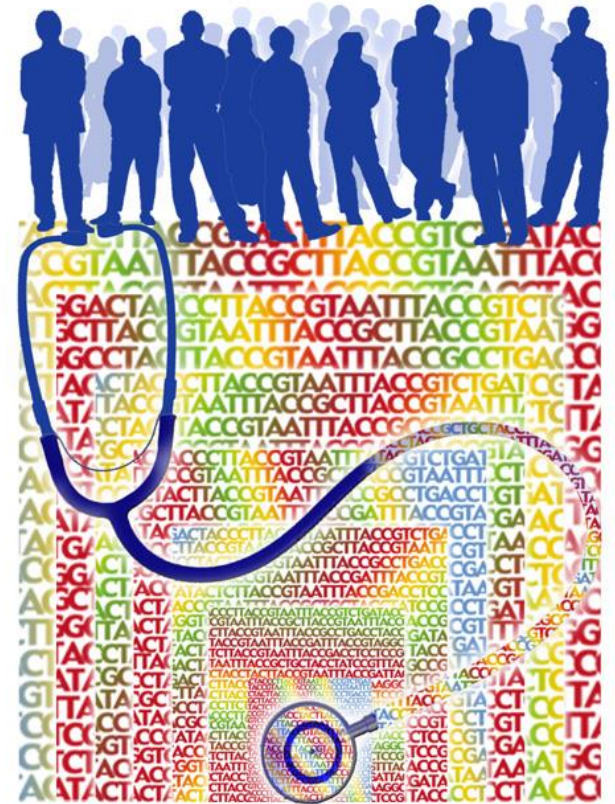


Precision Medicine

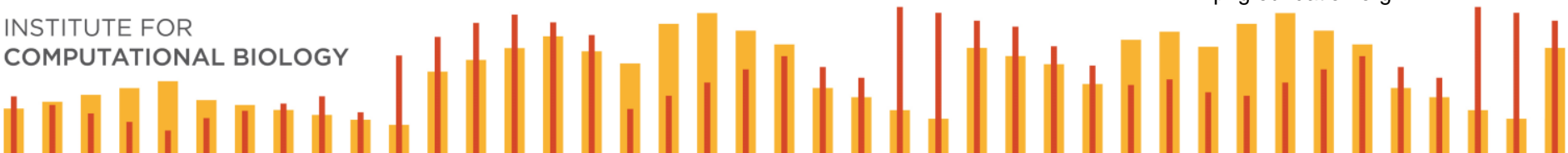
The Concept

Cost-effective Clinical Sequencing

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



www.phgfoundation.org

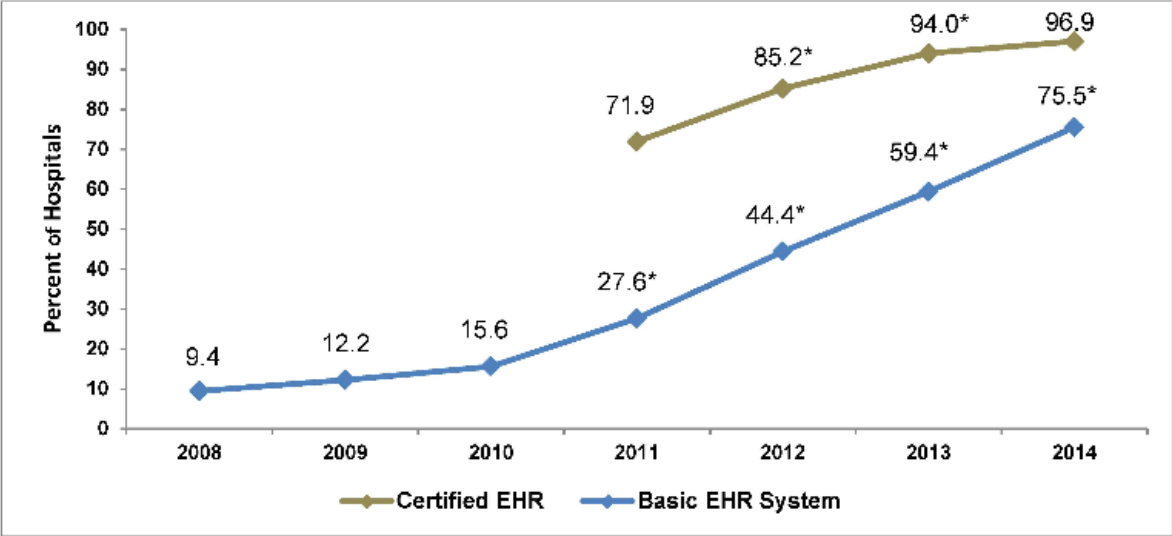


Precision Medicine

The Concept



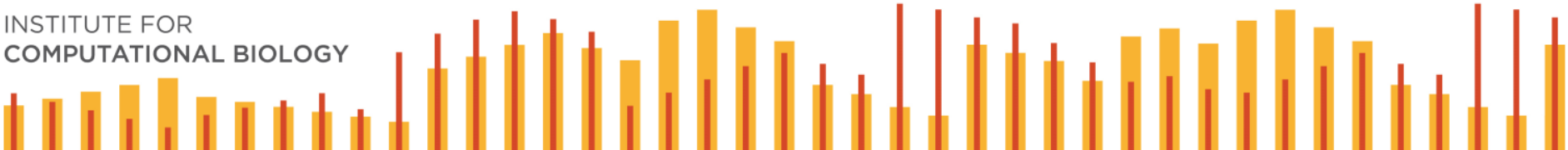
Figure 1: Percent of non-Federal acute care hospitals with adoption of at least a Basic EHR with notes system and possession of a certified EHR: 2008-2014



The rapid rise of EHRs:

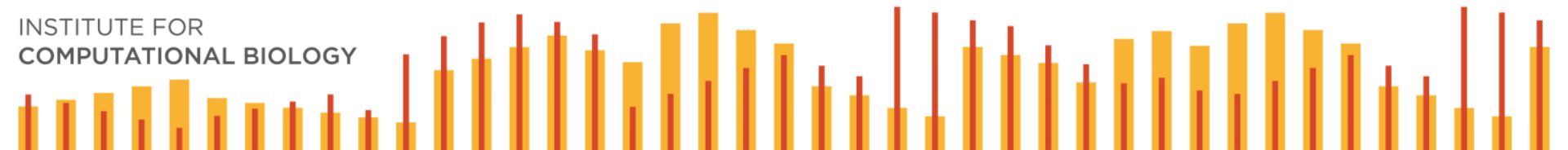
As of 2014, 3 out of 4 US hospitals have at least a basic EHR

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

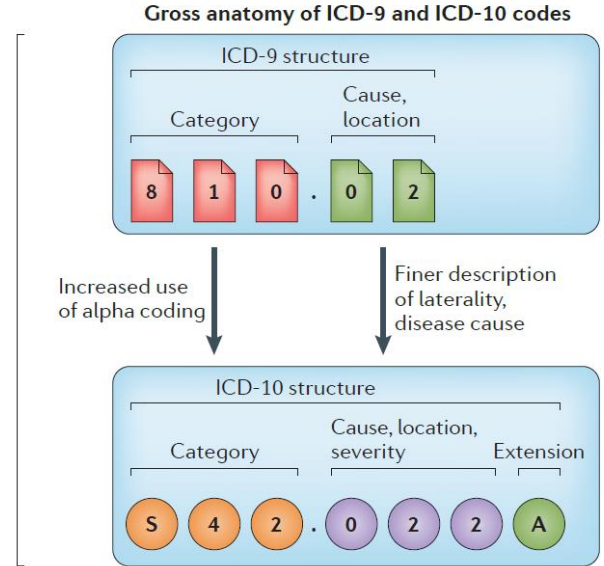


WHAT'S IN AN EMR?

- 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 x Medications x Labs x ICD9 Codes x Comments x

Open Filters Open Highlight

High Value Documents Other Documents Problem Lists

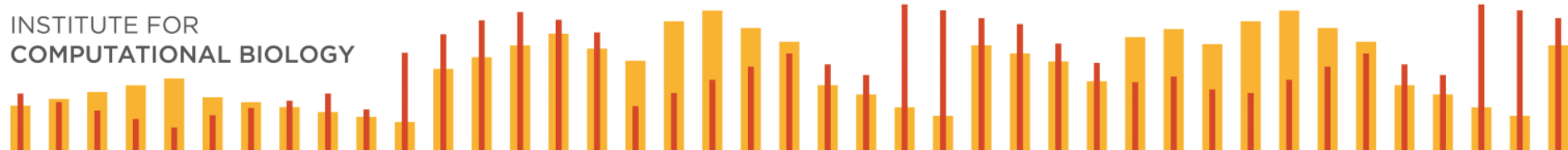
- Synthroid 0.15 mg qd
ALLERGIES:
- NKDA
FAMILY MEDICAL HISTORY:
- Mother(70): Deceased, CHF, hypertension
- Father(40): Deceased, stomach ulcers, colitis
- Siblings(x10): Four deceased, diabetes, CAD, MI, hypertension
- Daughter(39): Unknown blood disorder
SOCIAL HISTORY:
- Martial status: Single
- 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)?



