JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

JCPH Forum
Sep 2017

Population Health Informatics

Challenges, Opportunities, and Case Studies

Hadi Kharrazi, *MHI MD PhD*

Johns Hopkins University
School of Public Health – Department of Health Policy and Management
School of Medicine – Division of Health Sciences and Informatics

kharrazi@jhu.edu

Overview

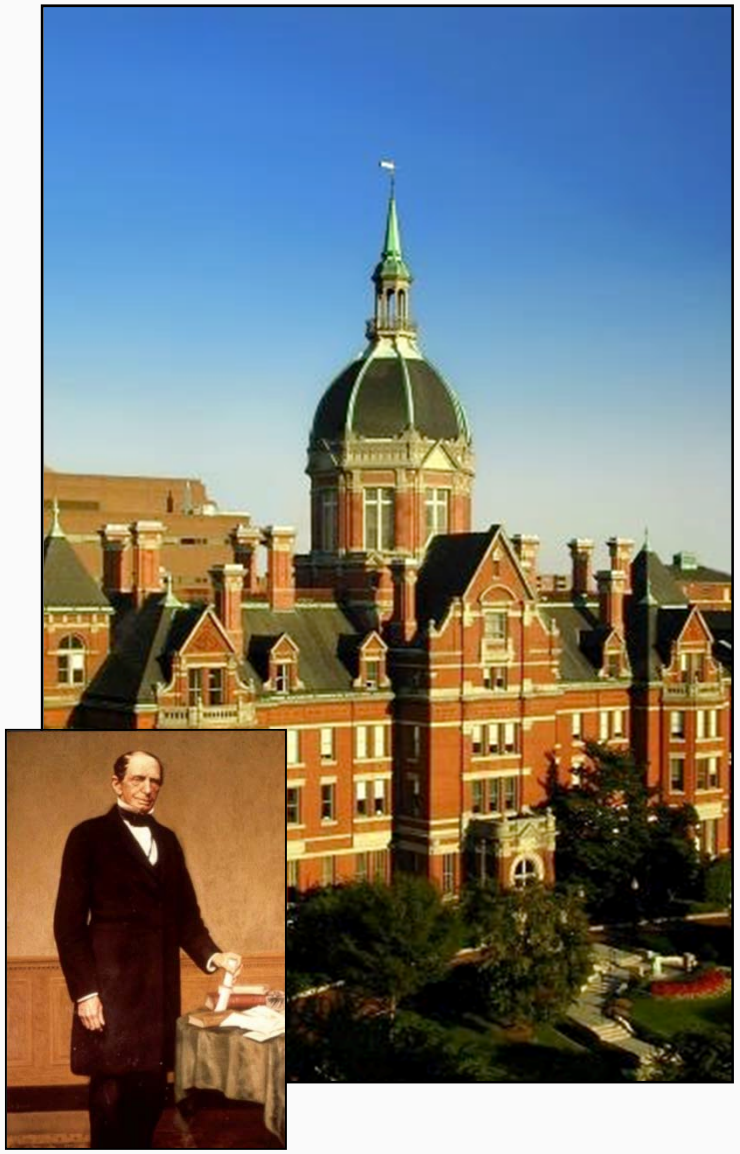
- JHU CPHIT
  - JHSoM
  - JHSPH
  - CPHIT
- Population Health Informatics
  - Emerging Field
  - Data Sources
  - Data Types
  - Data Spectrum
  - Data Analytic
  - Drivers *(incentives; mandates; standards)*
  - Results of the National Population Health Informatics Workshop
- CPHIT Portfolio
  - EHR-based Utilization Prediction (eACG)
  - Predicting Elderly Falls
  - Geriatric Frailty & EHRs
  - VHA Obesity Trajectory
  - Population Health Metrics
  - Pop Health IT Curriculum and Workforce Training
- Discussion
  - Challenges & Opportunities

Johns Hopkins University Center for Population Health IT (CPHIT)

JHU → JHSoM (School of Medicine)

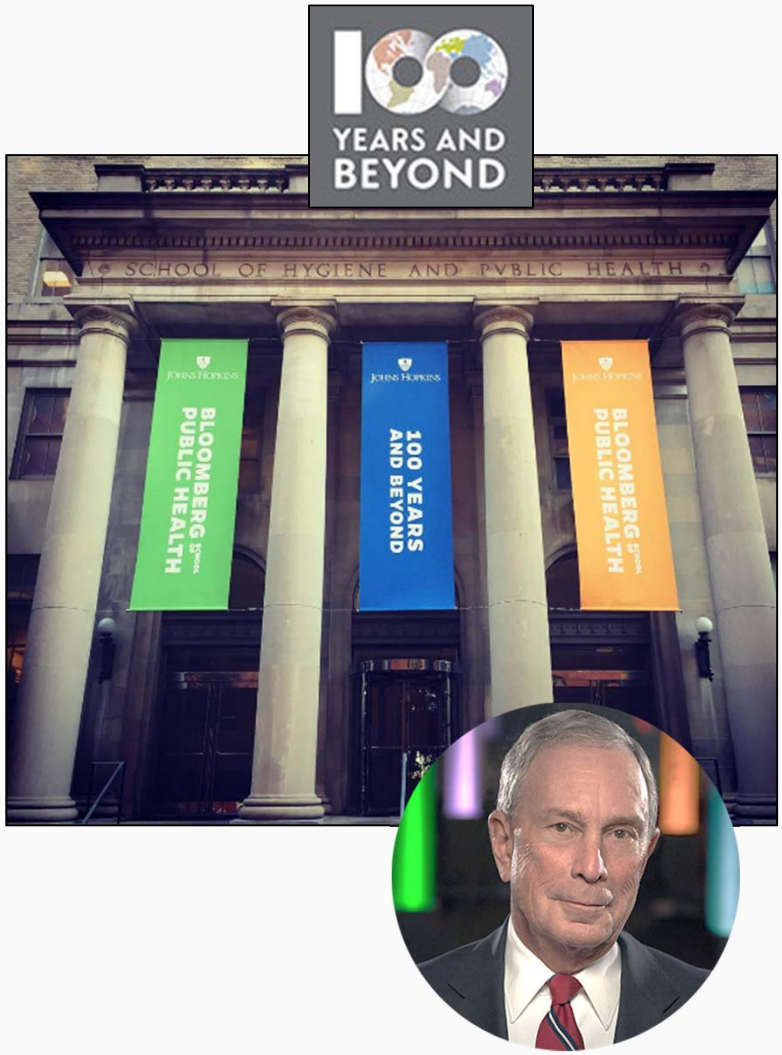
- Johns Hopkins, the Quaker merchant, banker and businessman, left $7 million in 1873 to create The JHU/Hospital to create new models and standards for medical education and health care.

- JHU/JHM/JHSoM
  - 6 hospitals + 35 outpatient clinics
  - $8 billion in operating revenues
  - 2800 full time faculty + 1200 part time faculty (40k with staff members)
  - 2.8 million-plus annual outpatient visits
  - 360,000-plus annual ED visits
  - 115,000-plus annual hospital admissions
  - $2 billion in total research and development spending based on NSF rankings (highest for 35 years) → $420 mil NIH grants to SoM
  - Managed care health plan: JHHC (368k lives)



JHU → JHSPH (School of Public Health)

- JHU/JHSPM
  - Johns Hopkins Bloomberg School of Public Health (founded 1916)
  - 670 faculty + 709 part time
  - Research in 130 countries
  - Research budget: $500+ mil
  - Degrees: 12 graduate including the largest MPH
  - Centers: 60+
  - Highlights:
    - First School of Public Health in the US
    - Largest school of public health in the world
    - Receives 20 percent of all grants and contracts awarded to the 57 accredited U.S. schools of public health
    - Ranked No. 1 by U.S. News & World Report since 1994



JHU → CPHIT

The Johns Hopkins

Center for Population Health Information Technology

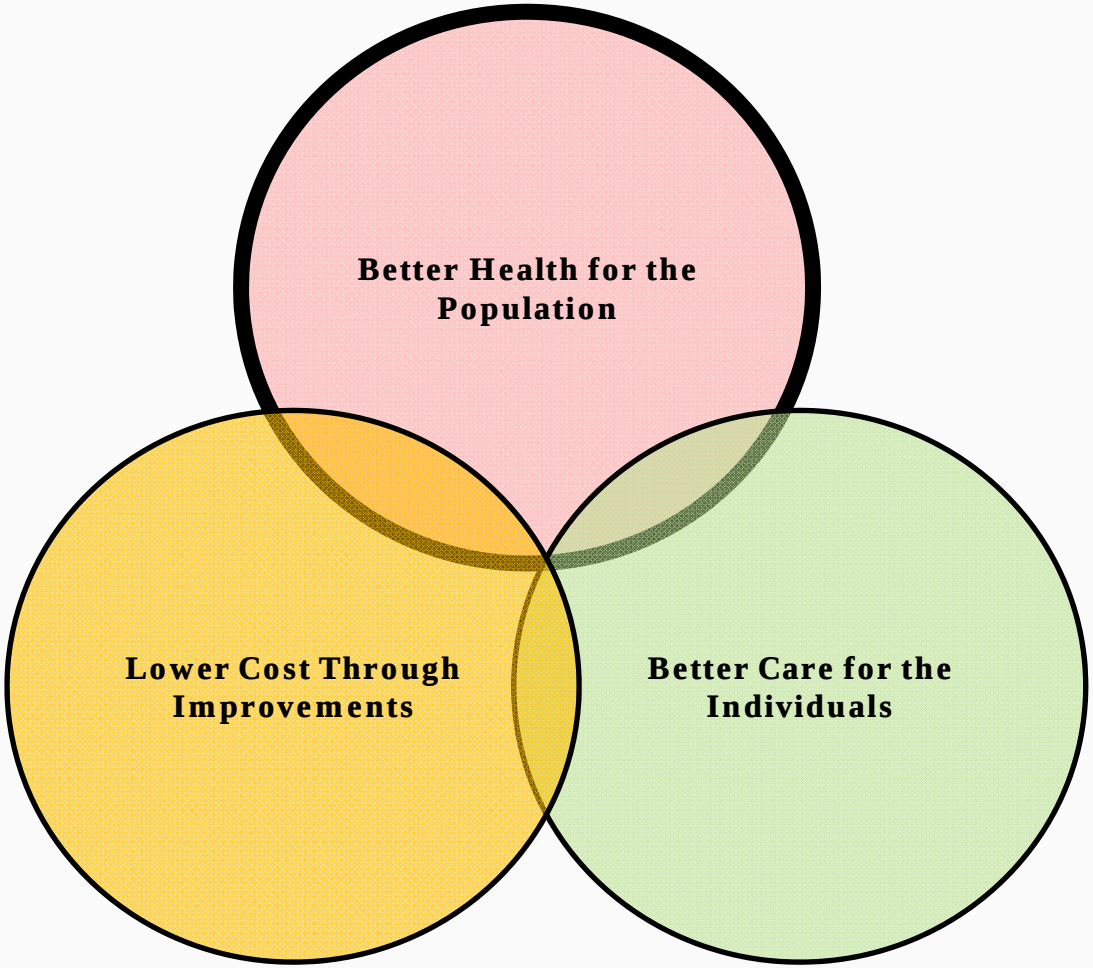
(CPHIT, or "see-fit")

- The **mission** of this innovative, multi-disciplinary R&D center is to improve the health and well-being of populations by advancing the state-of-the-art of Health IT across public and private health organization.
- CPHIT **focuses** on the application of electronic health records (EHRs), mobile health and other e-health and HIT tools targeted at communities and populations.
- Director: Dr. Weiner
- Research Director: Dr. Kharrazi

www.jhsph.edu/cphit

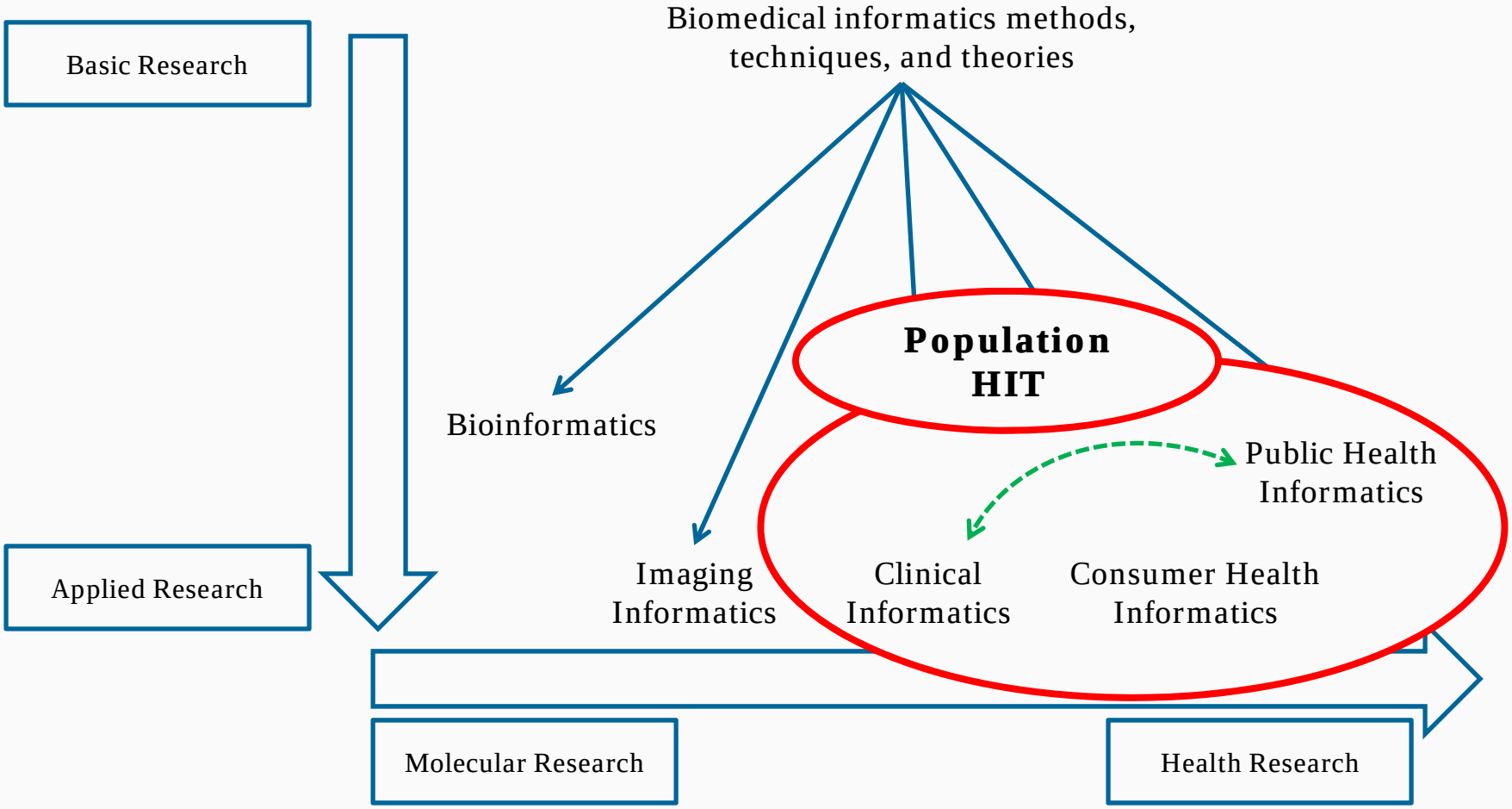
Population Health Informatics

Population Health Informatics → Emerging Field



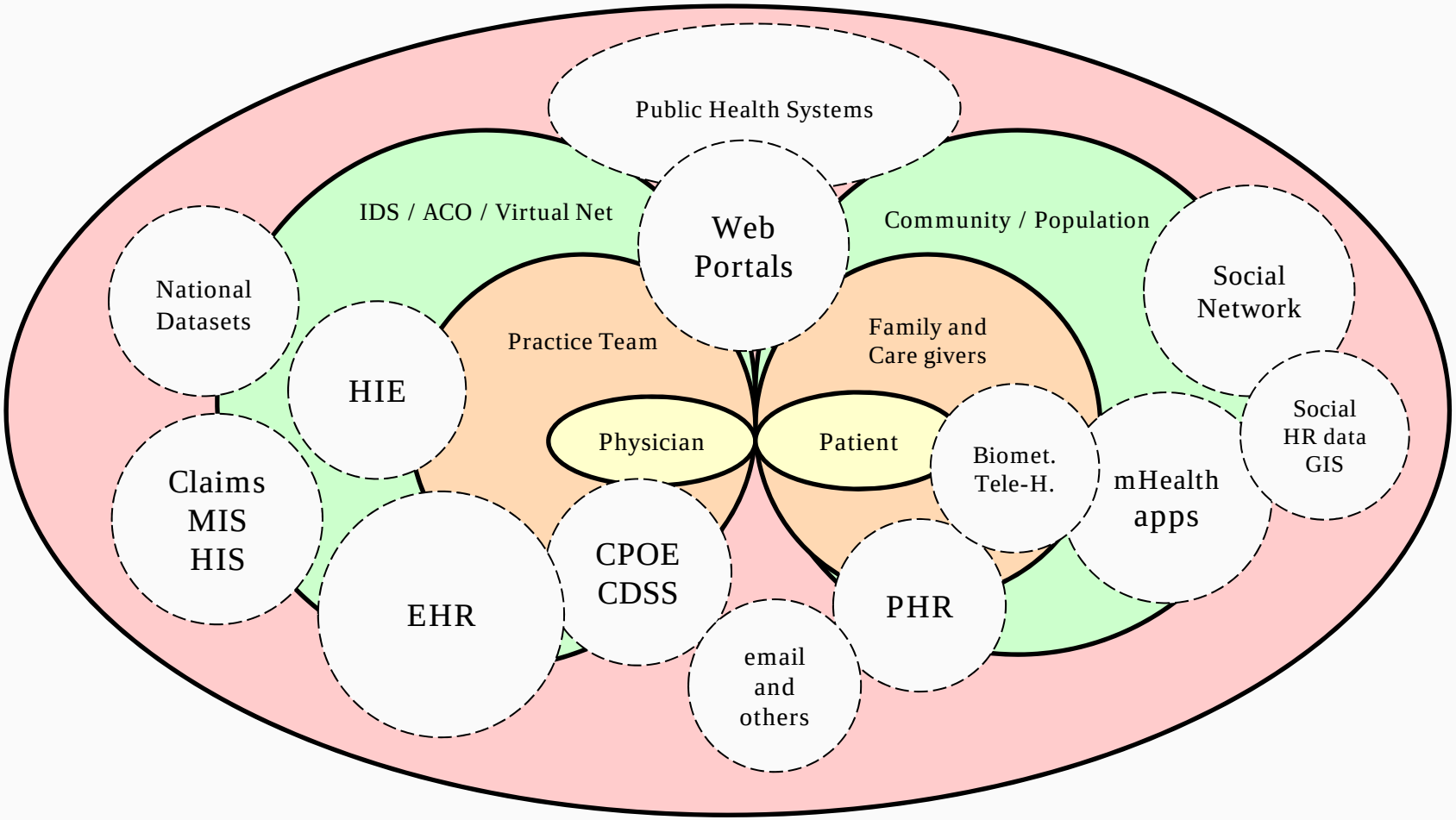
Triple Aims developed by the Institute for Healthcare Improvement (IHI)

Population Health Informatics → Emerging Field



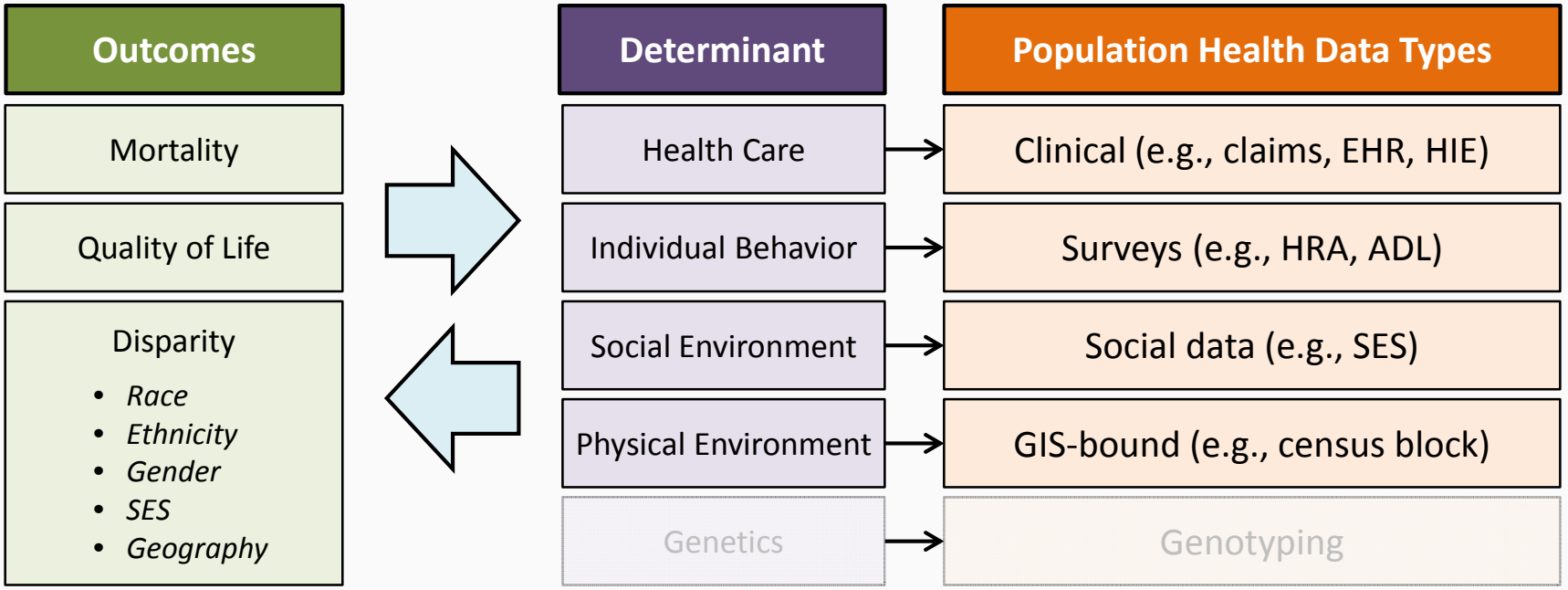
Biomedical informatics as a basic science

Population Health Informatics → Data Sources



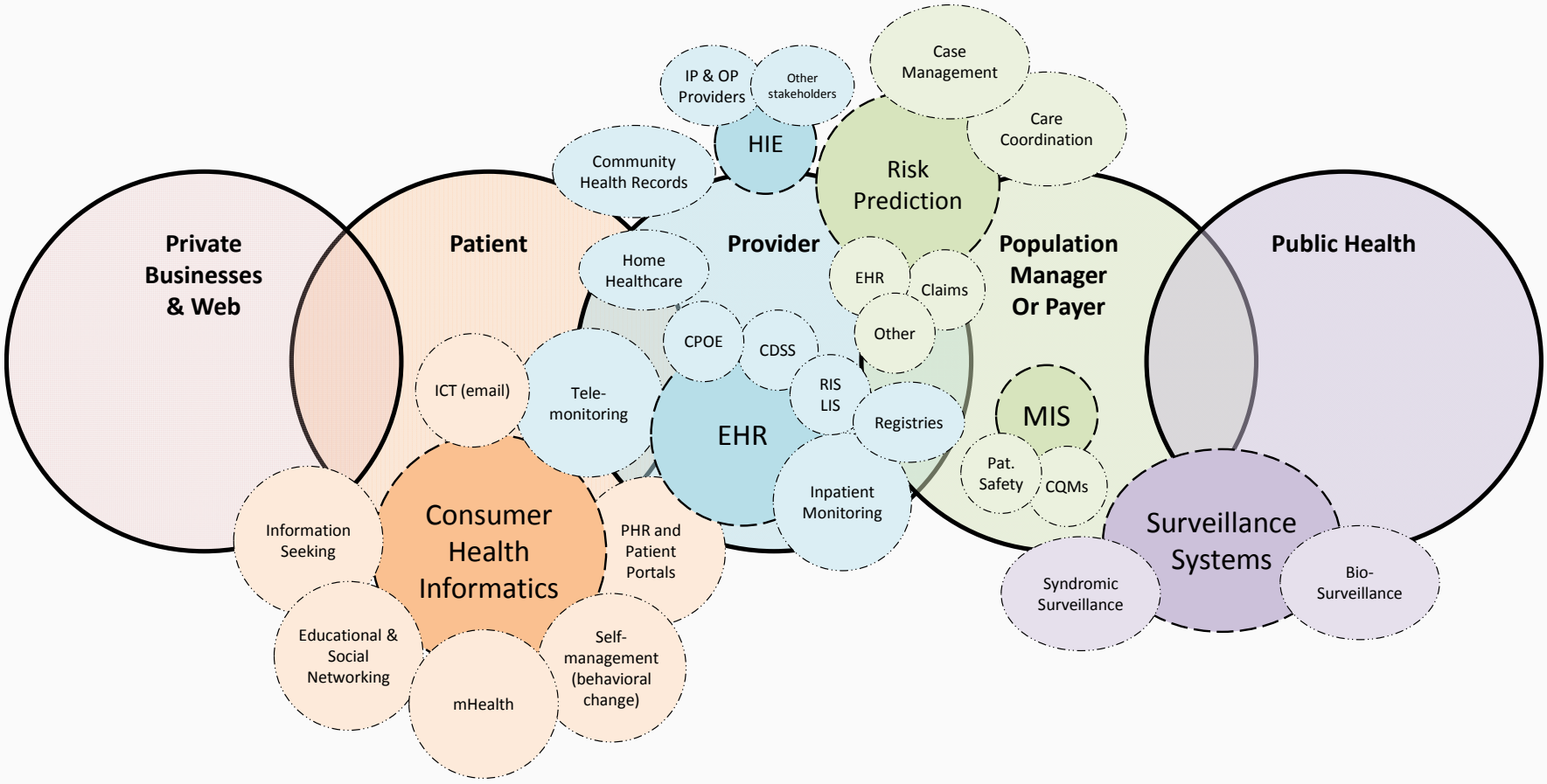
Weiner, 2012 http://www.ijhpr.org/content/1/1/33

Population Health Informatics → Data Types



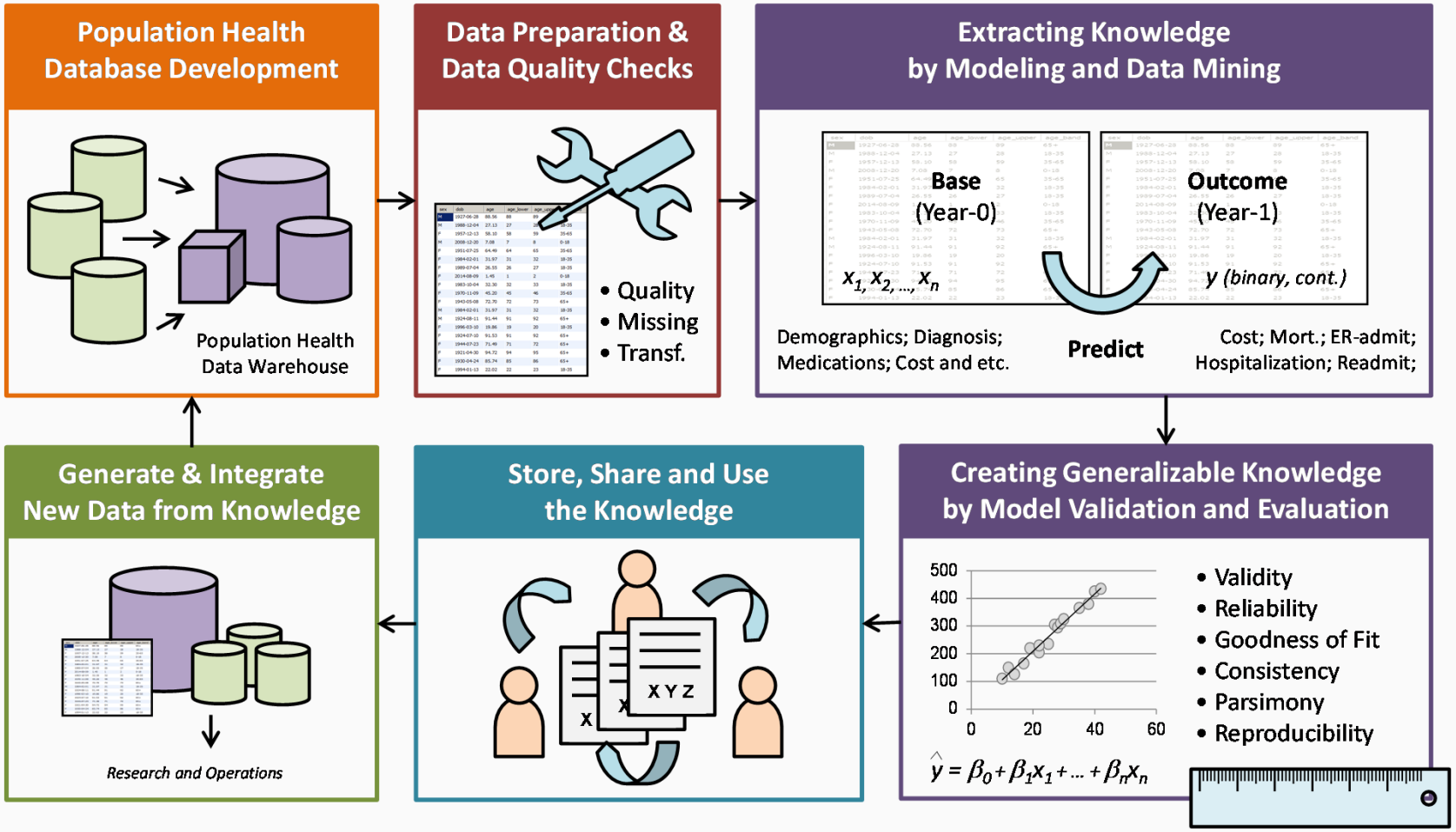
Health Determinants and Population Health Data Types

Population Health Informatics → Data Spectrum



Variety and Continuum of Information Systems in PopHI

Population Health Informatics → Data Analytic Cycle



Overall Population Health Knowledge Management Process

Population Health Informatics → Drivers

- **Incentives**:

  ARRA / HITECH →
  (1) EHR adoption [Meaningful Use (MU) measures];
  (2) state-wide HIEs;
  (3) Beacon Communities

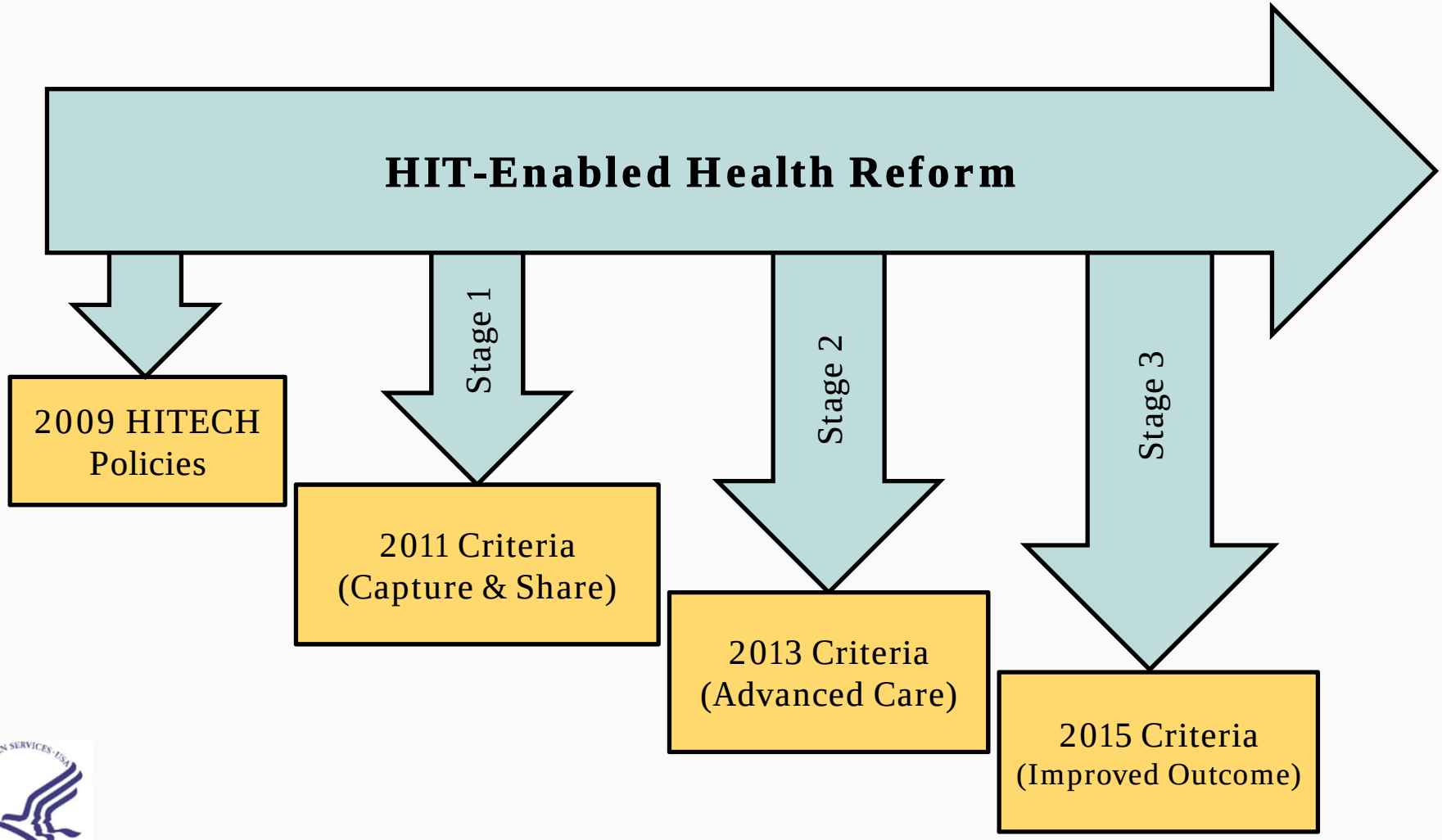
- **Mandates**:

  ACA → Payment Reforms → PCMH and ACO initiatives → Value-based → Capitated models → Population health + MACRA

- **Facilitation**:

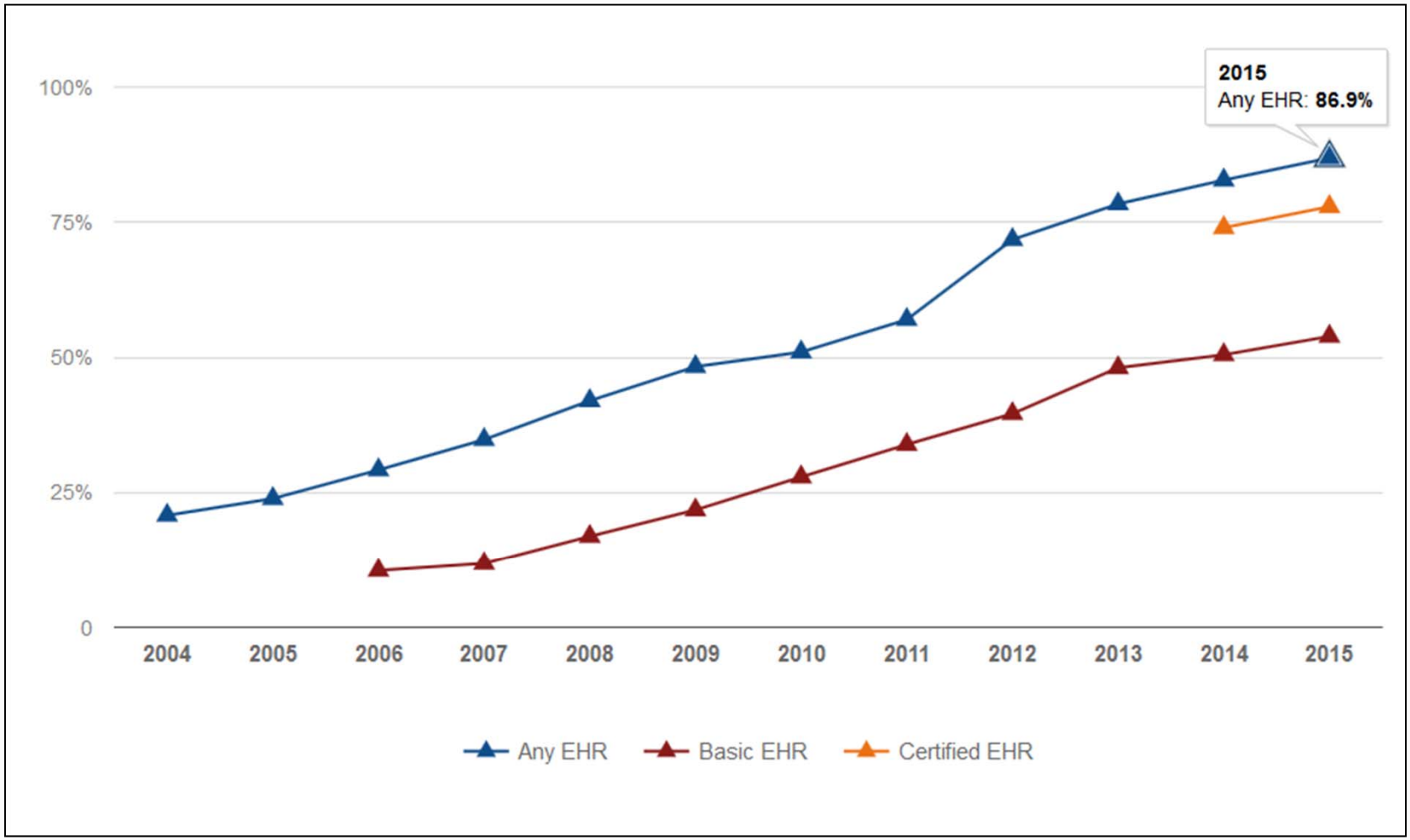
  ONC → Data Standards, integration, and sharing → Distributed models

Population Health Informatics → Drivers → EHR adoption



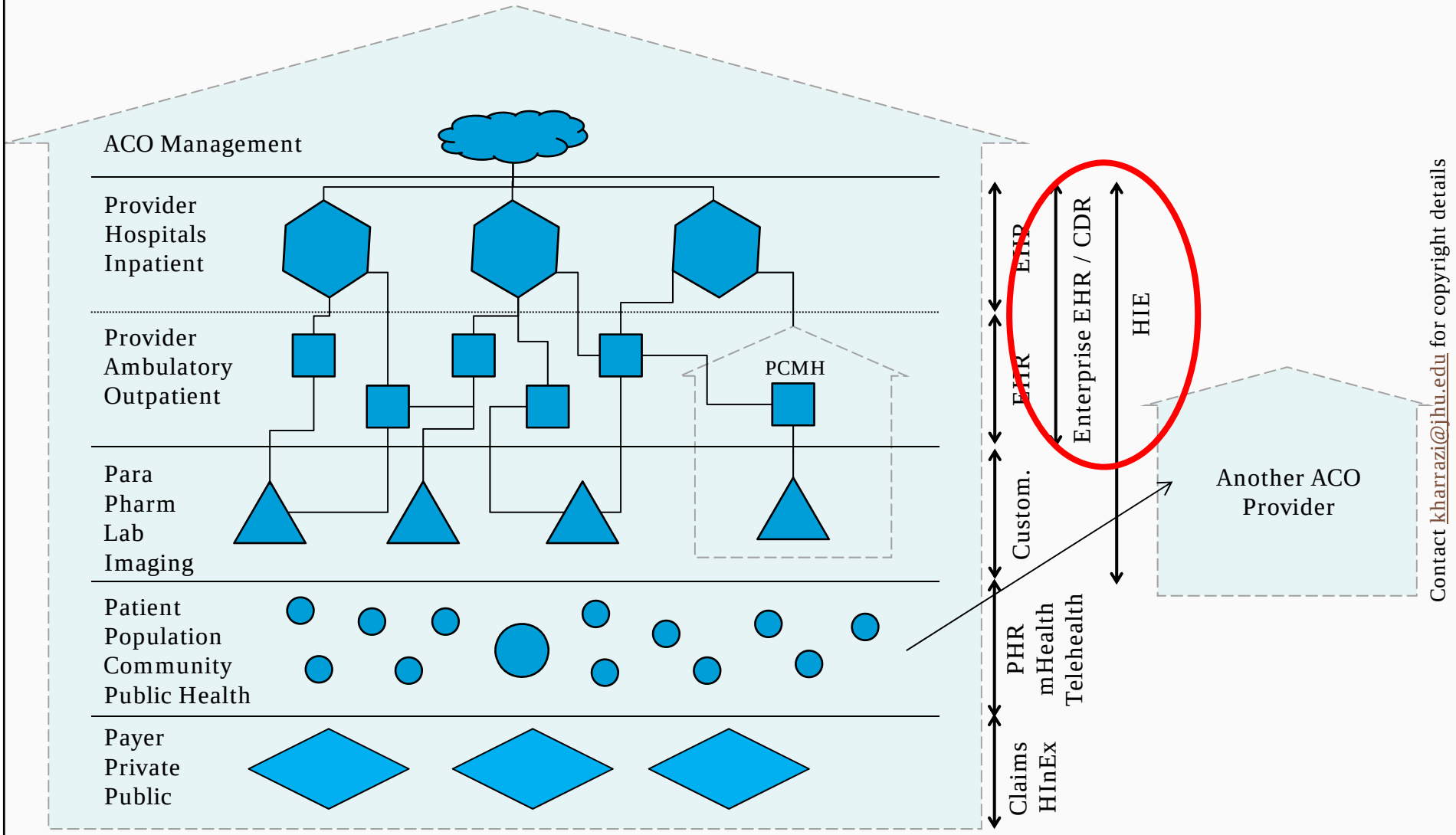
Drivers → ARRA → HITECH / MU of EHRs + HIEs

Population Health Informatics → Drivers → EHR adoption *(cont.)*

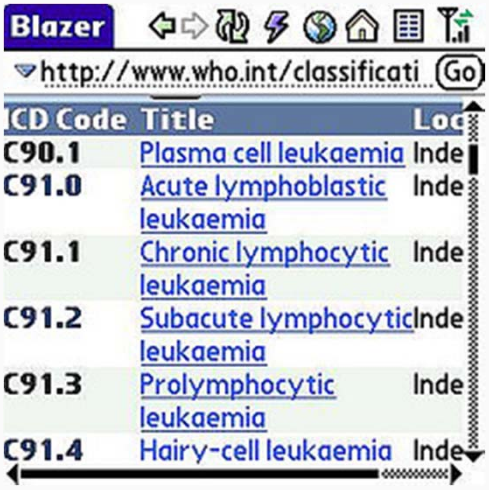


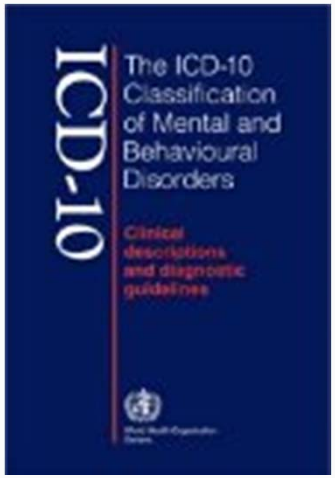
Percentage of EHR systems among office-based physicians
https://dashboard.healthit.gov/quickstats/pages/physician-ehr-adoption-trends.php

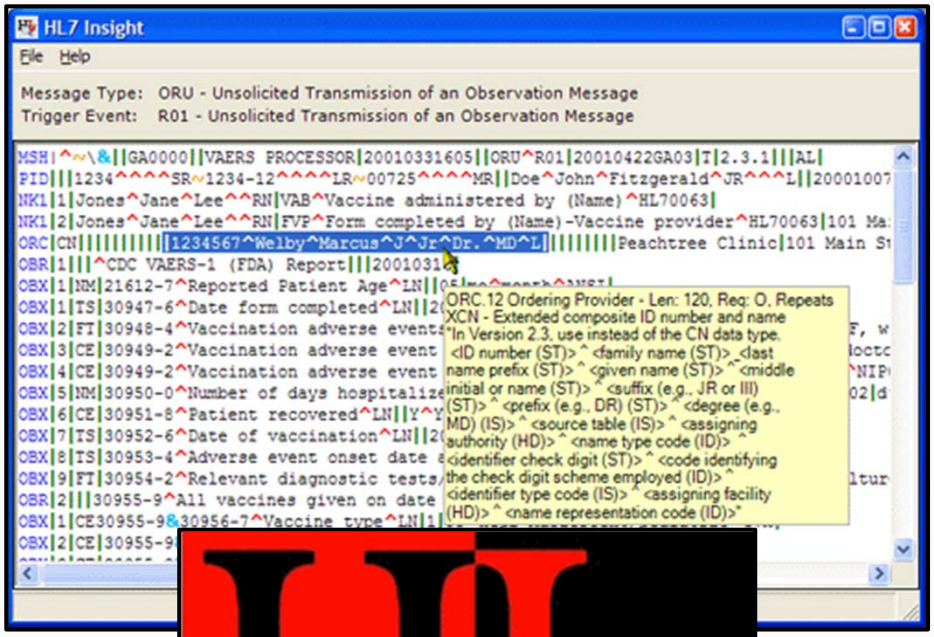
Big Data and Healthcare → Drivers → Value-base Care (ACOs) *(cont.)*

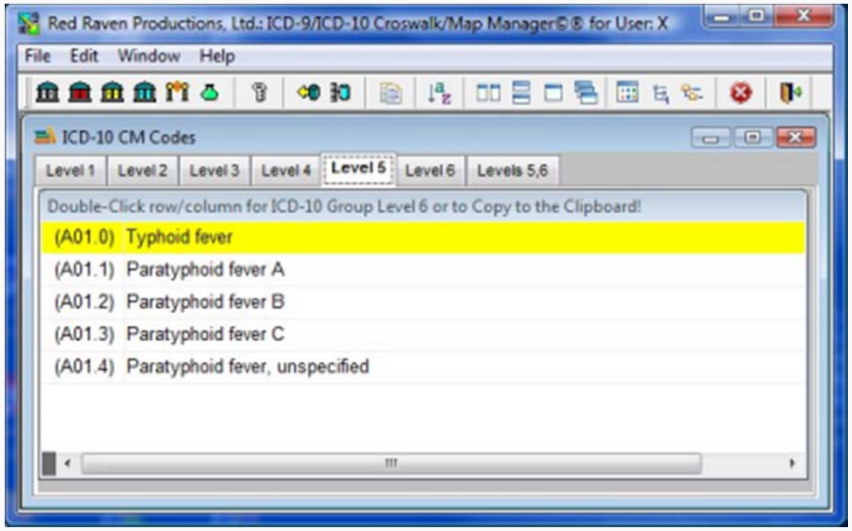


Population Health Informatics → Drivers → Data Standards









Population Health Informatics → National Workshop on Pop Health IT

A proposed national research and development agenda for population health informatics: summary recommendations from a National Expert Workshop

RECEIVED 6 August 2015
REVISED 17 December 2015
ACCEPTED 21 December 2015

Hadi Kharrazi[1,2], Elyse C Lasser[1], William A Yasnoff[2,3], John Loonsk[1], Aneel Advani[1], Harold P Lehmann[2], David C Chin[1], Jonathan P Weiner[1,2]

ABSTRACT

Objective The Johns Hopkins Center for Population Health IT hosted a 1-day symposium sponsored by the National Library of Medicine to help develop a national research and development (R&D) agenda for the emerging field of population health informatics (PopHI).

Material and Methods The symposium provided a venue for national experts to brainstorm, identify, discuss, and prioritize the top challenges and opportunities in the PopHI field, as well as R&D areas to address these.

Results This manuscript summarizes the findings of the PopHI symposium. The symposium participants' recommendations have been categorized into 13 overarching themes, including policy alignment, data governance, sustainability and incentives, and standards/interoperability.

Discussion The proposed consensus-based national agenda for PopHI consisted of 18 priority recommendations grouped into 4 broad goals: (1) Developing a standardized collaborative framework and infrastructure, (2) Advancing technical tools and methods, (3) Developing a scientific evidence and knowledge base, and (4) Developing an appropriate framework for policy, privacy, and sustainability. There was a substantial amount of agreement between all the participants on the challenges and opportunities for PopHI as well as on the actions that needed to be taken to address these.

Conclusion PopHI is a rapidly growing field that has emerged to address the population dimension of the Triple Aim. The proposed PopHI R&D agenda is comprehensive and timely, but should be considered only a starting-point, given that ongoing developments in health policy, population health management, and informatics are very dynamic, suggesting that the agenda will require constant monitoring and updating.