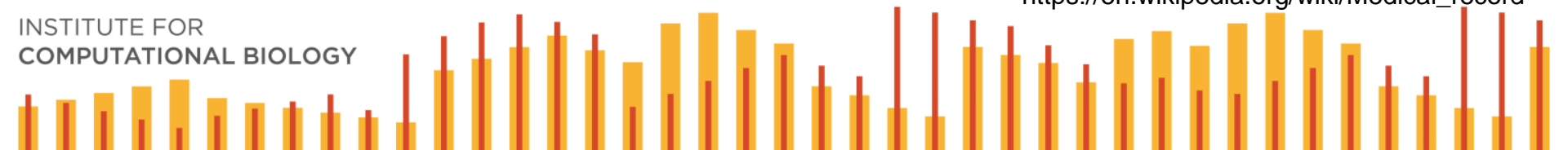
- A digital version of a patient's paper chart
- EHRs go beyond the data collected in the provider's office and include a more comprehensive patient history.
 - Designed to contain and share information from all providers involved in a patient's care.
 - Data can be created, managed, and consulted by 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

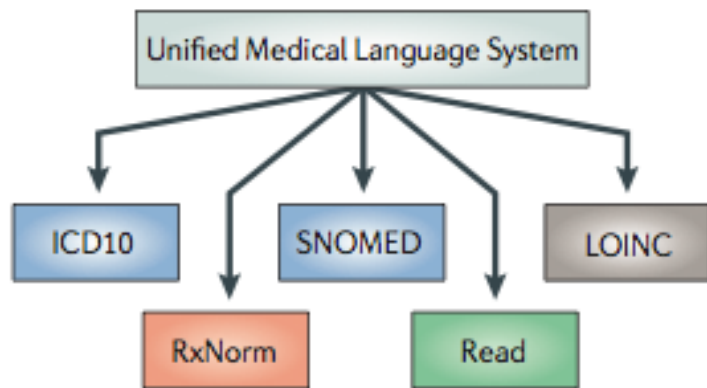


Mrs Jones is an 84-year-old African-American woman admitted from the emergency department with complaint of **crushing substernal** pain ... past medical history is significant for a 20 year history of **type 2 diabetes mellitus** controlled with **oral hypoglycaemics**, **2 ppd** history smoking ...

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

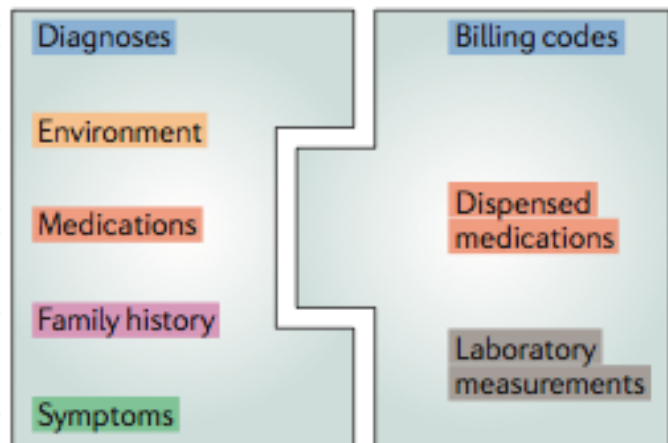
Mrs Jones was discharged on a 1,500 ml fluid restriction, **nitroglycerin** 0.4 mg/spray 1-2 spray po. **Aciphex** 20 mg (20 mg tablet DR take 1) **PO**

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

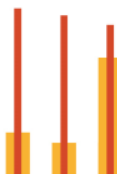


UMLS maps a data type from one vocabulary to another

Standardized vocabularies in which all EHR-data items are represented

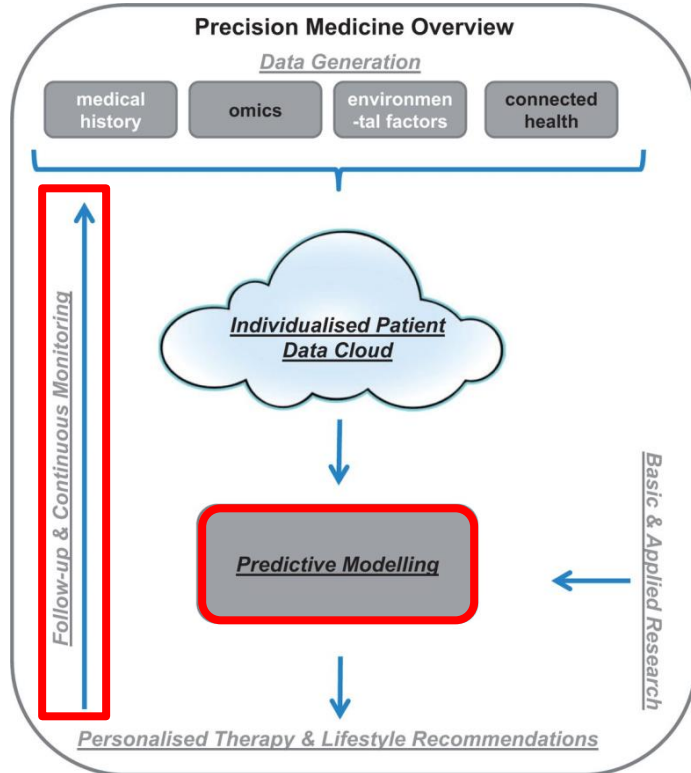


EHR-derived database composed of natively codified and NLP-codified data types



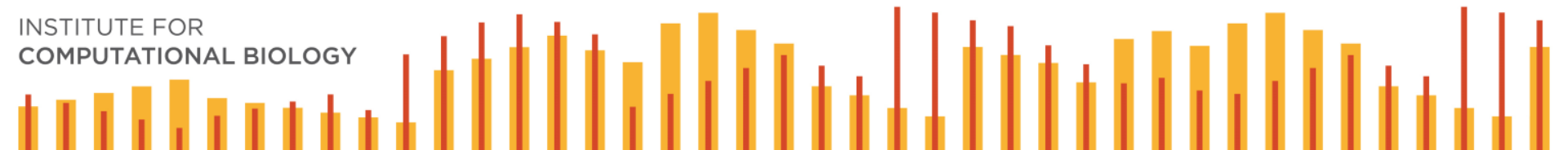
Precision Medicine

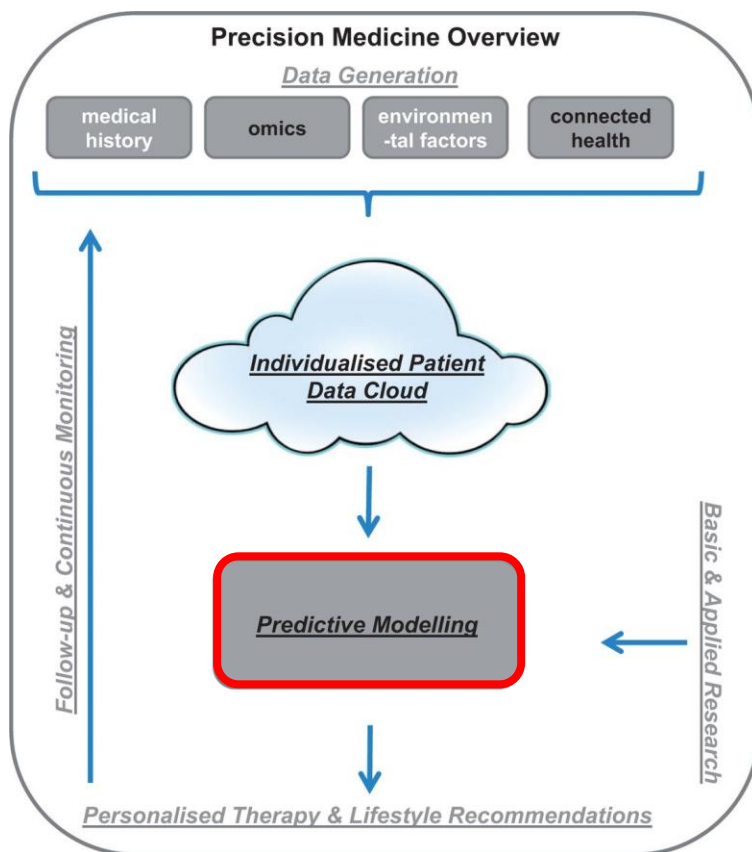
The Concept



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

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

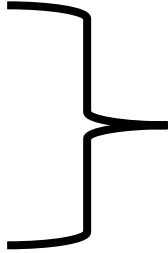




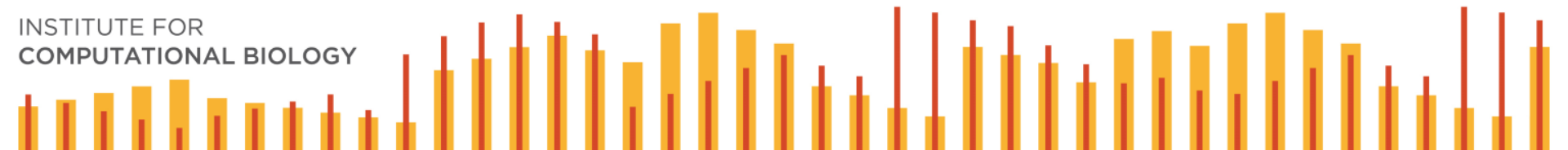
Duffy DJ *Brief Bioinform* 17(3):494-504 (2016)

Need More Data?