JAMIA Paper Reporting on the Findings from the National Workshop

Population Health Informatics → National Workshop on Pop Health IT *(cont.)*

❖ Population Health Definition

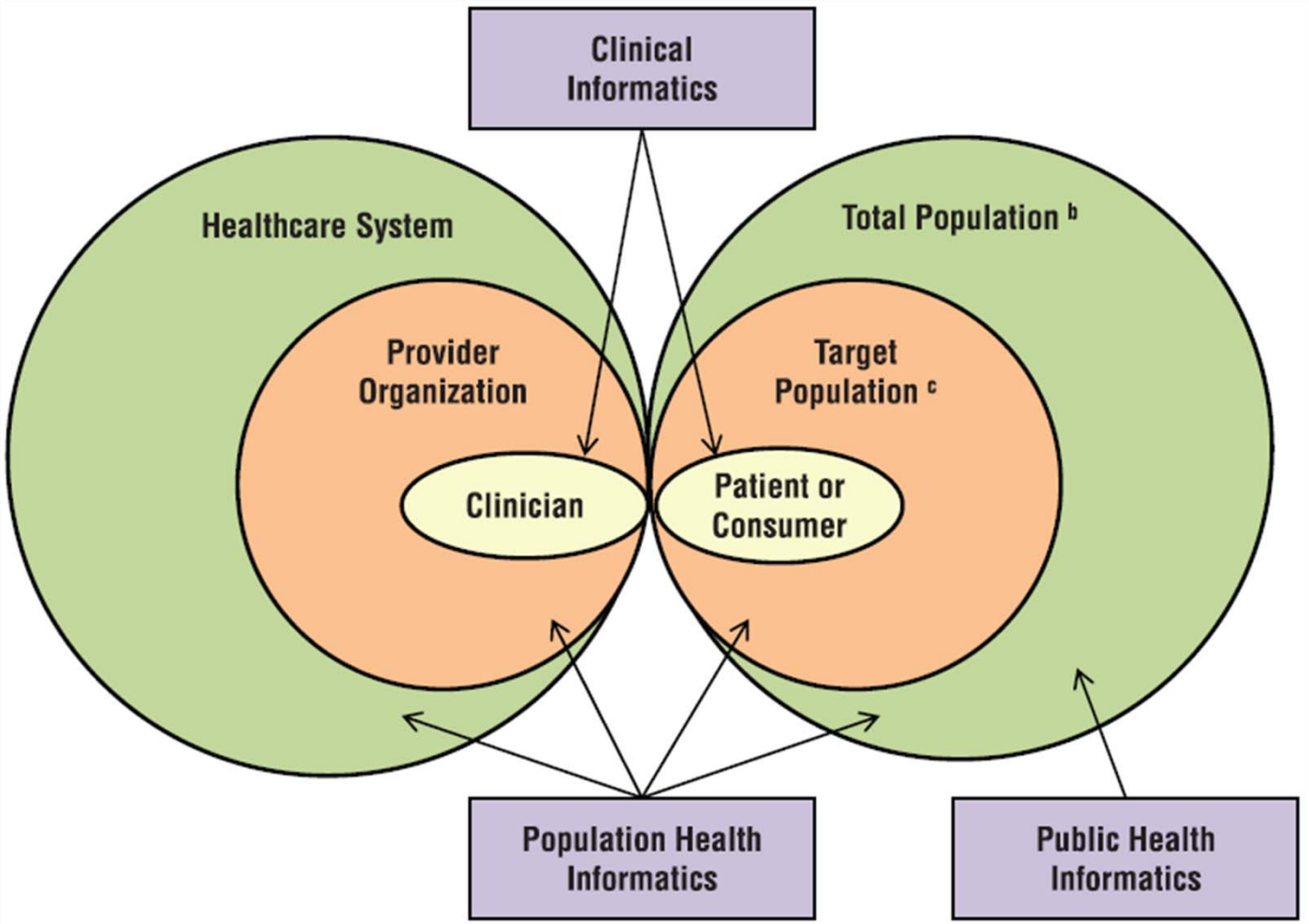
Population health comprises organized activities for assessing and improving the health and well-being of a defined population. Population health is practiced by both private and public organizations. The target "population" can be a specific geographic community or region, or it may represent some other "denominator," such as enrollees of a health plan, persons residing in a provider's catchment area, or an aggregation of individuals with special needs. The difference between population health and public health is subtle, and there is not always a full consensus on these definitions. That said, public health services are typically provided by government agencies and include the "core" public health functions of health assessment, assurance, and policy-setting. In the United States, most actions of public health agencies represent population health, but a considerable proportion, if not the majority, of population health services are provided by private organizations.

Population Health Informatics → National Workshop on Pop Health IT *(cont.)*

Domain / Context	Population Health Informatics	Public Health Informatics	Clinical Informatics
Common Intervention Targets [a]	• Total population • Target populations • Provider organization • Healthcare systems	• Total population	• Clinician • Patient or consumer • Provider organization • Target population
Main Operational Goal	• Outreach and prevention • Care integration • Disease management	• Assessment • Prevention	• Treatment • Rehabilitation
Action Arm	• Population health organization • Care management organizations	• Public health agencies • Non-for-profit and non-governmental organizations	• Clinical organizations
Key Stakeholders	• Provider and payer systems • Government and community	• Federal, state, and local governments	• Providers • Consumers
Key Information Challenges	• Capturing non-medical info • Information system inter-operability across sectors	• Expanding public health IT systems • Medical and public health interoperability	• Decision support • EHR interoperability

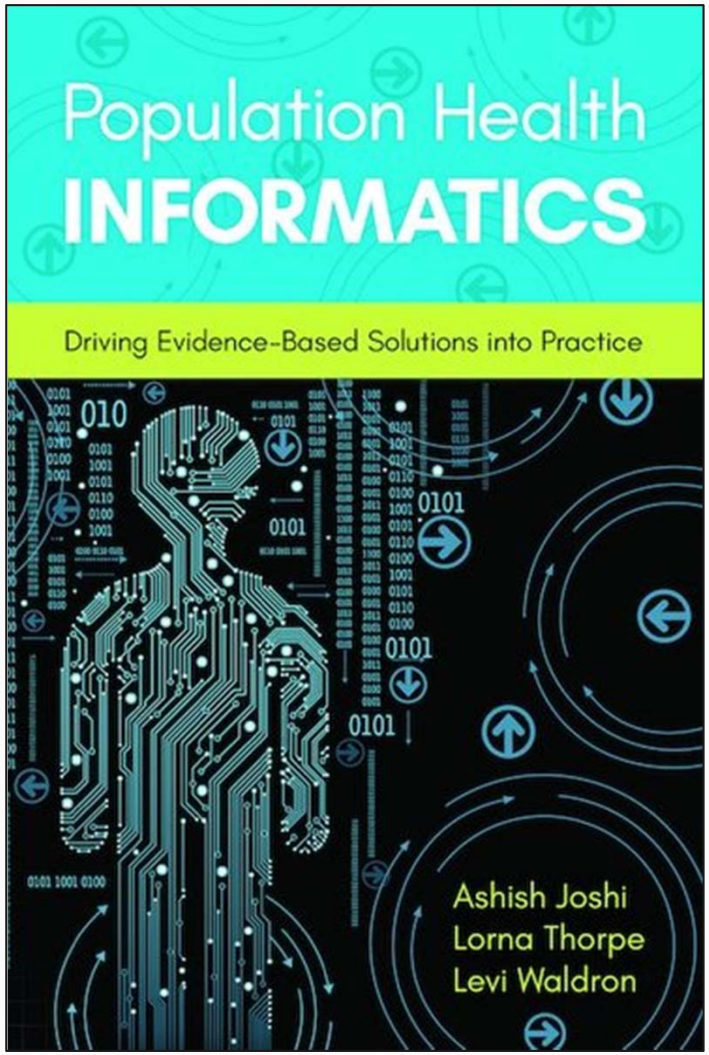
Population Health Informatics vs Public Health Informatics vs Clinical Informatics
(table of contexts and differences)

Population Health Informatics → National Workshop on Pop Health IT *(cont.)*



Population Health Informatics vs Public Health Informatics vs Clinical Informatics
(Venn diagram of population denominator)

Population Health Informatics → Growing domain…



CPHIT Research Portfolio

CPHIT Portfolio *(cont.)*

- **Research Portfolio** *(selected list)*
- **EHR-based Utilization Prediction (eACG)**: Developing a wide range of EHR-based population focused predictive modeling tools
  - EHR ICD and e-prescribing Rx input (coordinated with claims)
  - Non-claims data: Lab Data; BMI and Vitals; and Social History
- **Geriatric Frailty**: Developing a new geriatric/frailty "e-risk" score utilizing structured and unstructured EHR data and Claims
- **Predicting Elderly Falls**: Developing analytics with Baltimore City Heath Department for regional collaboration to identify and predict elder's fall injuries in the community using social, medical, and public health data

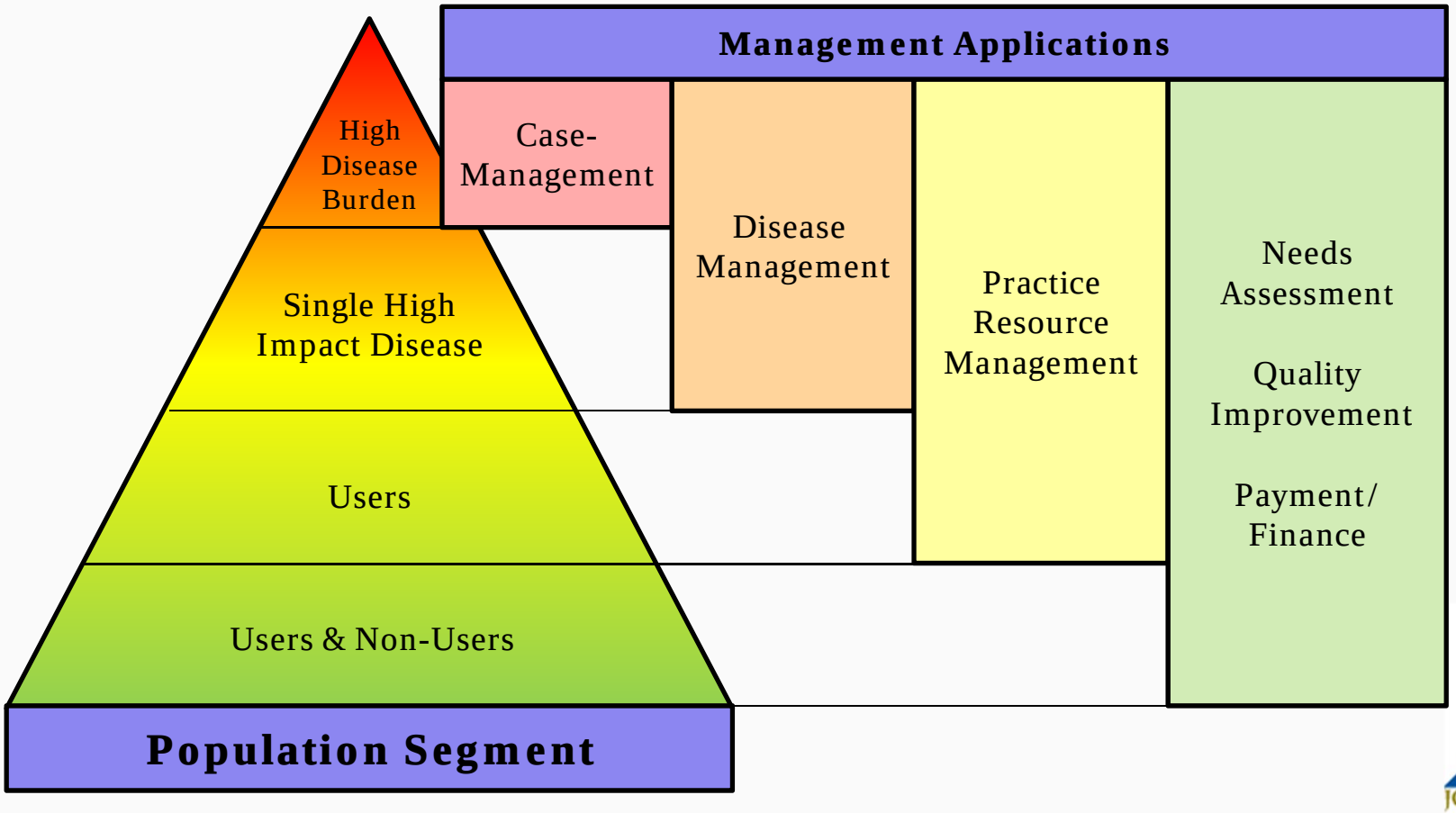
CPHIT Portfolio *(cont.)*

- **VHA Obesity**: Collaborating with the Veteran Health Administration
  - Developed population health analytic framework
  - Linking geo and social data for obesity trend analysis
  - Adding social/geo framework to PCMH (PACT) case finding program
- **Pop e-Measures**: Collaborating with Maryland State Health Department, HIE and Hospital Commission
  - Assist in the statewide pop health digital measurement infrastructure
  - Develop population health focused measures
- **Opioid**: Using multiple novel sources of data to develop predictive models to identify persons at risk for opioid overdose
- **Consumer Data**: Linking consumer/marketing data with medical data to identify health outcomes (e.g., consumer reports)

CPHIT Portfolio → EHR-based Utilization Prediction (eACG)

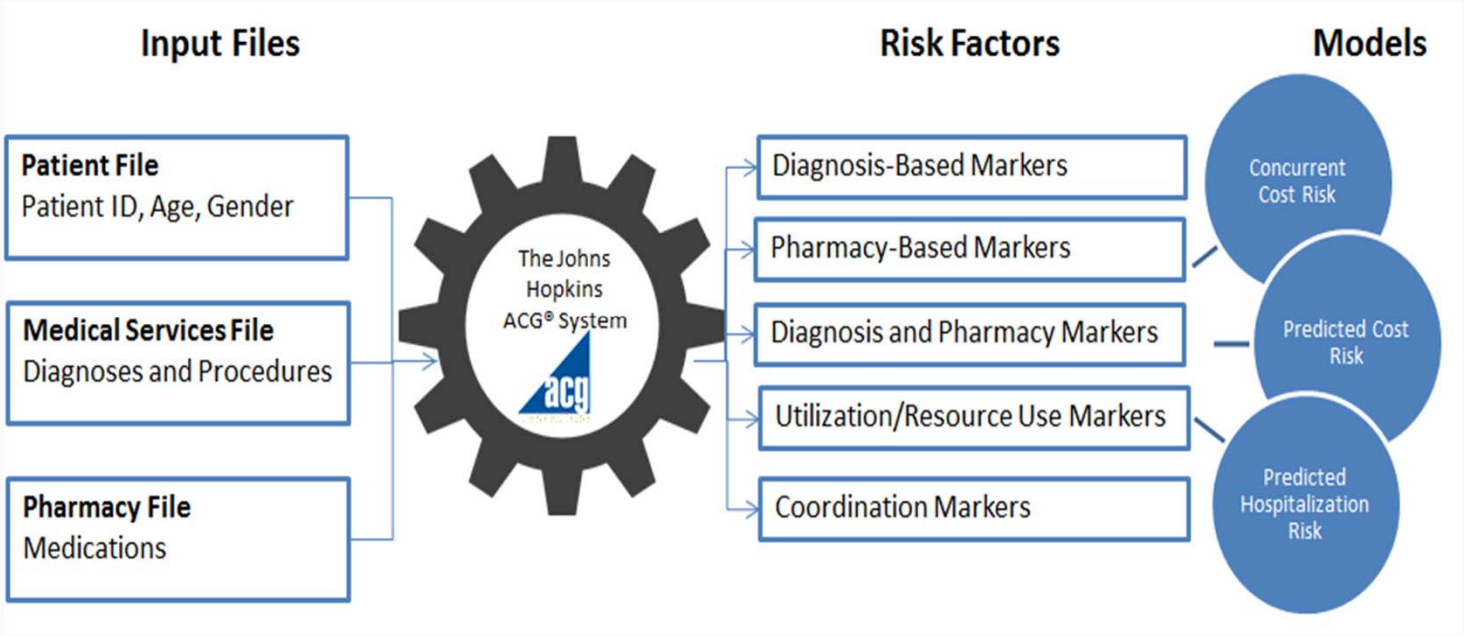
acg.jhsph.edu

- ACG system offers a unique approach to measuring morbidity that improves accuracy and fairness in evaluating provider performance, identifying patients at high risk, forecasting healthcare utilization and setting equitable payment rates.



CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*

- Billions of dollars per year are now routinely exchanged using ACGs in almost every US State and in 16 + nations. Healthcare of as many as 100+ million patients is actively managed and monitored using ACGs on all continents. Over 700+ peer reviewed articles have been published that apply and evaluate ACGs.
- **Other EHR-only data**: Lab results; Vital Signs; Medical and Family history...
- **New applications**: Real time population health / community surveillance; Real time clinical action for individual consumer; Functional Status / Frailty



CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*

	Data Source[a]	
Characteristic	**Claims**	**EHR**[b]
Purpose	Reimbursement	Clinical care
Scope	All providers, including out of network providers, for a given patient	Network providers of a patient
Data consistency	High consistency across sources	Lower consistency across sources
Data structure	Most of data is structured	Considerable unstructured data
Coding standard	Strict adherence to coding systems	Variable adherence to coding systems
Provider coverage	All providers accepting the insurance	Limited to providers using same EHR
Coding limit	Limited to encoded data	Provides ability to enter free text
Member limitation	Limited to insured patients	Insured and uninsured patients
Coverage limitation	Non-covered items are missing	Includes data on non-covered items
Data type	Limited (mainly enrollment, Dx, Rx)	Additional data types (see below)
Data Availability	**Claims**	**EHR**[b]
Demographics[a]	Yes	Yes
Race/ethnicity	Limited	Limited
Diagnosis[a]	Yes	Yes
Procedures	Yes	Yes
Eligibility	Yes	Limited
Medications[a]	Pharmacy data (drugs dispensed)	Prescriptions ordered & MedRec data
Socioeconomic data	Zip-code derived	Coded and zip-code derived
Family history	Not available	Yes
Problem list	Not available	Yes
Procedure results	Not available	Yes
Laboratory results	Not available	Yes
Vital signs	Not available	Yes
Behavioral risk factors	Not available	Limited
Standardized surveys	Limited	Limited

Claims vs EHR

CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*

ORIGINAL ARTICLE

Comparing Population-based Risk-stratification Model Performance Using Demographic, Diagnosis and Medication Data Extracted From Outpatient Electronic Health Records Versus Administrative Claims

Hadi Kharrazi, MD, PhD, Winnie Chi, PhD,* Hsien-Yen Chang, PhD,* Thomas M. Richards, MSc,* Jason M. Gallagher, MBA,† Susan M. Knudson, MA,† and Jonathan P. Weiner, DrPH**

Background: There is an increasing demand for electronic health record (EHR)–based risk stratification and predictive modeling tools at the population level. This trend is partly due to increased value-based payment policies and the increasing availability of EHRs at the provider level. Risk stratification models, however, have been traditionally derived from claims or encounter systems. This study evaluates the challenges and opportunities of using EHR data instead of or in addition to administrative claims for risk stratification.