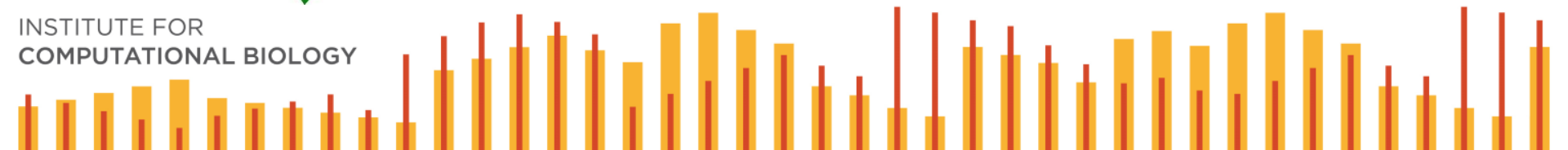
- More genetic data
- Other 'omic data
- Co-morbidities



- Longitudinal data
- Environmental exposures

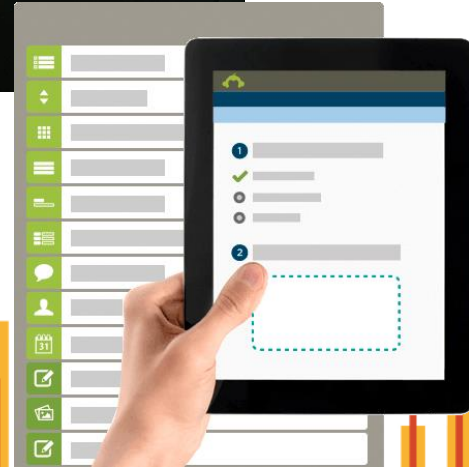


EHRs and Exposure Data Consistently Inconsistent



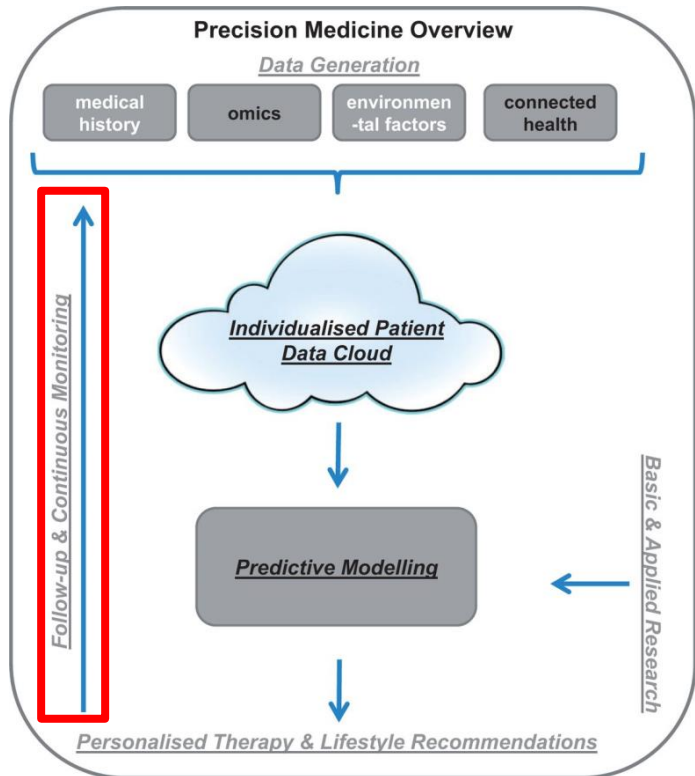
Emerging Tools for Patient-Provided Data

mHealth, Wearables, and Patient Portals



INSTITUTE FOR
COMPUTATIONAL BIOLOGY

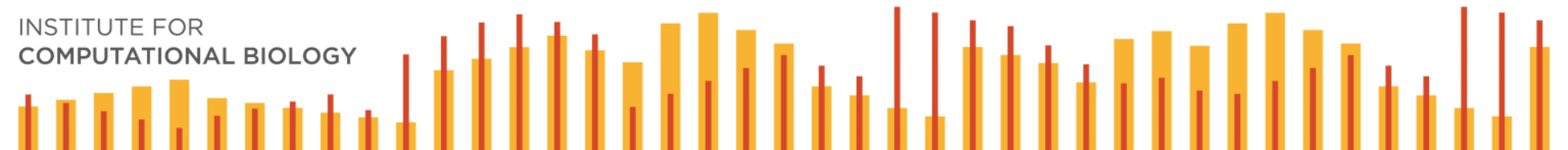




Duffy DJ *Brief Bioinform* 17(3):494-504 (2016)

EHRs and Follow-Up

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

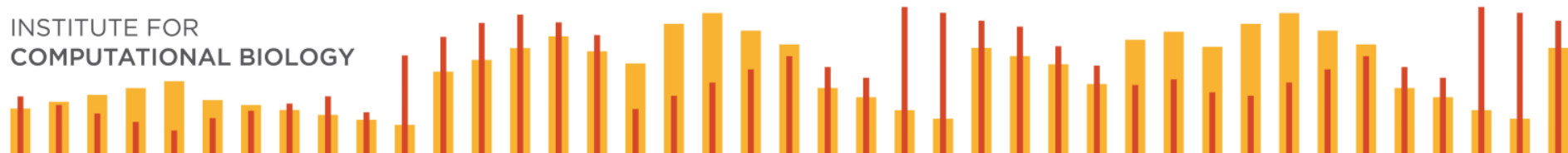
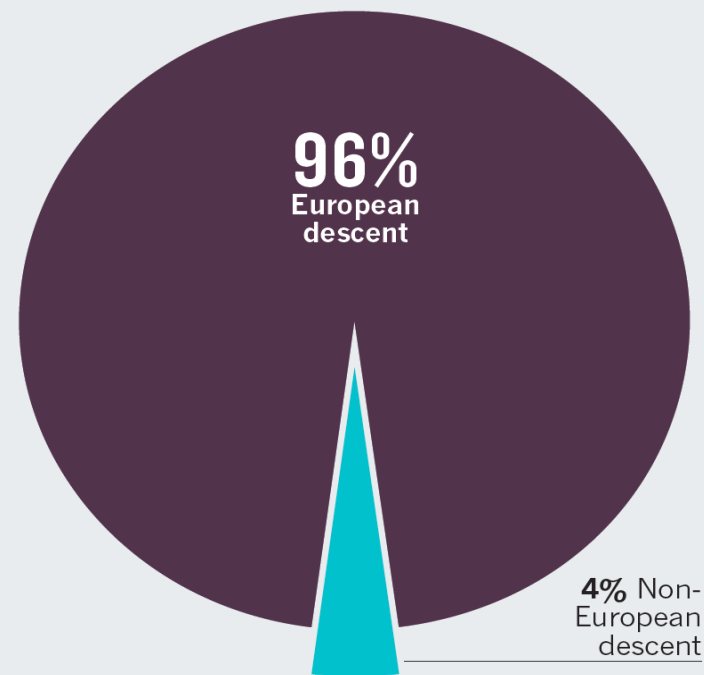


Genomic data based mostly on studies of European-descent

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

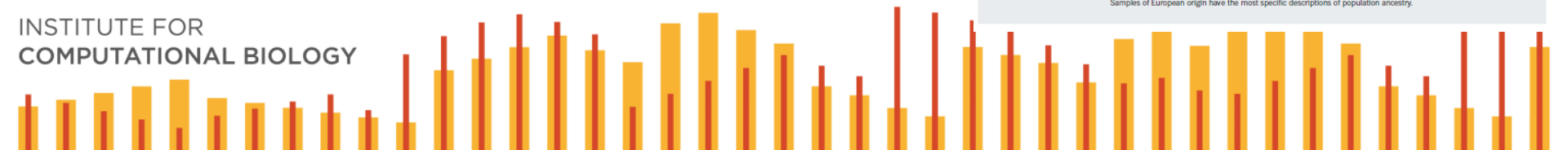
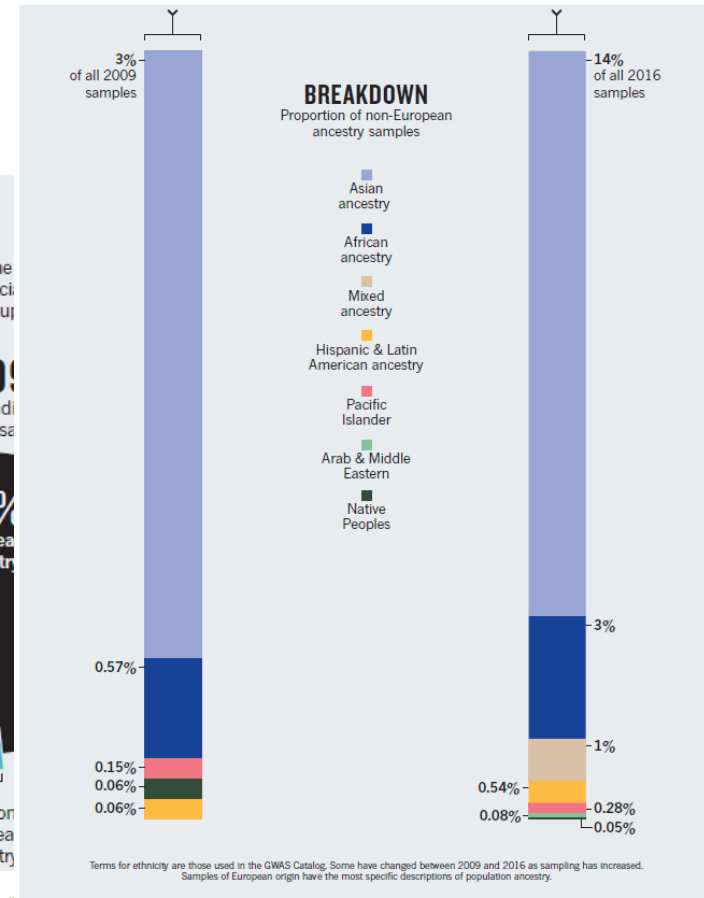
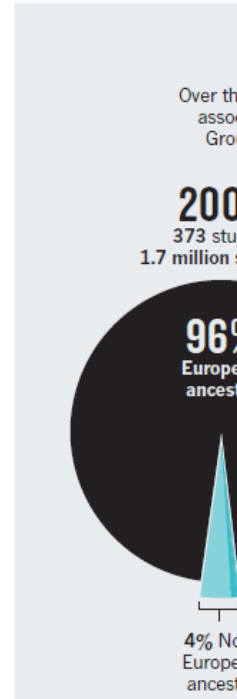
SAMPLING BIAS

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



Getting Better?

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



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^{1,2,*}, David J. Friedman^{1,3,*}, Michael D. Ross⁴, Laurence Lecordier⁵, Pierrick Uzureau⁵, Barry I. Freedman⁶, Donald W. Bowden^{7,8}, Carl D. Langefeld^{8,9}, Taras K. Oleksyk¹⁰, Andrea L. Uscinski Knob⁴, Andrea J. Bernhardt¹, Pamela J. Hicks^{7,8}, George W. Nelson¹¹, Benoit Vanhollebeke⁵, Cheryl A. Winkler¹², Jeffrey B. Kopp¹¹, Etienne Pays^{5,†}, Martin R. Pollak^{1,13,†}

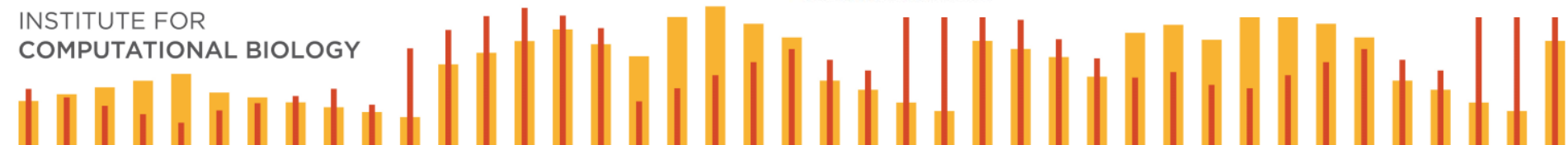
+ Author Affiliations

†To whom correspondence should be addressed. E-mail: mpollak@bidmc.harvard.edu (M.R.P.); 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



All of Us SM | The Precision Medicine Initiative[®]
THE FUTURE OF HEALTH BEGINS WITH YOU

1 Million Cohort

Large

Linked to EHR and payer databases

Oversample subgroups

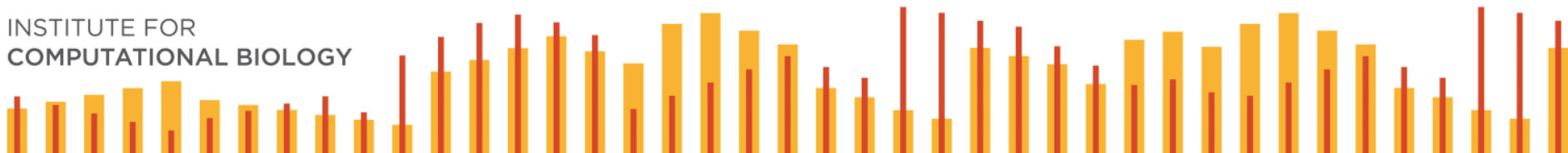
Recontact for return of results

Wide age range

Sophisticated lifestyle data collection

Rare and common disease

Extensive follow-up



Precision Medicine Initiative Cohort Program

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



Health Care Provider Organizations

10 Regional Centers,
including

University of Arizona,
Tucson in partnership
with Banner Health,
Phoenix
(PI: Dr. Akinlolu O. Ojo)



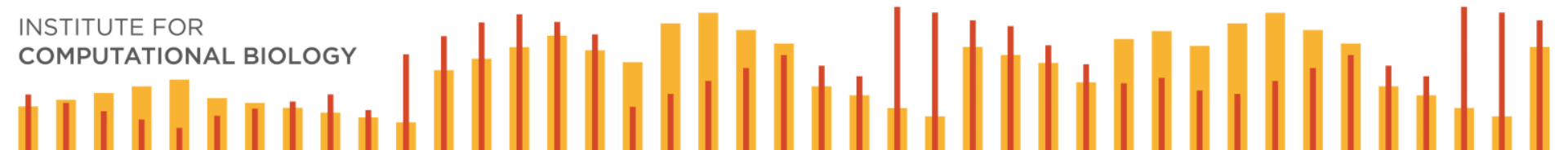
<https://allofus.nih.gov/about/program-components/health-care-provider-organizations>



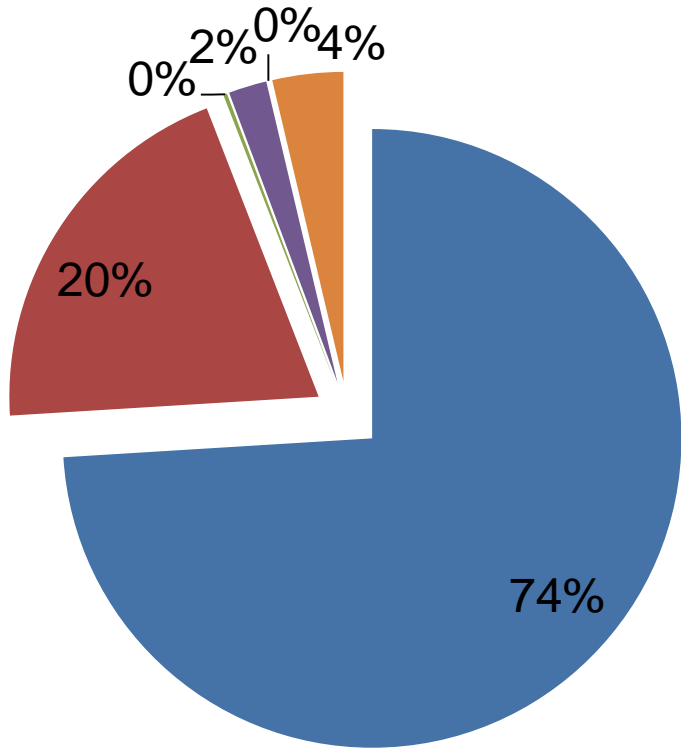
Local Precision Medicine CLE is All In!



INSTITUTE FOR
COMPUTATIONAL BIOLOGY



The CLE is All In! Diversity

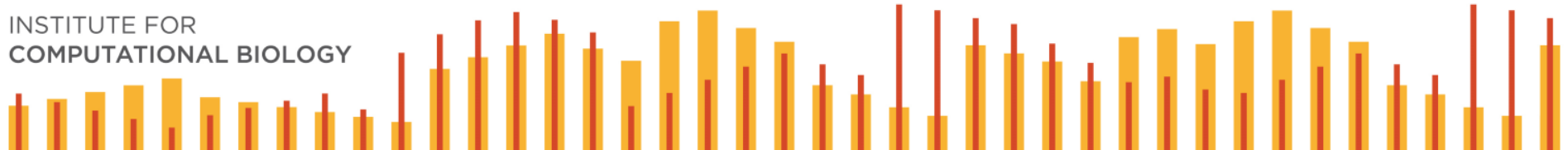


- White
- Black
- American Indian
- Asian
- Pacific Islander
- Other

28.3 % minority

4.7 % Hispanic

2010 US Census

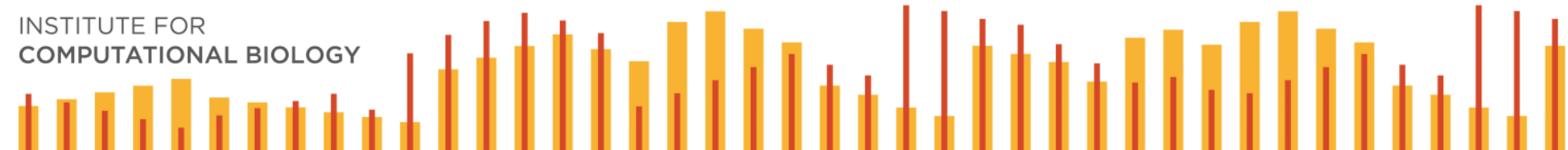


The CLE is All In!

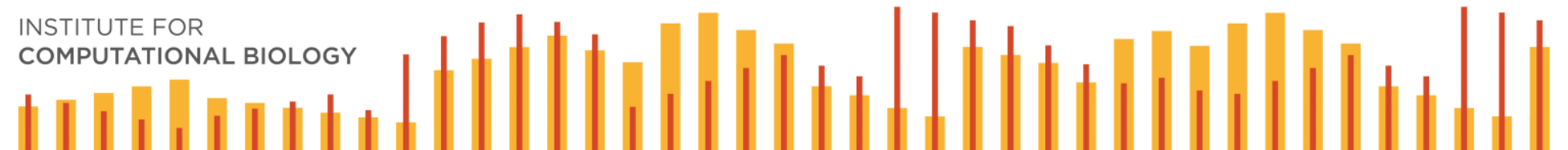
Diversity

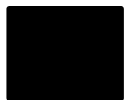
- Ranked #18 in population for US urban areas
- Median age is 40.5 years
- High urban (34.2%) versus suburban (12%) poverty rate