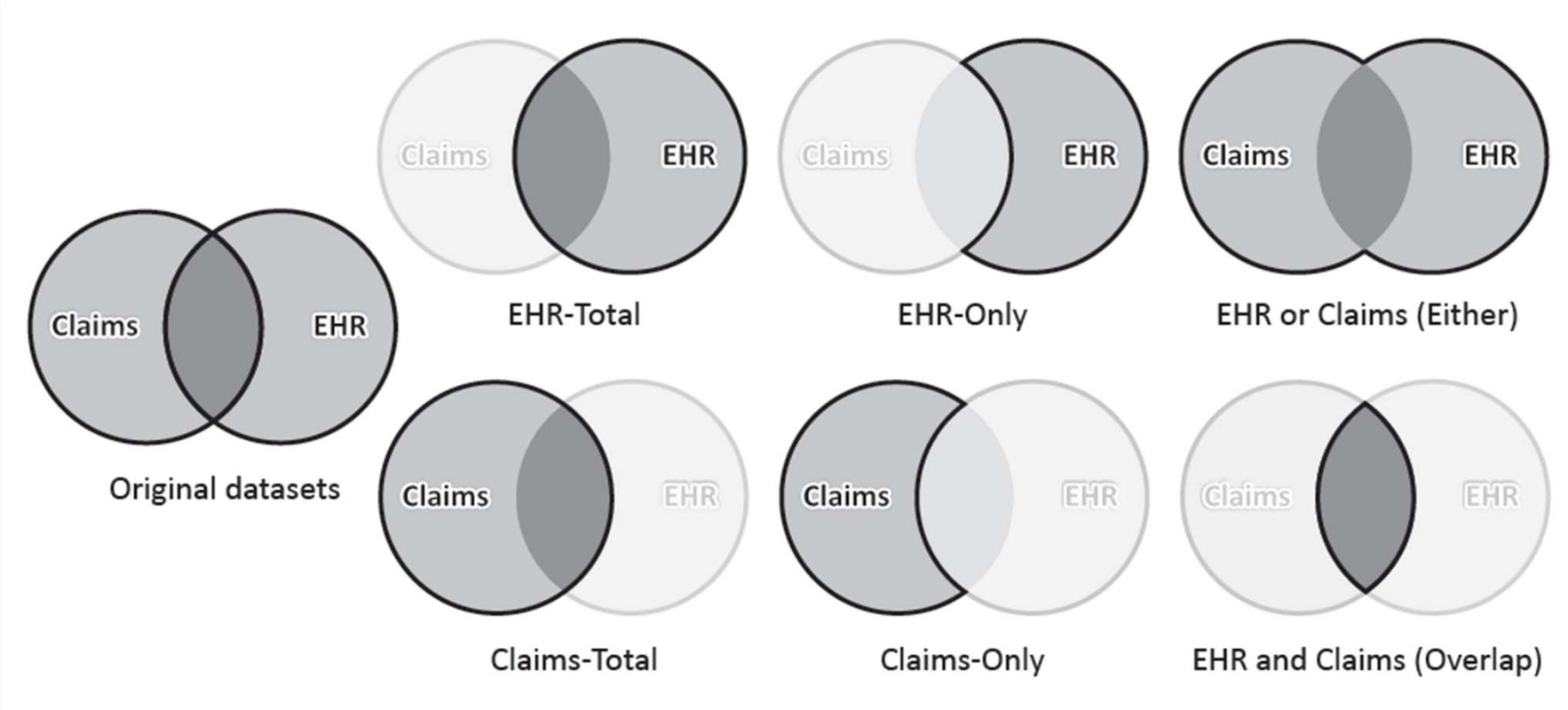
Discussion: The results show a promising performance of models predicting cost and hospitalization using outpatient EHR's diagnosis and medication data. More research is needed to evaluate the benefits of other EHR data types (eg, lab values and vital signs) for risk stratification.

Key Words: risk stratification, predictive modeling, electronic health records, administrative claims, The Johns Hopkins ACG System

(*Med Care* 2017;55: 789–796)

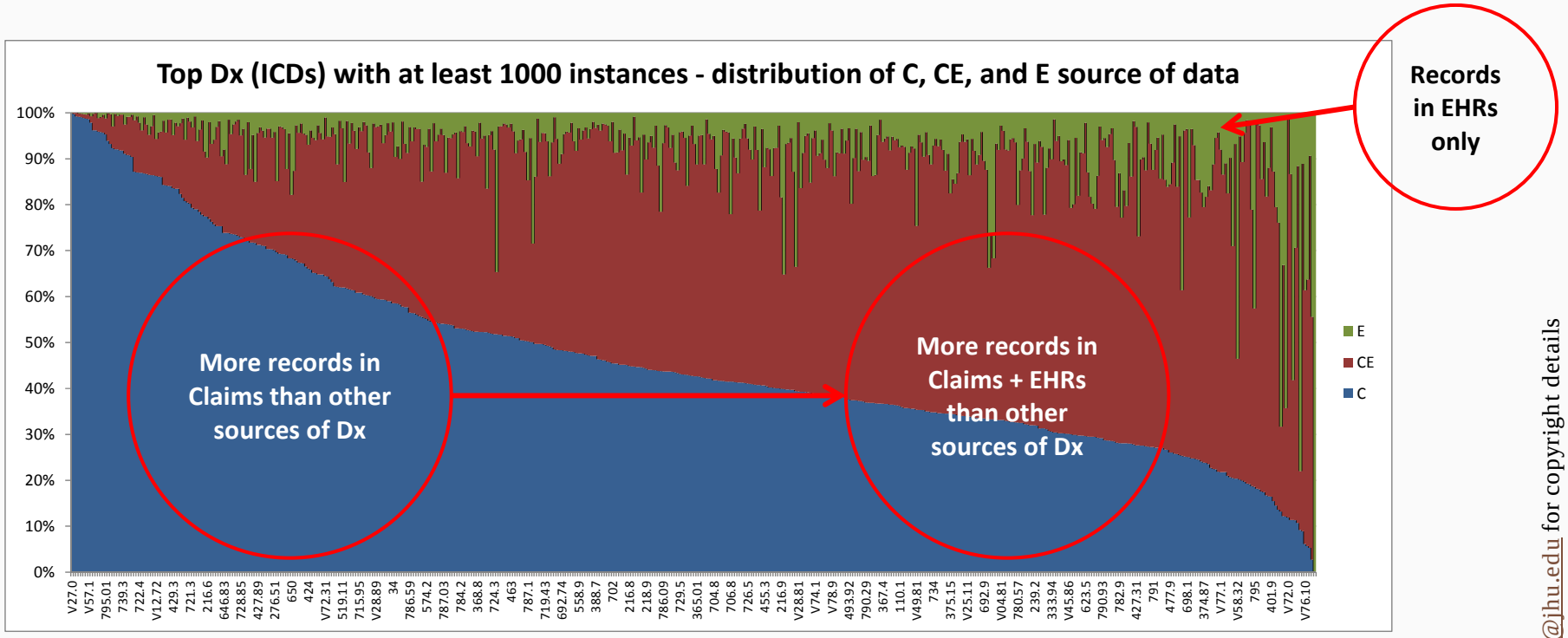
Comparing Claims and EHR for Risk Stratification

CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*



Comparing Various Overlaps of Claims and EHR for Risk Stratification

CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*



Comparing Diagnostic Data Found in Claims (C) vs EHRs (E)

CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*

TABLE 2. Comparison of Case Findings of Selected Conditions Based on Diagnostic and Medication Information Extracted From Administrative Claims, EHR Data, and a Combination of Both

Measure		N (% Population)			Overlaps (%)*				
Condition†	Dx and Rx	Claims	EHR	Claims or EHR	CLM-T	EHR-T	OVER	CLM-O	EHR-O
Diabetes	Dx (EDC)	4047 (4.73)	3872 (4.52)	4209 (4.92)	96.15	91.99	88.14	8.01	3.85
	Rx (RxMG)	3878 (4.53)	3632 (4.24)	4389 (5.13)	88.36	82.75	71.11	17.25	11.64
	Both (and)	3364 (3.93)	2912 (3.40)	3522 (4.12)	95.51	82.68	78.19	17.32	4.49
	Either (or)	4561 (5.33)	4592 (5.37)	5076 (5.93)	89.85	90.46	80.32	9.54	10.15
Hypertension	Dx (EDC)	10,152 (11.86)	10,225 (11.95)	11,656 (13.62)	87.10	87.72	74.82	12.28	12.90
	Rx (RxMG)	13,824 (16.15)	10,082 (11.78)	14,090 (16.46)	98.11	71.55	69.67	28.45	1.89
	Both (and)	9232 (10.79)	8132 (9.50)	10,600 (12.39)	87.09	76.72	63.81	23.28	12.91
	Either (or)	14,744 (17.23)	12,175 (14.23)	15,146 (17.70)	97.35	80.38	77.73	19.62	2.65
Depression	Dx (EDC)	5451 (6.37)	4853 (5.67)	6683 (7.81)	81.57	72.62	54.18	27.38	18.43
	Rx (RxMG)	15,814 (18.48)	9118 (10.65)	16,053 (18.76)	98.51	56.80	55.31	43.20	1.49
	Both (and)	4425 (5.17)	3126 (3.65)	5396 (6.31)	82.01	57.93	39.94	42.07	17.99
	Either (or)	16,840 (19.68)	10,845 (12.67)	17,340 (20.26)	97.12	62.54	59.66	37.46	2.88
Cancer	Dx (EDC)	2458 (2.87)	2210 (2.58)	2836 (3.31)	86.67	77.93	64.60	22.07	13.33
	Rx (RxMG)	806 (0.94)	356 (0.42)	827 (0.97)	97.46	43.05	40.51	56.95	2.54
	Both (and)	478 (0.56)	236 (0.28)	497 (0.58)	96.18	47.48	43.66	52.52	3.82
	Either (or)	2786 (3.26)	2330 (2.72)	3166 (3.70)	88.00	73.59	61.59	26.41	12.00

Cases found in EHR versus Claims:
Diabetes, Hypertension, Depression, Cancer

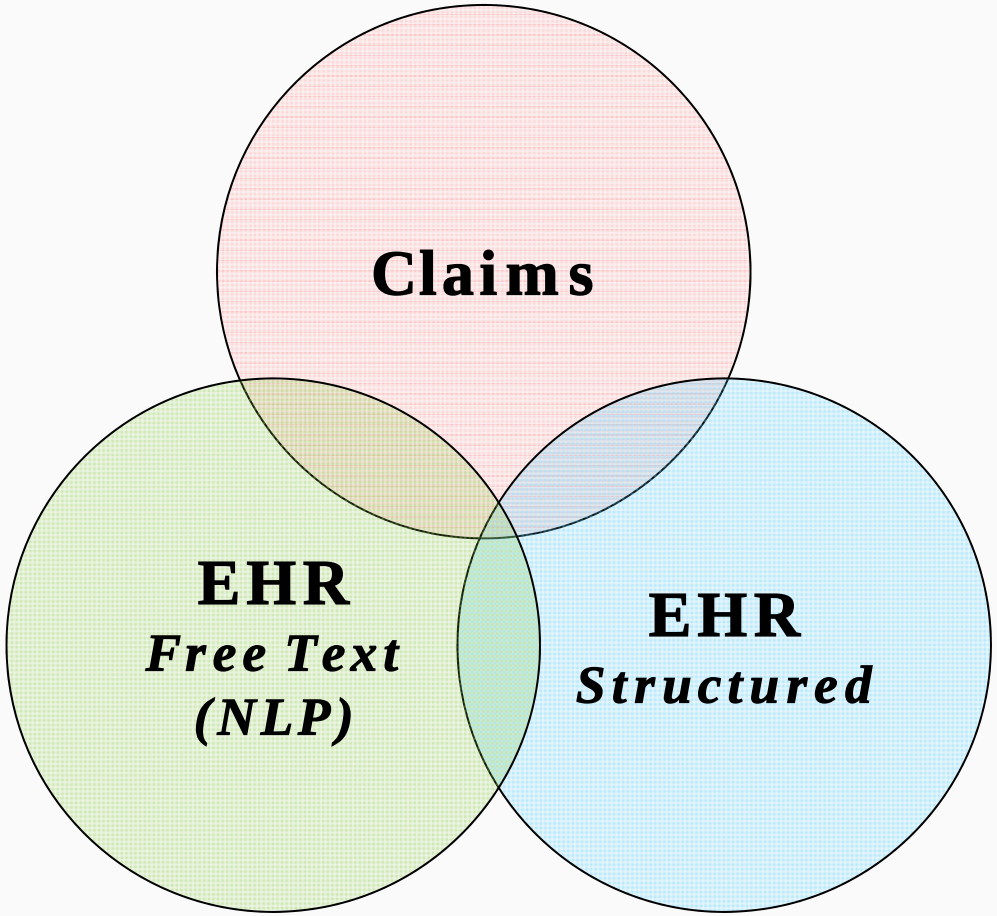
CPHIT Portfolio → EHR-based Utilization Prediction (eACG) *(cont.)*

TABLE 4. Comparison of Concurrent and Prospective Performance (Area Under the Curve) for Hospitalization and Top 1% Cost Using Claims, EHRs Data, or Both

		Data Source and Model Variables*									
		Dem.	Claims			EHR			Claims or EHR		
Outcome†	Year	Sex + Age	Dx	Rx	DxRx	Dx	Rx	DxRx	Dx	Rx	DxRx
Hosp.‡	Concurrent	0.640	0.864	0.804	0.875	0.775	0.765	0.809	0.861	0.839	0.897
	Prospective	0.656	0.779	0.736	0.783	0.732	0.686	0.736	0.783	0.730	0.783
Top 1% cost§	Concurrent	0.649	0.925	0.864	0.938	0.842	0.774	0.845	0.926	0.888	0.941
	Prospective	0.666	0.832	0.800	0.835	0.782	0.663	0.752	0.823	0.776	0.818

Model performance using EHR versus Claims

CPHIT Portfolio → Geriatric Frailty



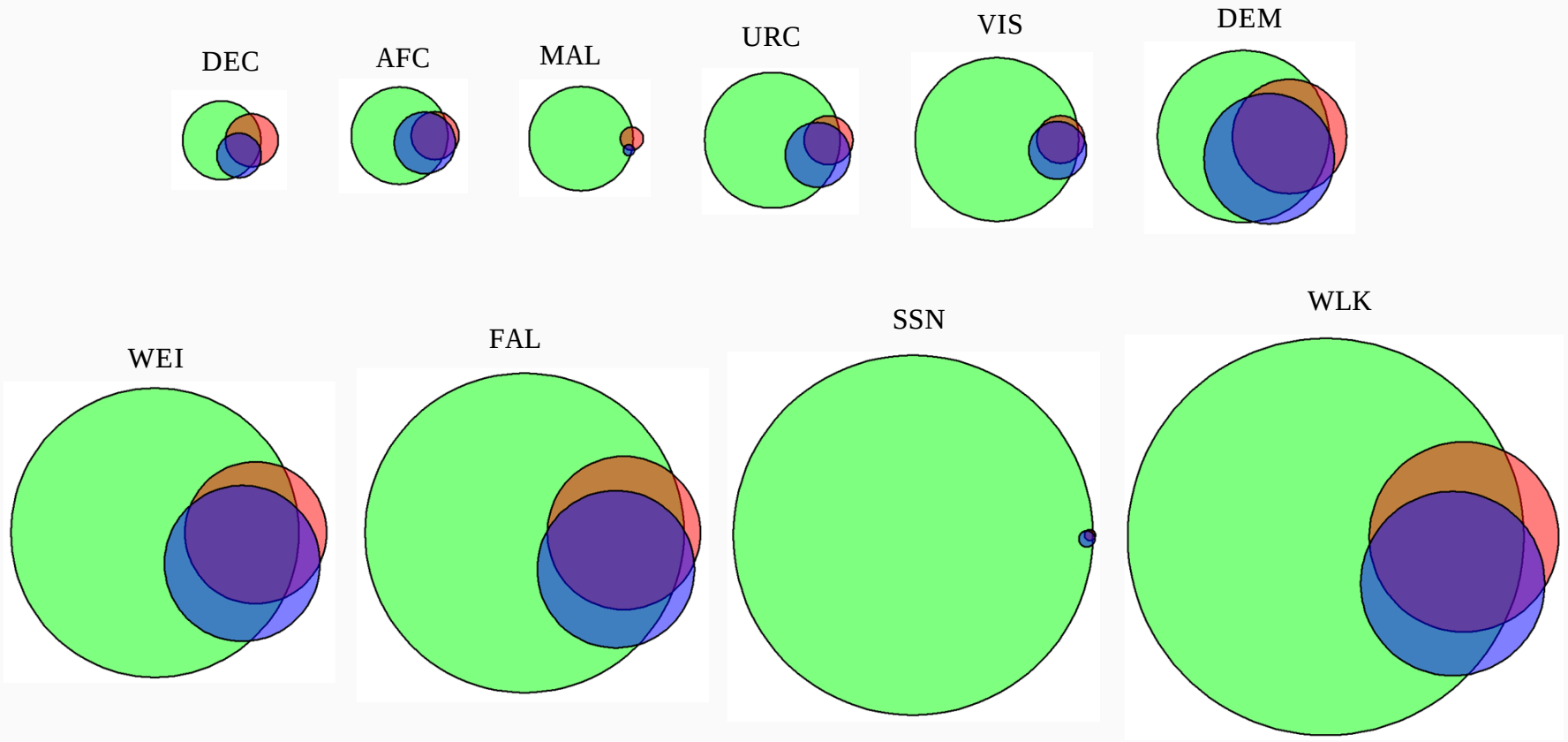
CPHIT Portfolio → Geriatric Frailty *(cont.)*

- This study evaluates the added-value of EHR's unstructured text in identifying geriatric risk factors. Claims and structured EHR data give an incomplete picture about several constructs that influence geriatric risk. There is a high possibility to miss patients with geriatric risk markers when excluding free-text from an analysis.