2010 US Census
2011 American Community Survey



	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

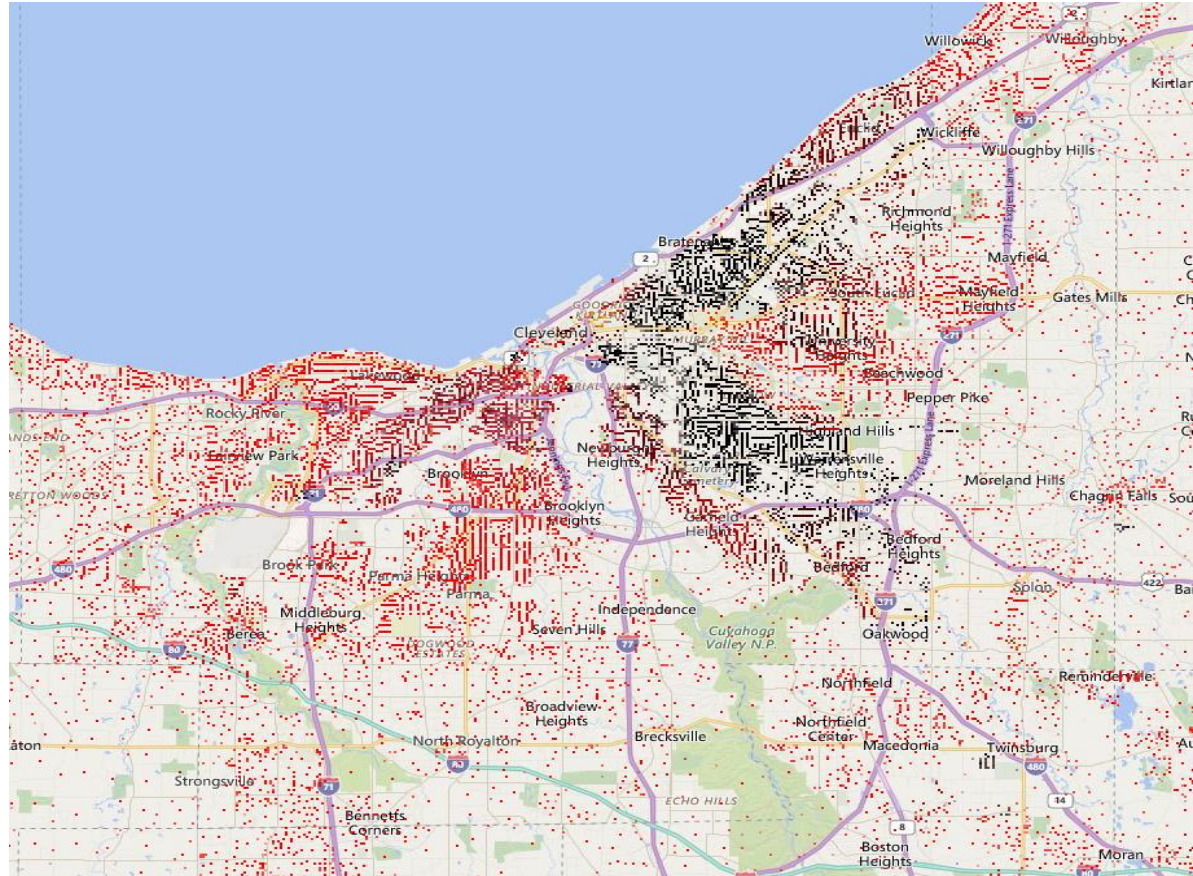




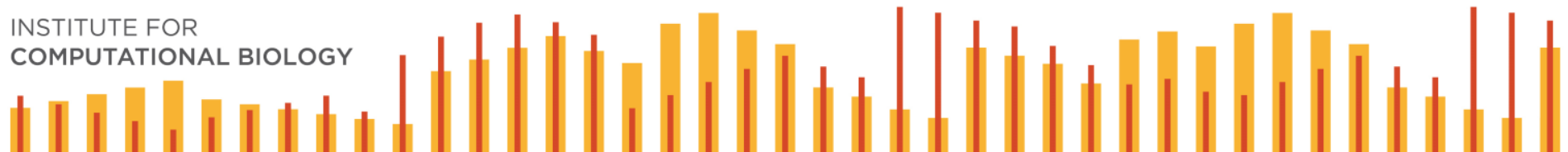
African
American



European
American



INSTITUTE FOR
COMPUTATIONAL BIOLOGY

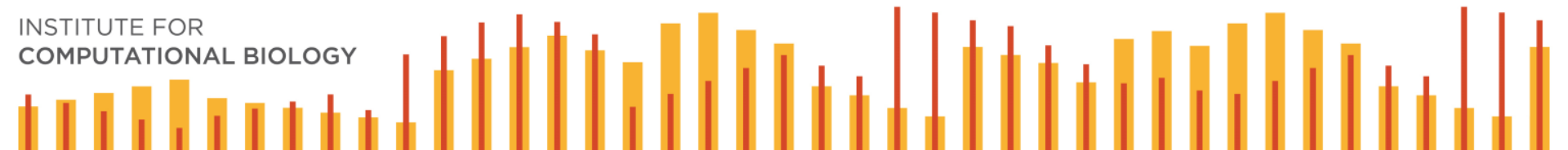


The CLE is All In!

Local Precision Medicine



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



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



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

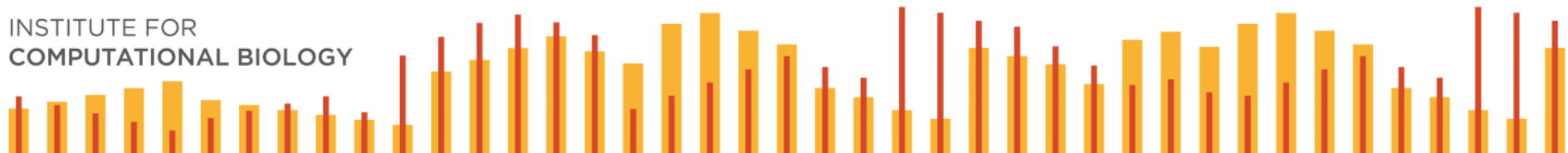


Epic installed in 2002; stage 7 (as of 2014)



VistA*** installed in 1970s (?)

INSTITUTE FOR
COMPUTATIONAL BIOLOGY



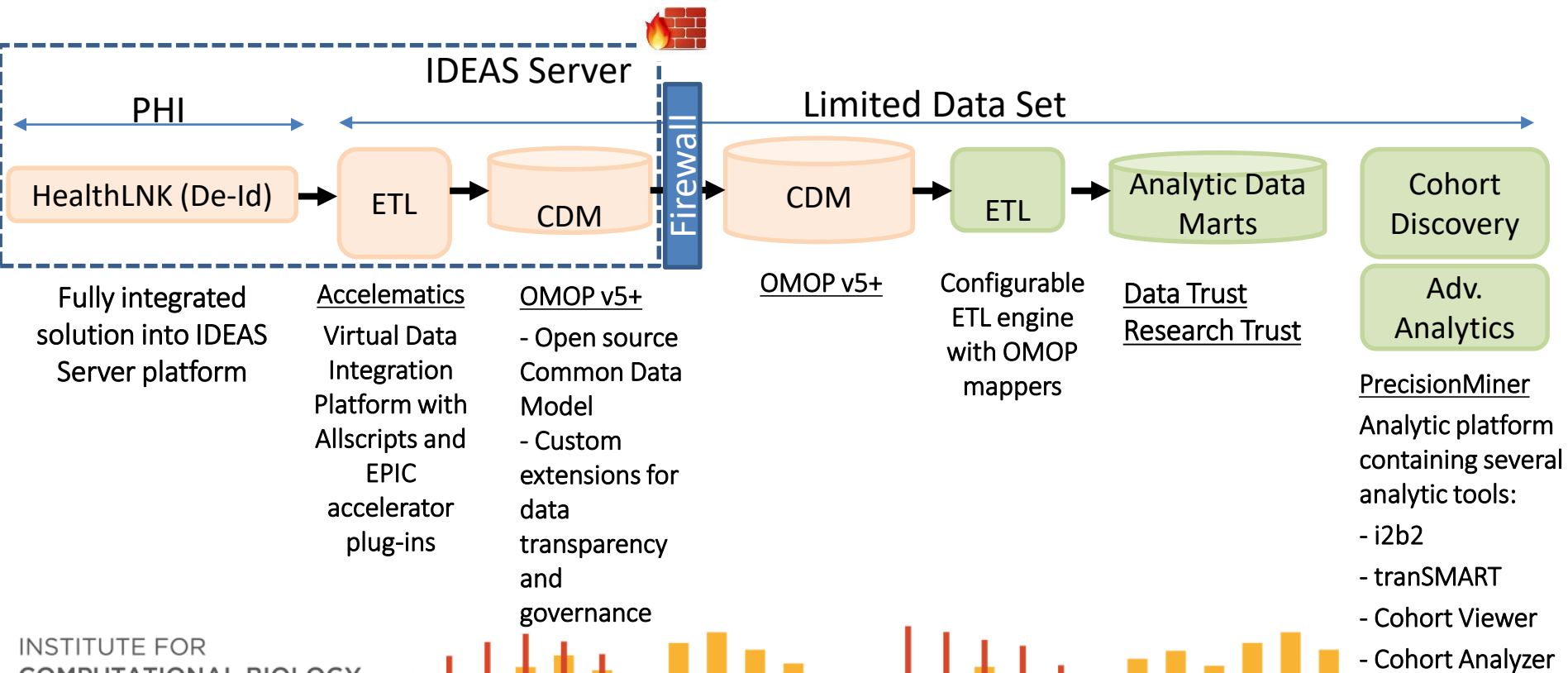
CLEARPATH:

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 Architecture



MetroHealth/ICB Pilot Study (MIPs)



David Kaelber, MD
Chief Medical
Informatics Officer



John Sedor, MD



John O'Toole, MD



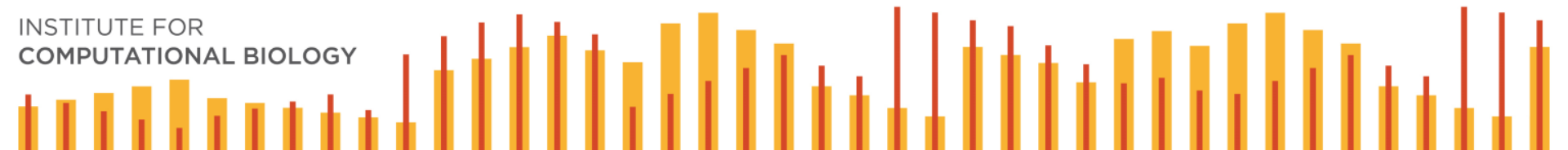
Will Bush, PhD, MS



Jessica Cooke
Bailey, PhD



INSTITUTE FOR
COMPUTATIONAL BIOLOGY



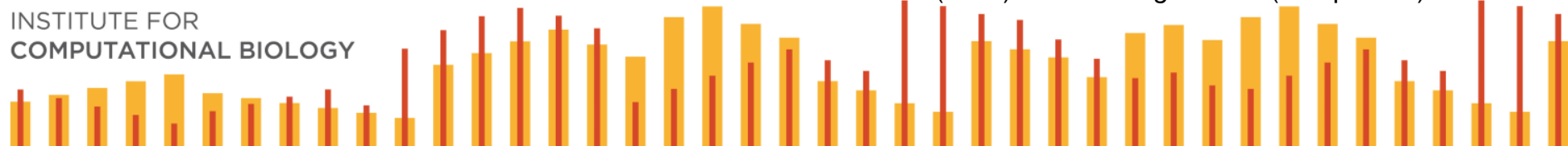
MetroHealth/ICB Pilot Study (MIPs)

- 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.

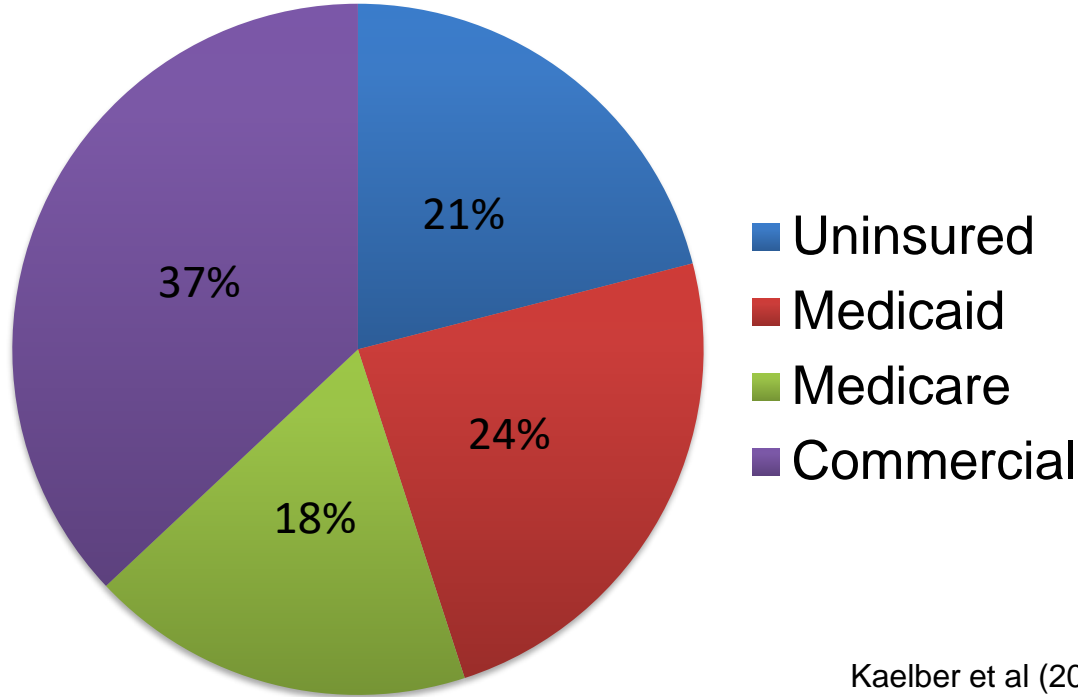


Kaelber et al (2013) *Am J Manag Care* 19(10 Spec No):SP337-43

INSTITUTE FOR
COMPUTATIONAL BIOLOGY

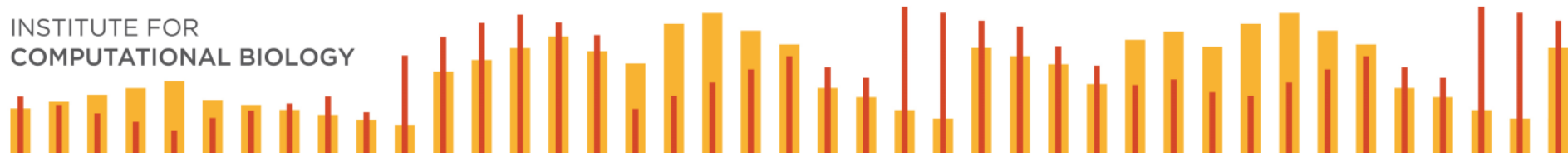


MetroHealth/ICB Pilot Study (MIPs)



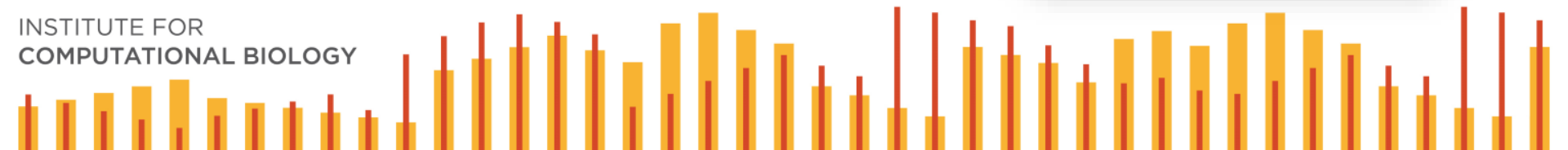
Kaelber et al (2013) *Am J Manag Care* 19(10 Spec No):SP337-43

INSTITUTE FOR
COMPUTATIONAL BIOLOGY



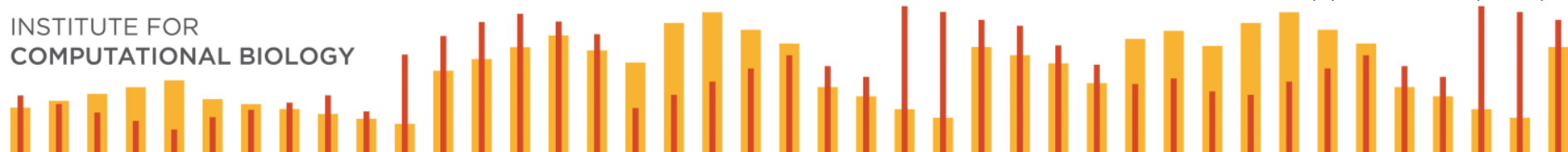
MetroHealth/ICB Pilot Study (MIPs)

- 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)



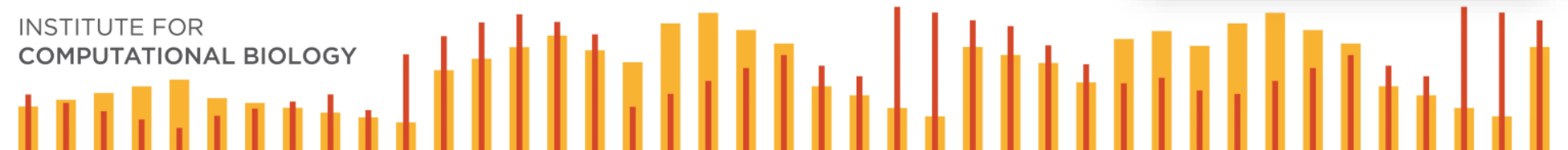
MetroHealth/ICB Pilot Study (MIPs)

Since February 2016:



101 patients surveyed at Nephrology Clinic

- >70% participation rate
- 50% African American
- 55% female
- Mean age 60.8 years
(range: 18 – 77 years)

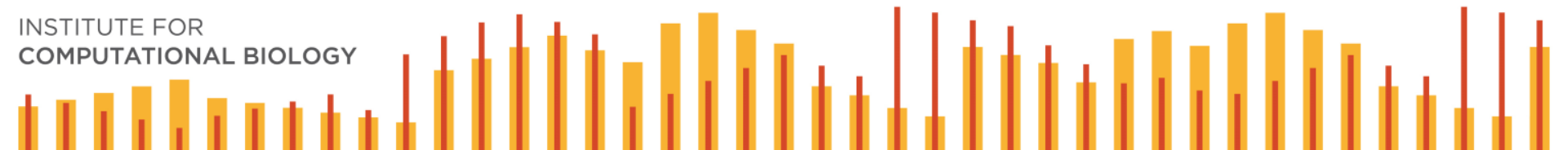


MetroHealth/ICB Pilot Study (MIPs)

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.



70% YES
68% YES



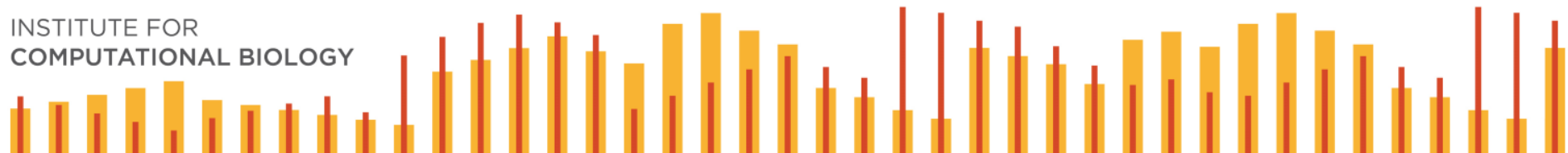
MetroHealth/ICB Pilot Study (MIPs)

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?