Construct	C%	E%	NLP%	C+E%	C+E+NLP%	E+NLP%
AFC	0.52	0.87	2.04	0.93	2.22	2.18
DEC	0.65	0.45	1.11	0.84	1.47	1.19
DEM	3.10	3.94	6.36	4.67	7.21	6.82
FAL	5.48	5.73	**22.10**	7.32	23.16	22.77
MAL	0.12	0.03	2.37	0.14	2.44	2.37
SSN	0.03	0.06	**28.02**	0.07	28.03	28.03
URC	0.55	0.97	3.10	1.09	3.38	3.30
VIS	0.53	0.76	4.81	0.85	4.94	4.92
WEI	4.68	5.60	9.11	6.65	11.96	11.18
WLK	8.39	7.88	**32.45**	10.80	34.93	34.02

Comparing rates of geriatric risk factors using claims and EHR data (structured and free text) within the larger population reference

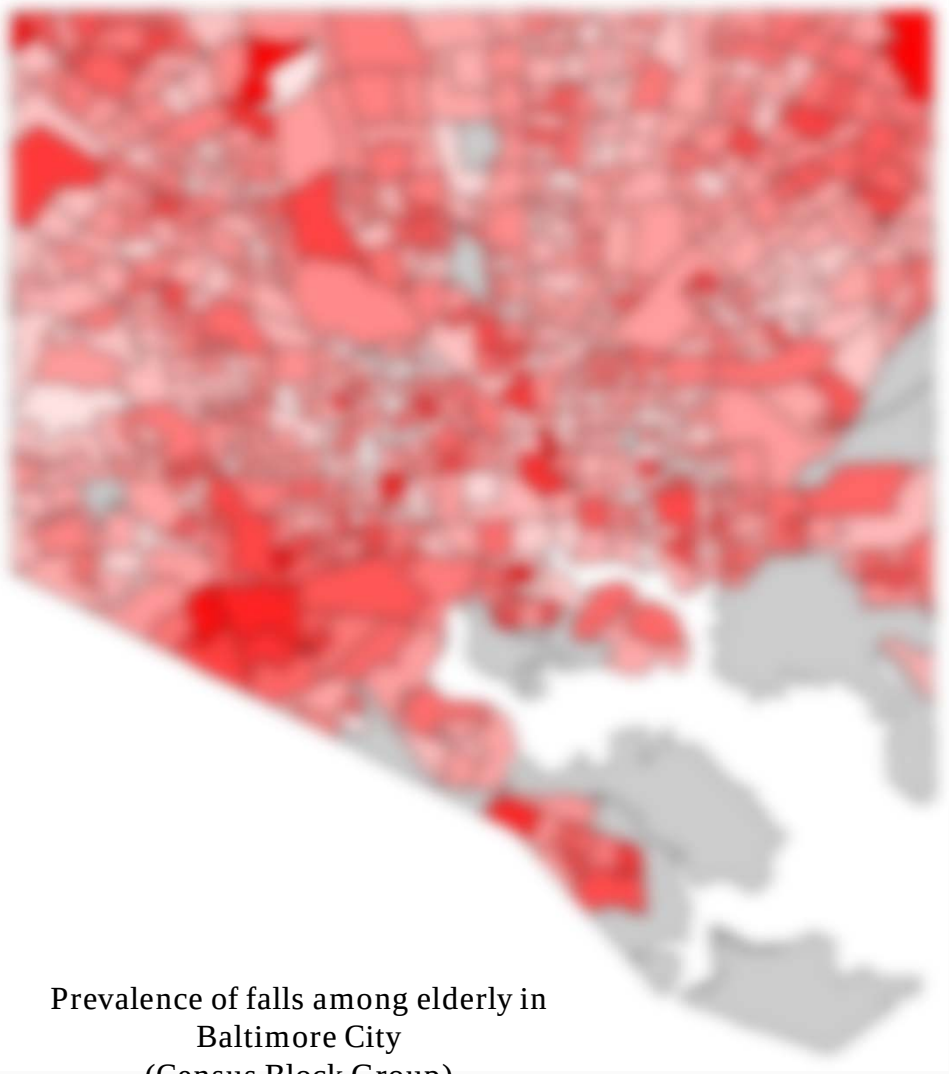
CPHIT Portfolio → Geriatric Frailty *(cont.)*



Added value of free text represented by the Venn diagram
Circle sizes represent the number of patients identified by each methodology/data-source
***Green**: EHR Free Text; **Blue**: EHR Structured; **Red**: Insurance Claims*

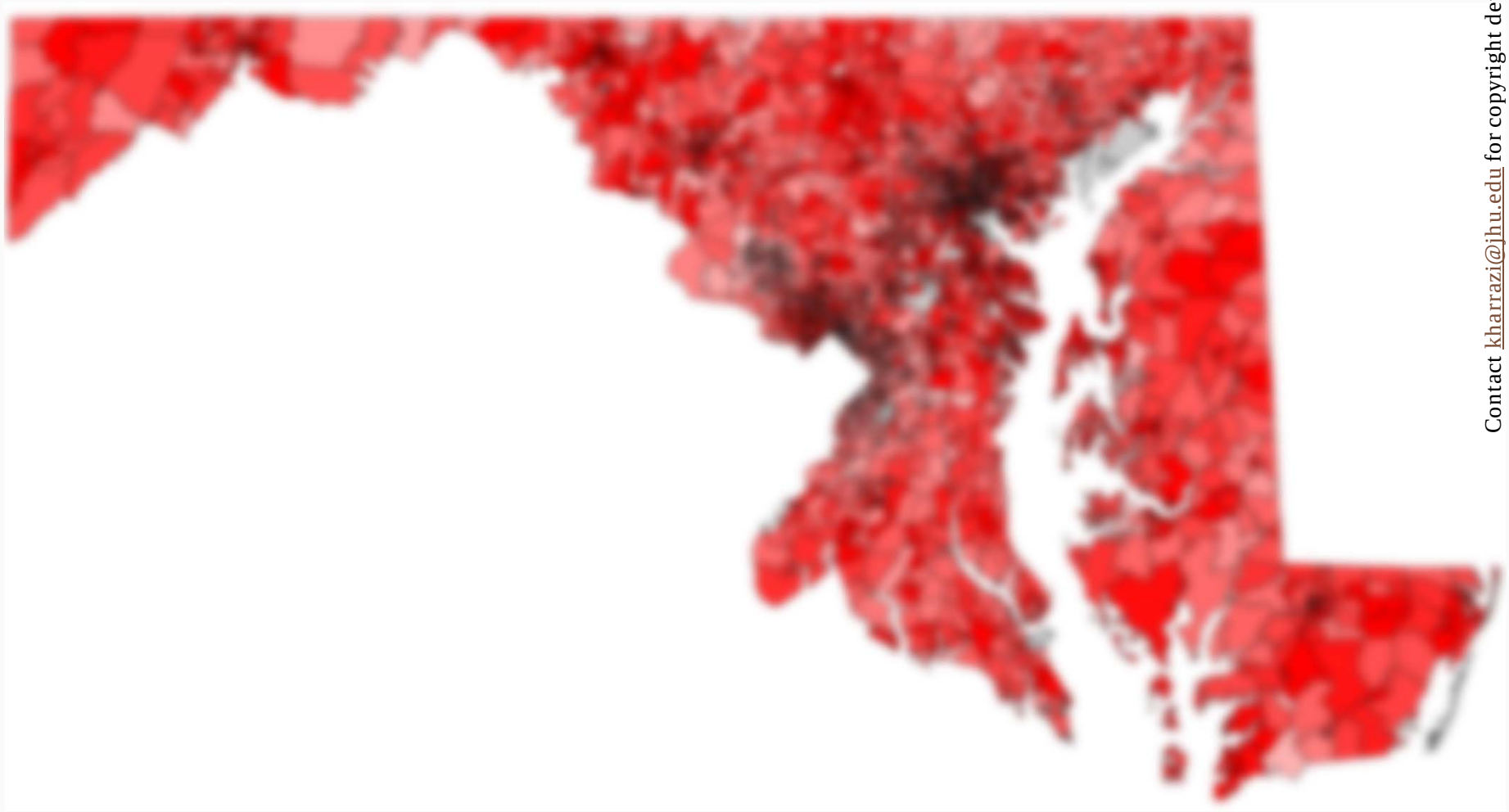
CPHIT Portfolio → Predicting Elderly Falls

- Baltimore Falls Reduction Initiative Engaging Neighborhoods & Data (B'FRIEND)
- B'FRIEND is a public-private partnership in Baltimore City based on innovative use of health data to decrease the rate of falls leading to an emergency room (ER) or hospital admission among elderly.
- Aim (1): Develop and validate a case identification methodology;
- Aim (2): Develop and validate a fall's risk prediction model
- Aim (3): Evaluate the fall risk score and disseminate results



Prevalence of falls among elderly in Baltimore City
(Census Block Group)

CPHIT Portfolio → Predicting Elderly Falls *(cont.)*



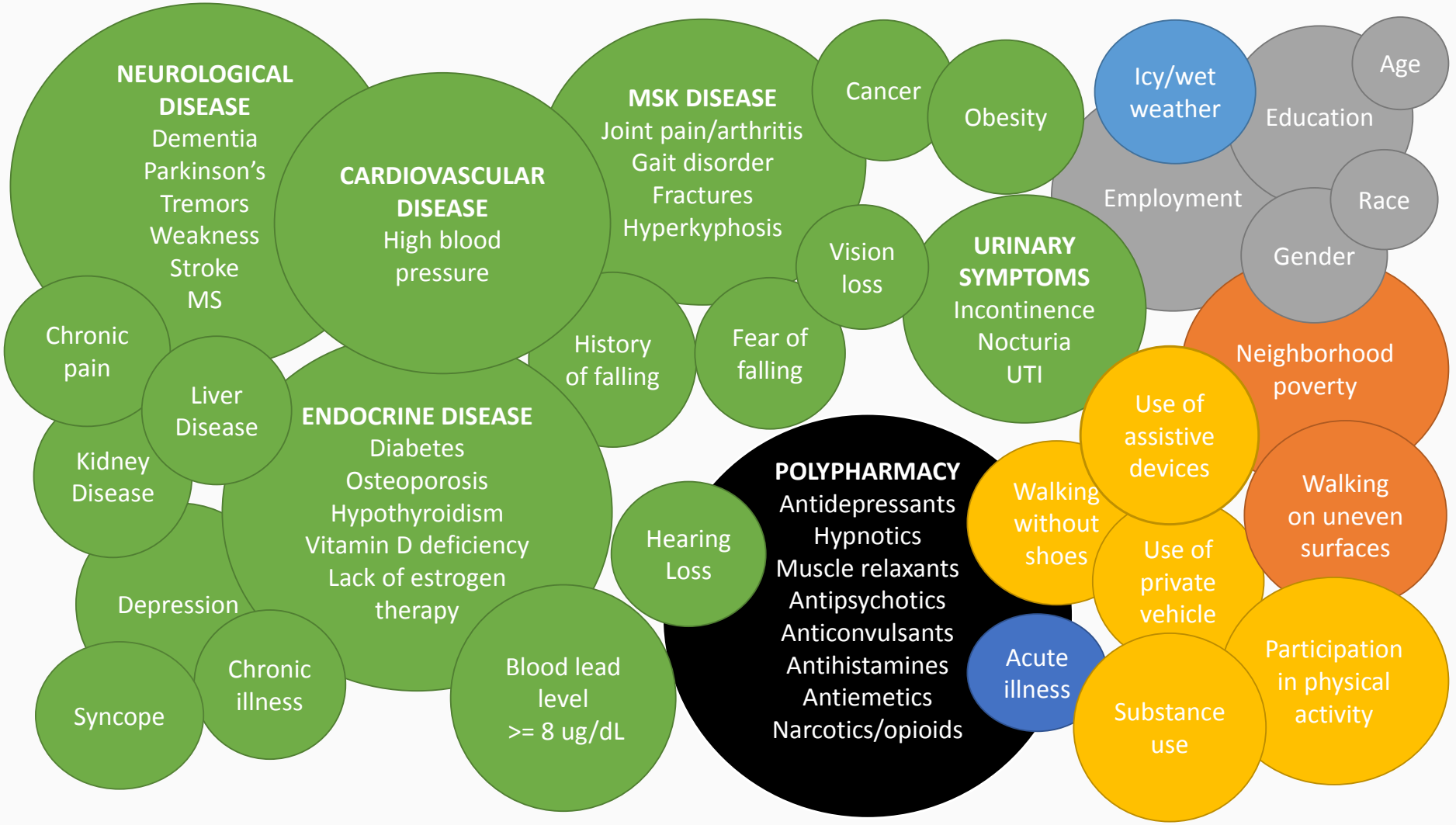
Prevalence of falls among elderly in Maryland (Census Block Group)

CPHIT Portfolio → Predicting Elderly Falls *(cont.)*

Predictors	Estimate	Std. error	z value	Pr(>|z|)	Significance	OR	2.50%	97.50%
History of fall	1.795	0.074	24.113	<2e-16	***	6.02	5.20	6.97
Fracture	0.604	0.104	5.821	5.85E-09	***	1.83	1.49	2.24
Substance Abuse	0.520	0.082	6.364	1.96E-10	***	1.68	1.43	1.97
Parkinson	0.337	0.178	1.895	0.058056	.	1.40	0.98	1.97
Kyphoscoliosis	0.322	0.153	2.102	0.035519	*	1.38	1.01	1.85
Sex (female)	0.173	0.046	3.736	0.000187	***	1.19	1.09	1.30
Depression	0.146	0.068	2.141	0.032238	*	1.16	1.01	1.32
Mental Illness	0.128	0.065	1.980	0.047652	*	1.14	1.00	1.29
Age	0.038	0.003	14.895	<2e-16	***	1.04	1.03	1.04
Charlson Index	-0.053	0.009	-5.711	1.12E-08	***	0.95	0.93	0.97
Vision	-0.211	0.057	-3.689	0.000225	***	0.81	0.72	0.91
Obesity	-0.251	0.076	-3.311	0.000931	***	0.78	0.67	0.90
Cardiovascular Disease	-0.313	0.050	-6.301	2.95E-10	***	0.73	0.66	0.81
Lower Urinary Tract Symptoms	-0.345	0.074	-4.656	3.23E-06	***	0.71	0.61	0.82
Hypertension	-0.357	0.050	-7.080	1.44E-12	***	0.70	0.63	0.77
Cancer	-0.441	0.081	-5.418	6.02E-08	***	0.64	0.55	0.75
Lower Back Pain	-0.495	0.067	-7.368	1.73E-13	***	0.61	0.53	0.69
Joint Trauma	-0.526	0.197	-2.674	0.007487	**	0.59	0.39	0.85
Lower Extremity Joint Surgery	-1.069	0.182	-5.870	4.36E-09	***	0.34	0.24	0.48
(Intercept)	-4.372	0.197	-22.249	<2e-16	***	0.01	0.01	0.02
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1								