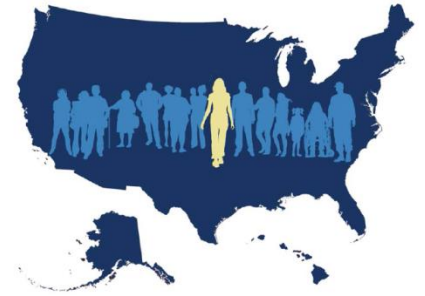
40% YES

46% YES



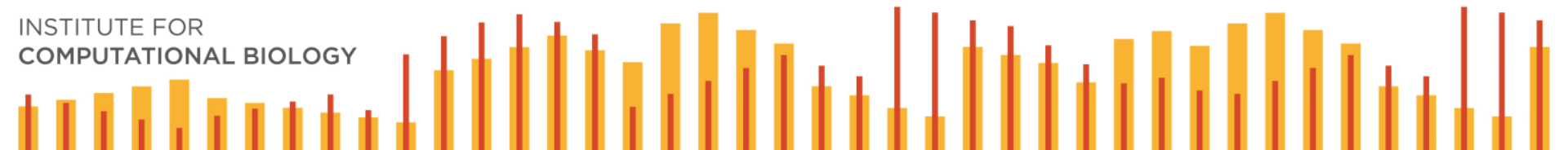
MetroHealth/ICB Pilot Study (MIPs)

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?



40% YES

52% 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

~19.8%

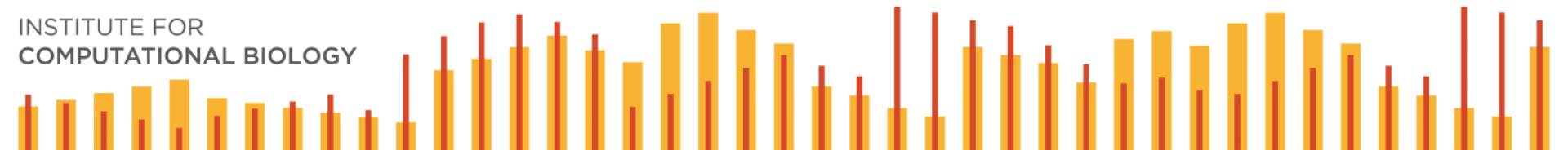
3

4

5

Very important

~51.5%



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

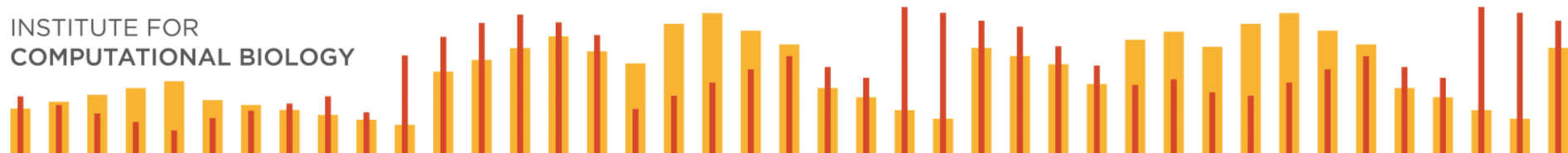
D – Information about your genes that has uncertain significance and won't change treatment

E – I do not want to receive any results



~29.7% EVERYTHING

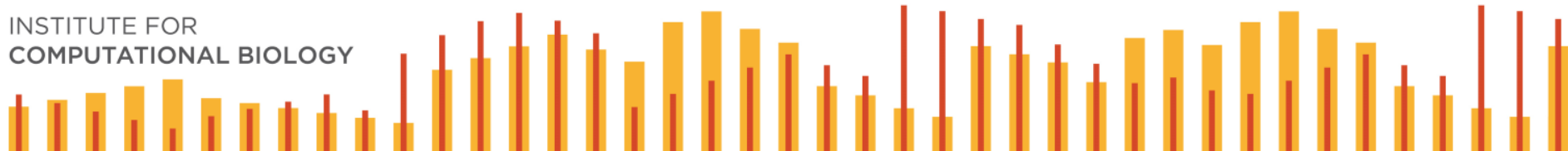
~10% NOTHING



MetroHealth/ICB Pilot Study (MIPs)

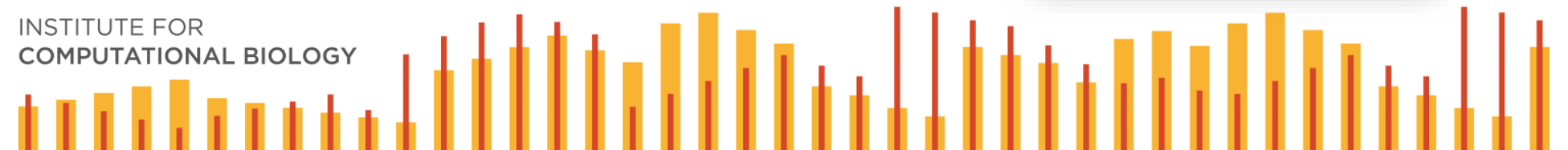
Survey Summary and Comparisons:

- Higher willingness to participate
70% v ~54%
- Same willingness to send data via phone app
46% v ~43% (to share social media data)
- More interested in genetic data
88% (*at least*) v 74%



MetroHealth/ICB Pilot Study (MIPs)

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



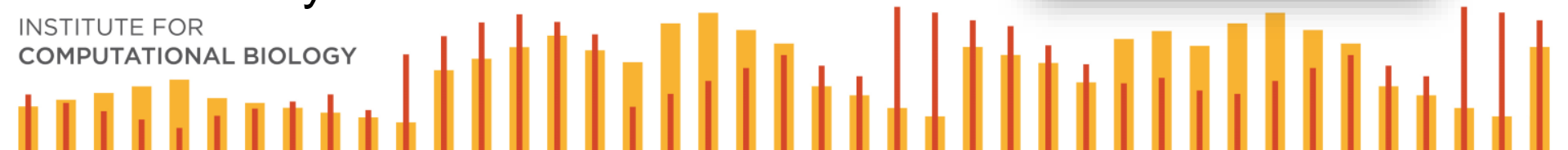
MetroHealth/ICB Pilot Study (MIPs)

Since February 2016:

135 donated biospecimens

3 EDTA tubes

- Two for aliquoting/extracting whole blood
- One to spin and save the plasma and buffy coat



MetroHealth/ICB Pilot Study (MIPs)

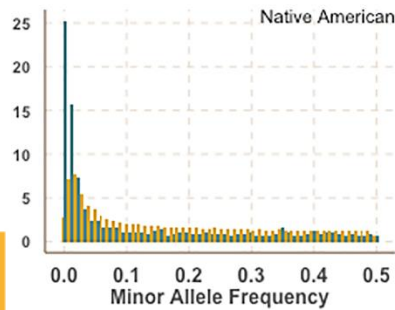
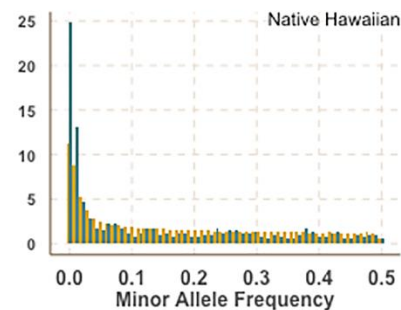
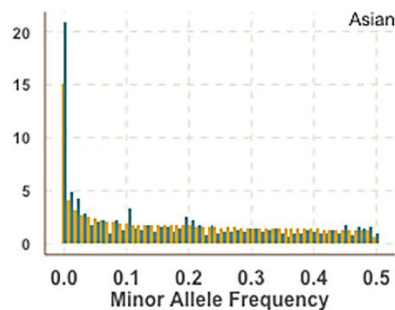
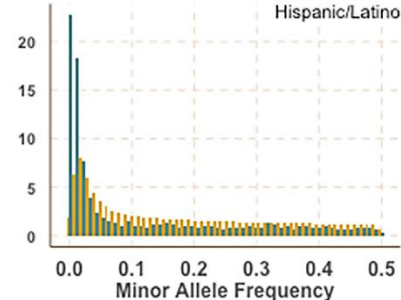
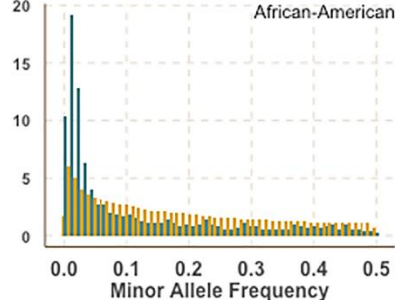
All genotyped on Illumina
Mega^{EX}

UNIVERSITY
OF MIAMI

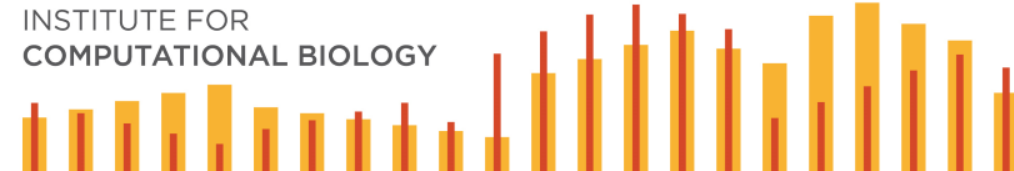


Bien et al *PLoS One* 11(12):e0167758 (2016)

INSTITUTE FOR
COMPUTATIONAL BIOLOGY



Bucket of Content
GWAS Backbone
Custom Content

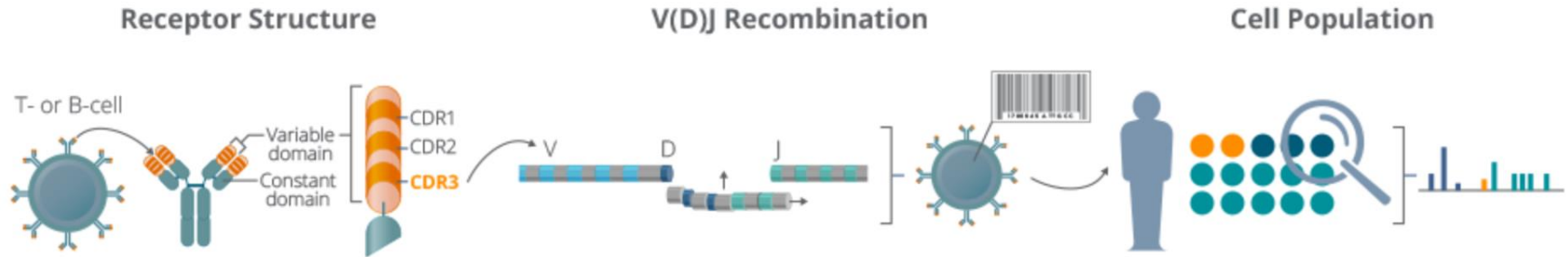


MetroHealth/ICB Pilot Study (MIPs)