Predictors and coefficients of the B'FRIEND model

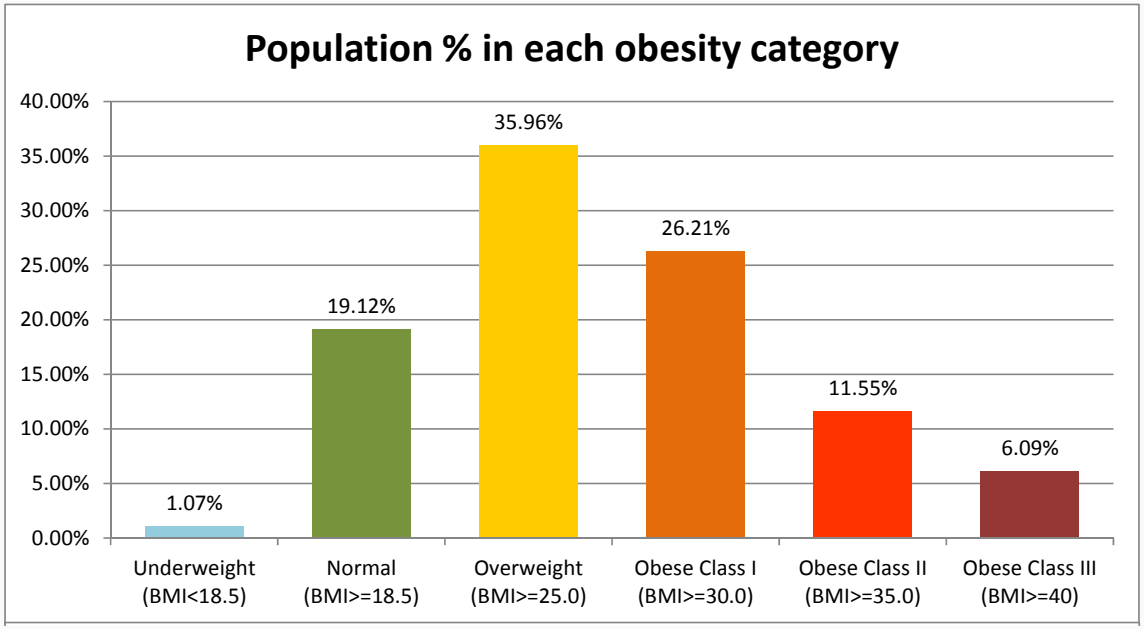
CPHIT Portfolio → Predicting Elderly Falls *(cont.)*



B'FRIEND data sources

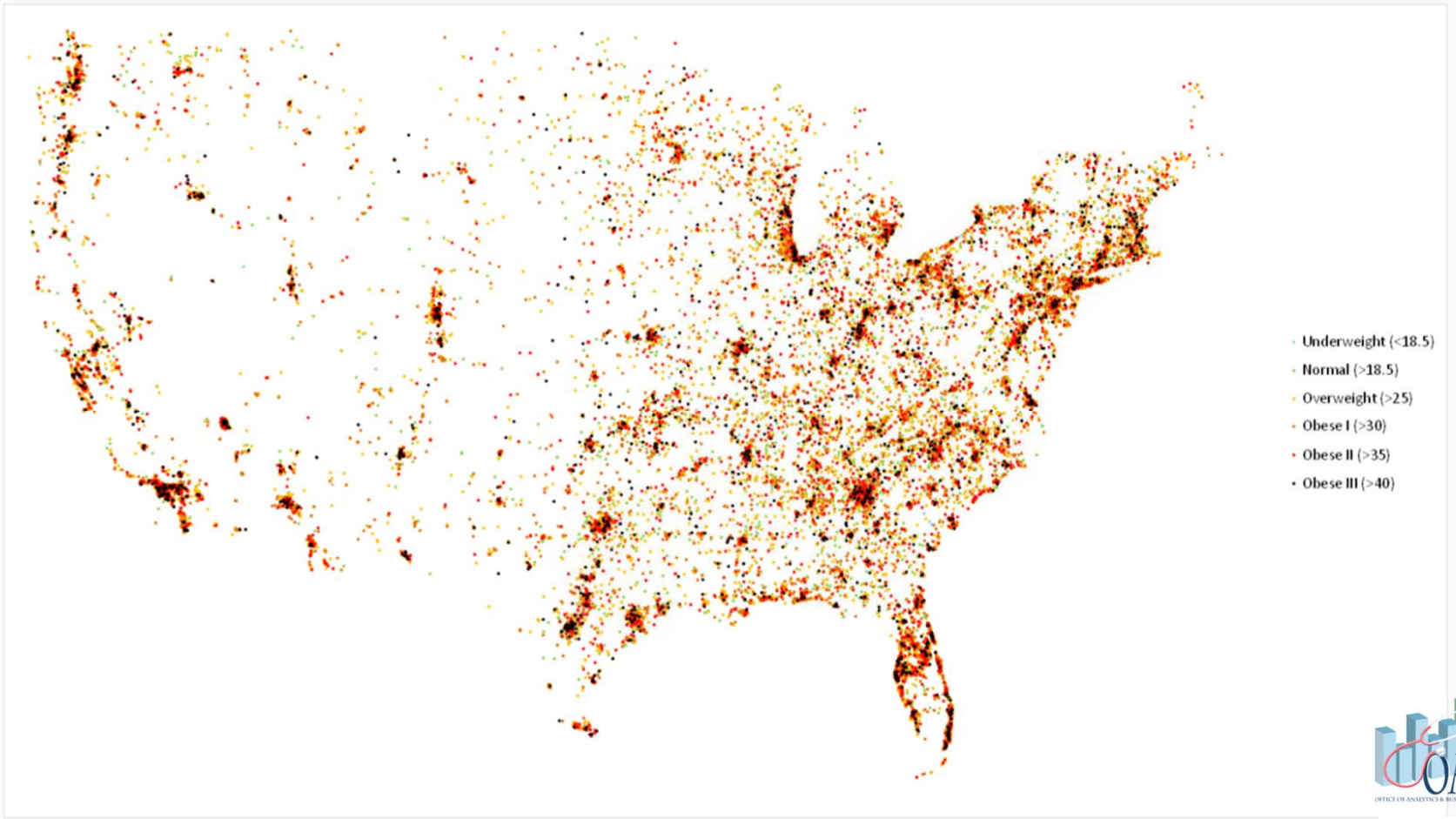
CPHIT Portfolio → VHA Obesity

- Aim (1): Contextualize obesity factors within **VHA's population health framework**
- Aim (2): Design a **scalable pop-health 'technical' platform** and develop a pilot for obesity
- Aim (3): Derive and evaluate "**VHA's Obesity Trajectory Population-based Risk Prediction Model**" to measure the GIS-clustered population-based factors that affect the management of obesity



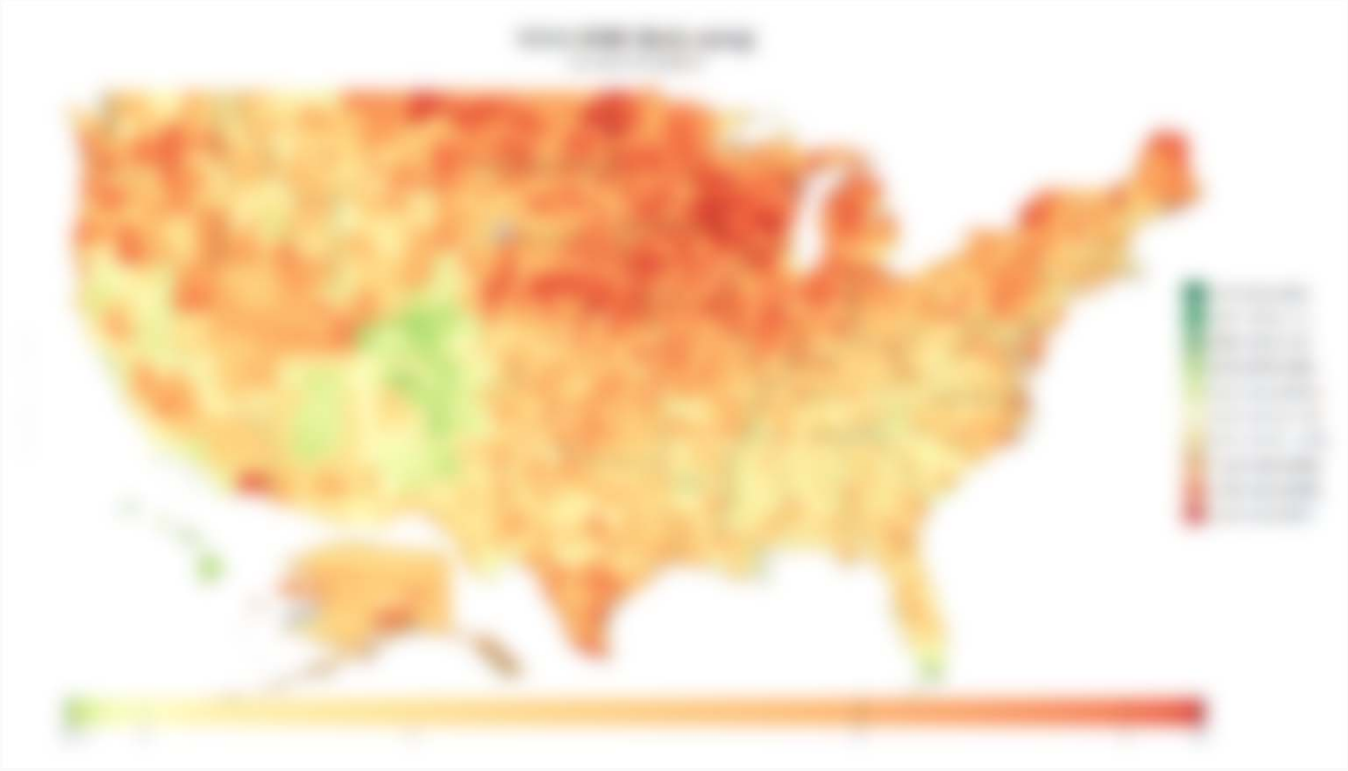
Prevalence of obesity among VHA population
(Limited to 29,322 visits occurred in one day of 2013; generated using CDW data)

Ongoing Projects → VHA Obesity *(cont.)*



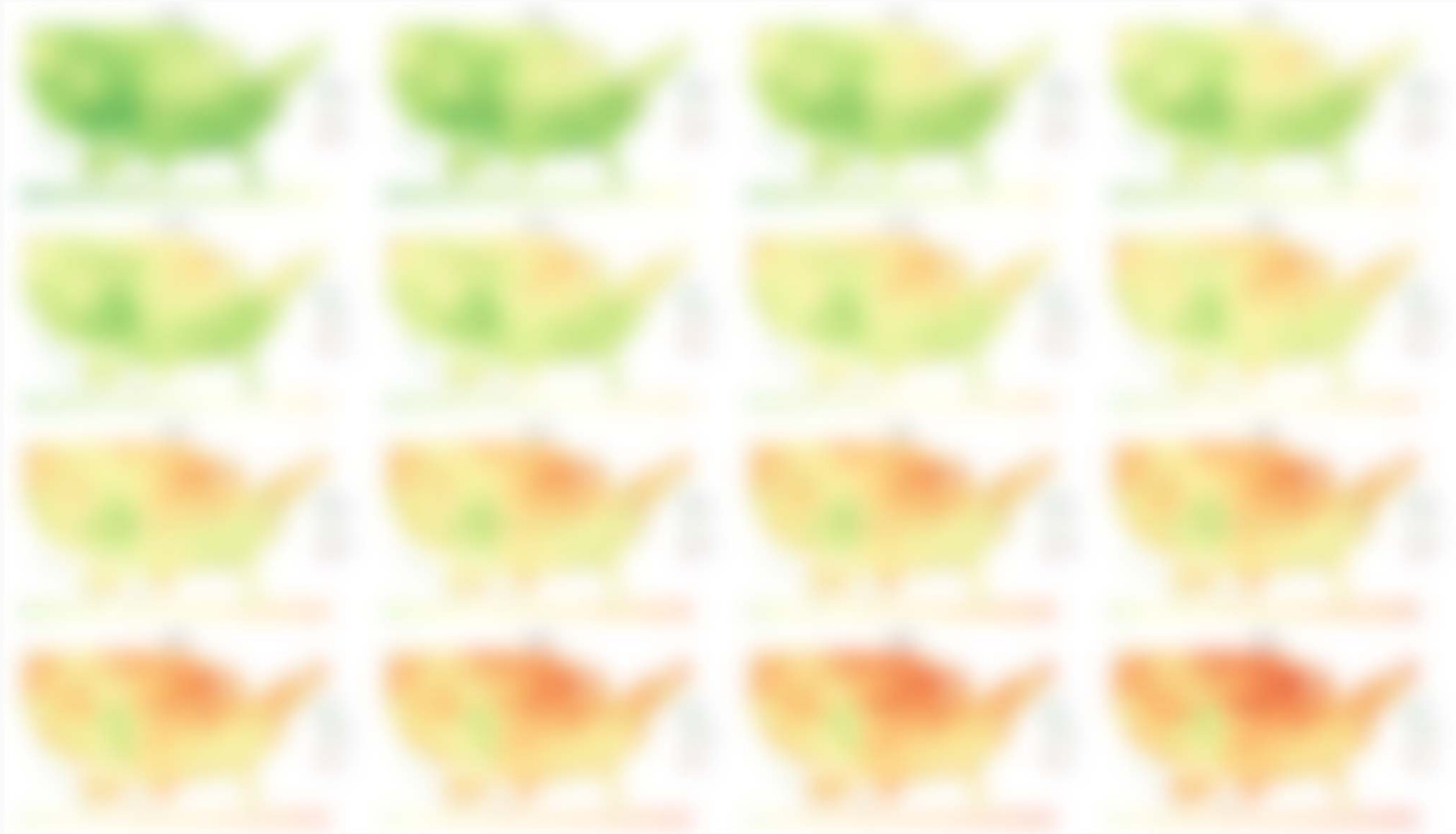
Geographic distribution of obesity among VHA population
(Limited to 29,322 visits occurred in one day of 2013; generated using CDW data)

Ongoing Projects → VHA Obesity (cont.)



County BMI (using MLM adjustment) for Males 2000-2015

Ongoing Projects → VHA Obesity *(cont.)*

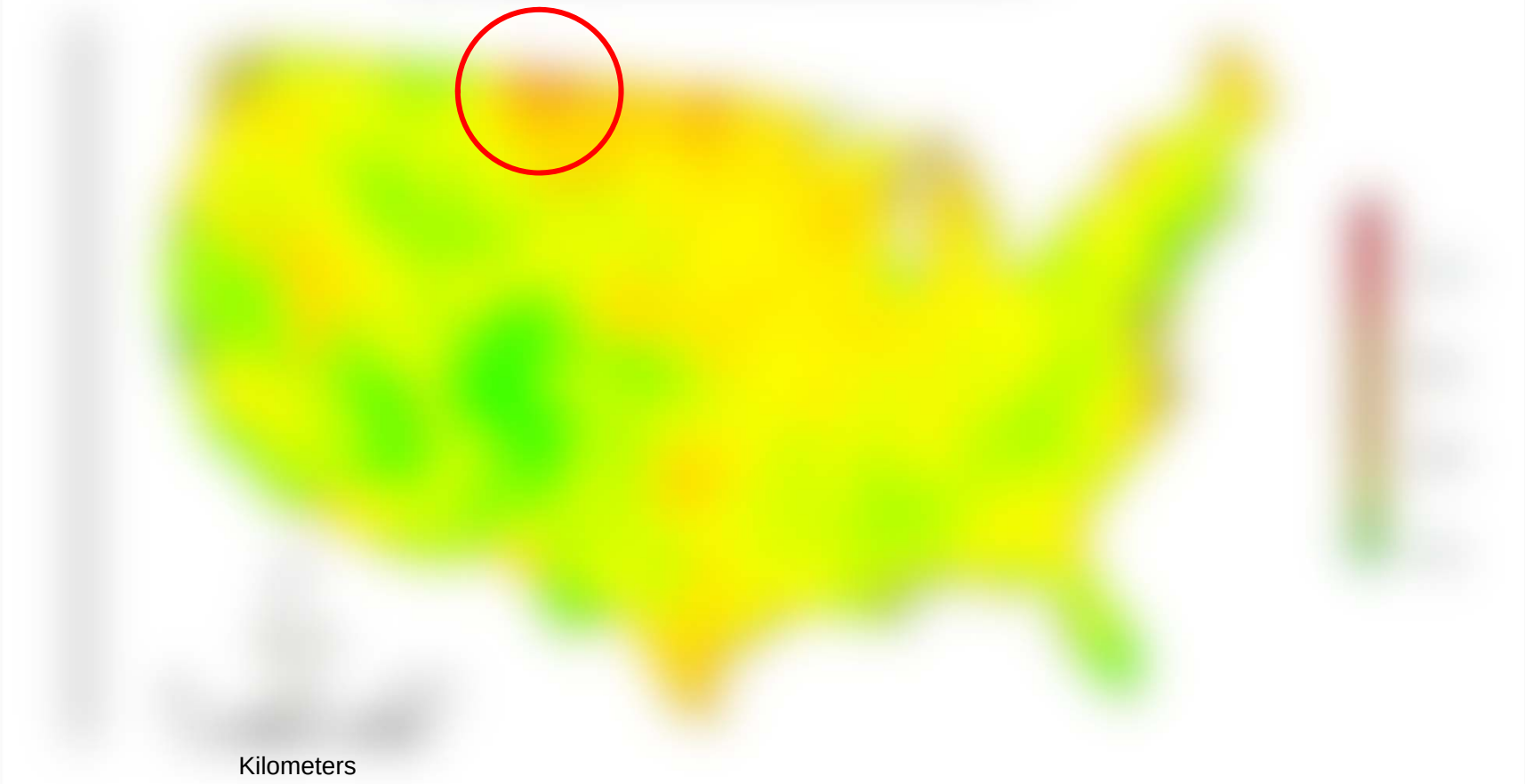


County BMI (using MLM adjustment) for Males 2000-2015

Ongoing Projects → VHA Obesity *(cont.)*

County BMI (using MLM adjustment) for Males 2015 (DC and Baltimore)