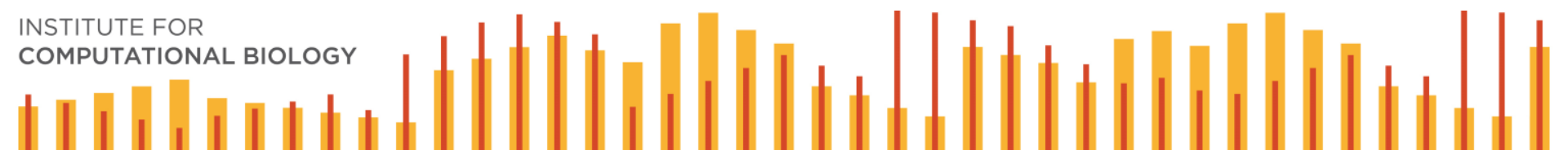
Subset sequenced using immunoSEQ



RECEPTOR PROFILING

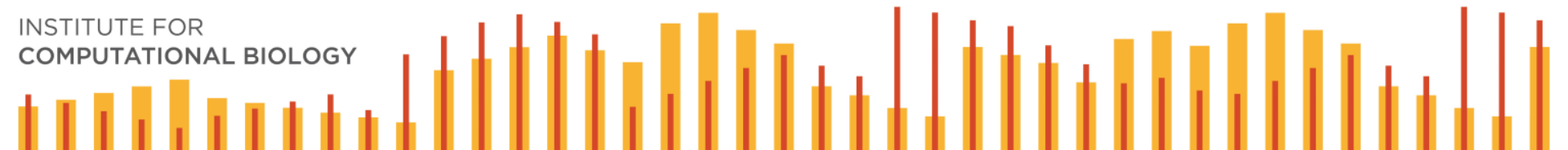


INSTITUTE FOR
COMPUTATIONAL BIOLOGY



Precision Medicine Research is Multidisciplinary

Biostatistics
Computer Genomics
Medicine Omics
Human Biomedical
Clinical Data Big
Bioethics Science
Bioinformatics
Informatics
Genetics



What Degrees and Training Do You Need?



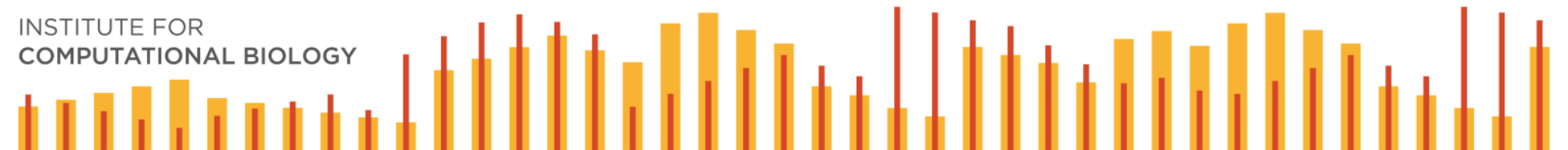
Genetics and Molecular Biology
Emory University
2000



EIS 2000-2002



Genome Sciences
University of Washington
2002-2006



What Degrees and Training Do You Need?

PhDs in Human Genetics

Logan



Research
Scientist
(Academia)

Janina



Industry
Scientist

Matt



Staff
Scientist
(Industry)

Jennifer



Technical
Writer

Nicole

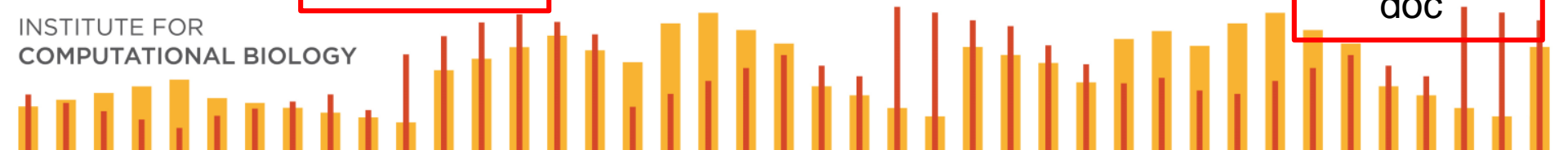


Staff
Scientist
(Industry)

Brittany



NIH Post-
doc



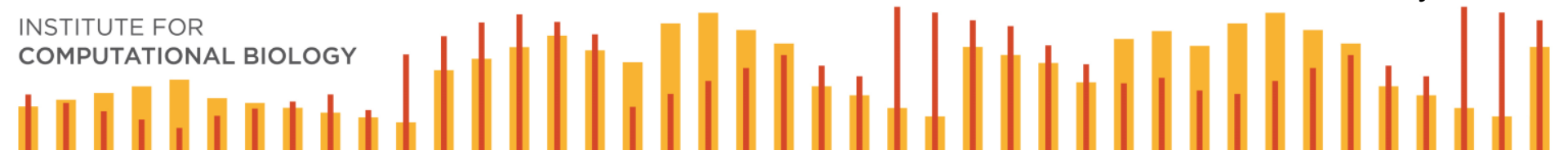
What Degrees and Training Do You Need? Bioethics and Human Genetics



Aaron Goldenberg, PhD
Associate Professor
Case Western Reserve University



Jessica Cooke Bailey, PhD
Assistant Professor
Case Western Reserve University



What Degrees and Training Do You Need?

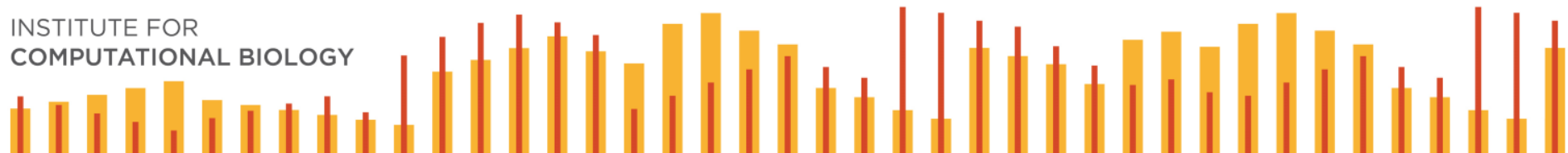
Bioinformatics, Computational Biology, Computer Science



Sarah Pendergrass, PhD, MS
Assistant Professor
Geisinger Health System



Will Bush, PhD, MS
Assistant Professor
Case Western Reserve University



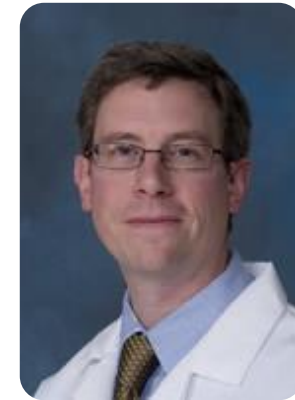
What Degrees and Training Do You Need? Biomedical Informatics and Doctorate of Medicine



David Kaelber, MD
Chief Informatics Officer
MetroHealth System

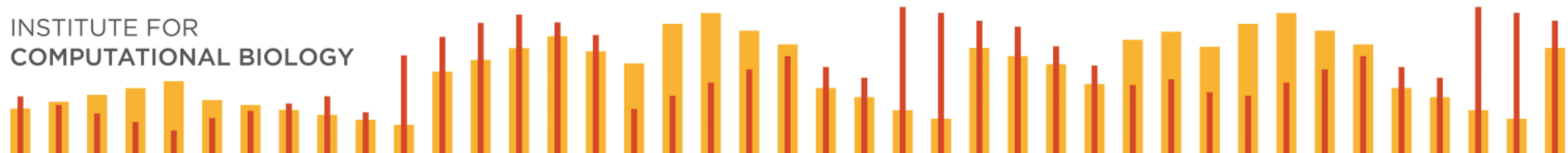


John Sedor, MD



John O'Toole, MD

Nephrologists
Cleveland Clinic



What Degrees and Training Do You Need?

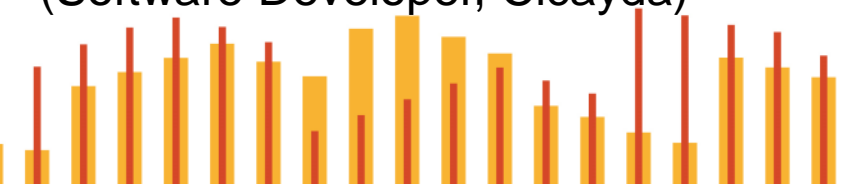
Biostatistics, Data Management/Architects

Kristin Brown-Gentry, MS
(Health Outcomes Scientist, Magellan Health)

Robert Goodloe, MS
(Consultant Statistician, Eli Lilly)

Bob McClellan, BS
(Senior Application Developer,
Vanderbilt University Medical Center)

Jonathan Boston, BS
(Software Developer, Cicayda)



CWRU/ICB Collaborators



Jonathan L.
Haines, PhD



William S. Bush,
PhD, MS



Jill Barnholtz-
Sloan, PhD



Jessica Cooke
Bailey, PhD



Devin Tian, PhD



Mark Beno, MSM

MetroHealth Collaborators

David Kaelber, MD

John Sedor, MD

John O'Toole, MD



INSTITUTE FOR
COMPUTATIONAL BIOLOGY



Questions?