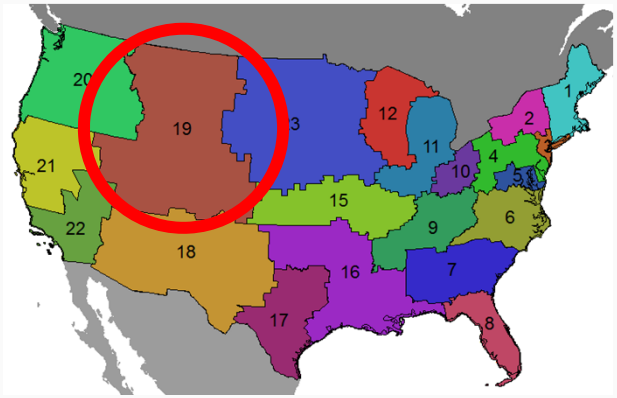
Ongoing Projects → VHA Obesity *(cont.)*



Spatial Intensity Male 2015

Ongoing Projects → VHA Obesity *(cont.)*

Spatial intensity of cases



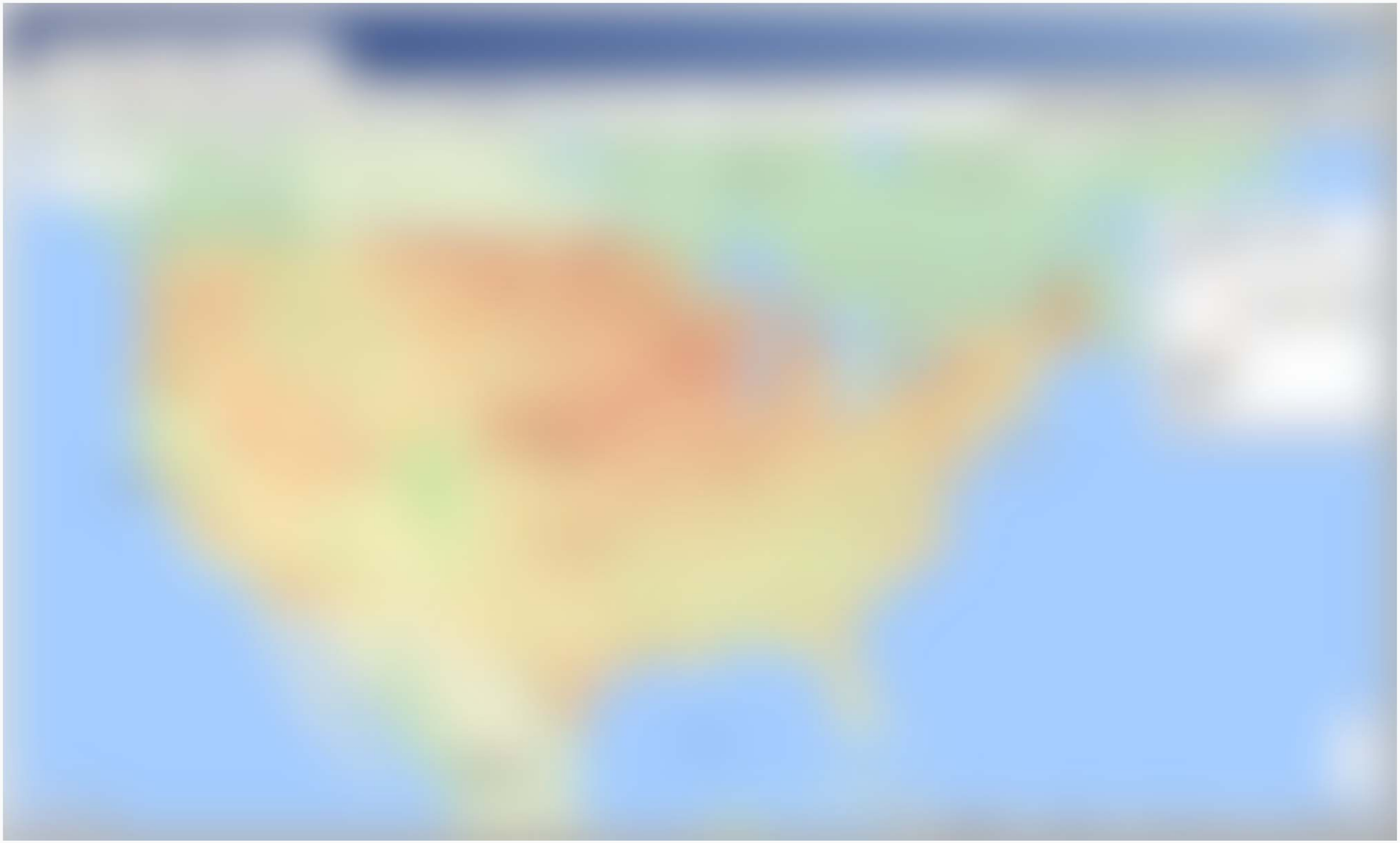
Spatial odds for obesity

Spatial intensity of non-cases

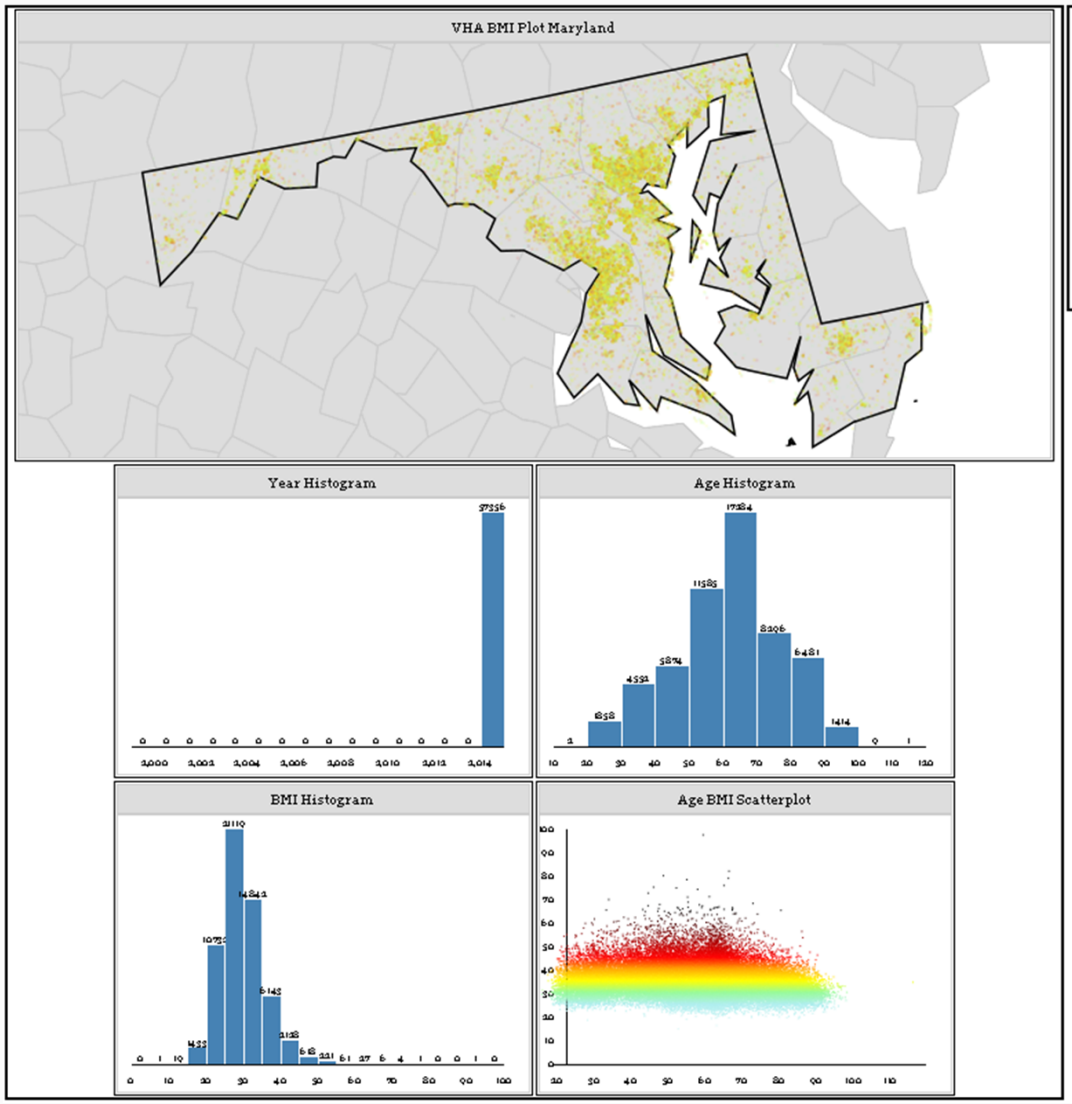
Bandwidth: 50 km

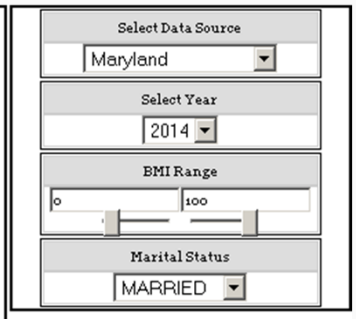
VISN #19: VA Rocky Mountain Network: (Colorado, Montana, Utah, Wyoming)

Ongoing Projects → VHA Obesity *(cont.)*



Ongoing Projects → VHA Obesity *(cont.)*





Name	Owner
VHA Corporate Data Warehouse	VHA
American Community Survey	Census
Census 2010	Census
National Health and Nutrition Examination Survey	CDC
Food Access Research Atlas + Others	USDA
National Vital Statisitcs Report	CDC
Reference USA	RefUSA
Open Street Map	OpenMap
Moderate Resolution Imaging Spectroradiometer	NASA
Consumer Expenditure Survey	BLS
Uniform Crime Reporting Statistics (FBI)	FBI
Maryland Food Systems	MD
USDA Detailed Maps Baltimore	USDA
ArcGIS Internal Datasets	ESRI
Satellite data	Google

Interactive Web-based Real-time Geo-Temporal Exploration of Obesity Data
(Showing averages of 2014 for MD)

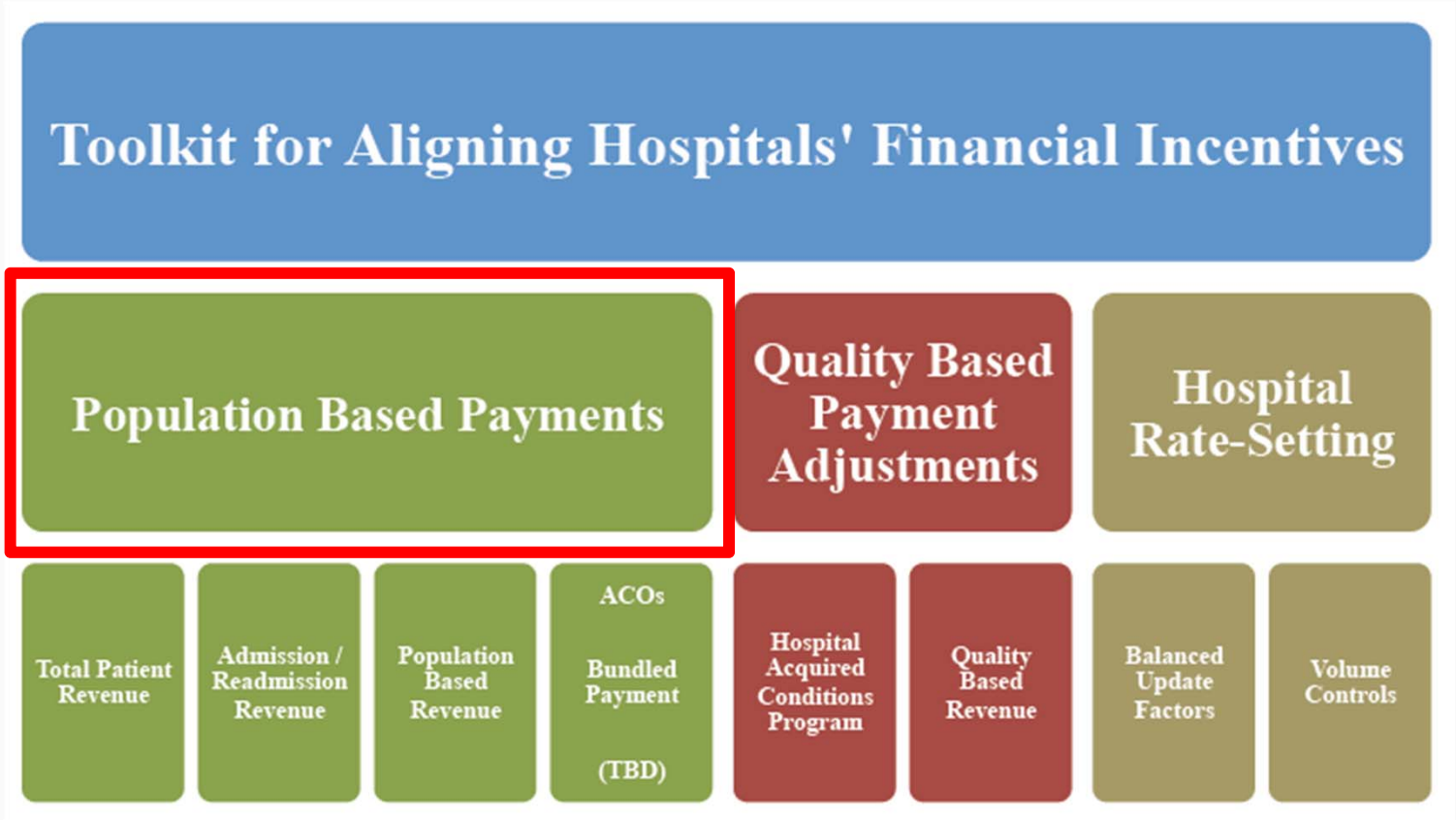
CPHIT Portfolio → Population Health Metrics

- **ACA** → Triple aims → ACO/PCMH

- **CMS** → Value-based Purchasing
  - Bundled models → Salaried models
  - Shared savings (MSSP) and/or Global payments (Pioneers)
  - Blended payments (Capitation + FFS)

- **MD's DHMH** → Maryland's All-Payer In-Patient Waiver program
  - Limit the growth of per capita hospital cost @ 3.58% (10 year compound MD's GDP) → saving Medicare $330mil over the next 5 years
  - Shifting away from fee-for-service models into population-based payment models that reward providers for improving health outcomes → 95% of hospitals are on global budget
  - Immense need for population health management and IT infrastructure to operationalize it

CPHIT Portfolio → Population Health Metrics *(cont.)*



Maryland's All-Payer Waiver Program – Toolkit to align financial incentives

CPHIT Portfolio → Population Health Metrics *(cont.)*

Population Health IT Infrastructure ?

Global Payment Strategies

Method	Services Included	Hospital Participating	Estimated Percent of Revenue at Risk
Total Patient Revenue (TPR)	All regulated services	10	~100%
Admission/Readmission Revenue (ARR)	All-cause readmissions for 30 days	31	~10%
Population Based Revenue (PBR) and Other Global Models	Core services for specific DRGs in hospital community	TBD	~ 30% to 100% (estimated)
Quality Programs with Revenue at Risk (Quality Based reimbursement and Maryland Hospital Acquired Conditions)	All inpatient regulated services *State will expand to all regulated services in future years*	All	For each performance year, Maryland will place the same percentage of hospital revenue at risk as the national Medicare Value-Based Purchasing Program, Hospital Acquired Condition and Readmission Reduction programs a
Balanced Update Factors	All regulated services	All	N/A
Volume Controls	All regulated services under the models	All non-TPR/Global revenues	N/A

Maryland's All-Payer Waiver Program – Global Payment Strategies

CPHIT Portfolio → Population Health Metrics *(cont.)*

The NEW ENGLAND JOURNAL of MEDICINE

Perspe

Maryland's All-Payer Approach to Delivery-System R

Rahul Rajkumar, M.D., J.D., Ankit Patel, J.D., Karen Murphy, Ph.D., John M. Colmers, M.P.H.
Jonathan D. Blum, M.P.P., Patrick H. Conway, M.D., and Joshua M. Sharfstein, M.D.

On January 10, 2014, the Centers for Medicare and Medicaid Services (CMS) and the State of Maryland jointly announced the launch of a statewide model that will transform Maryland's health

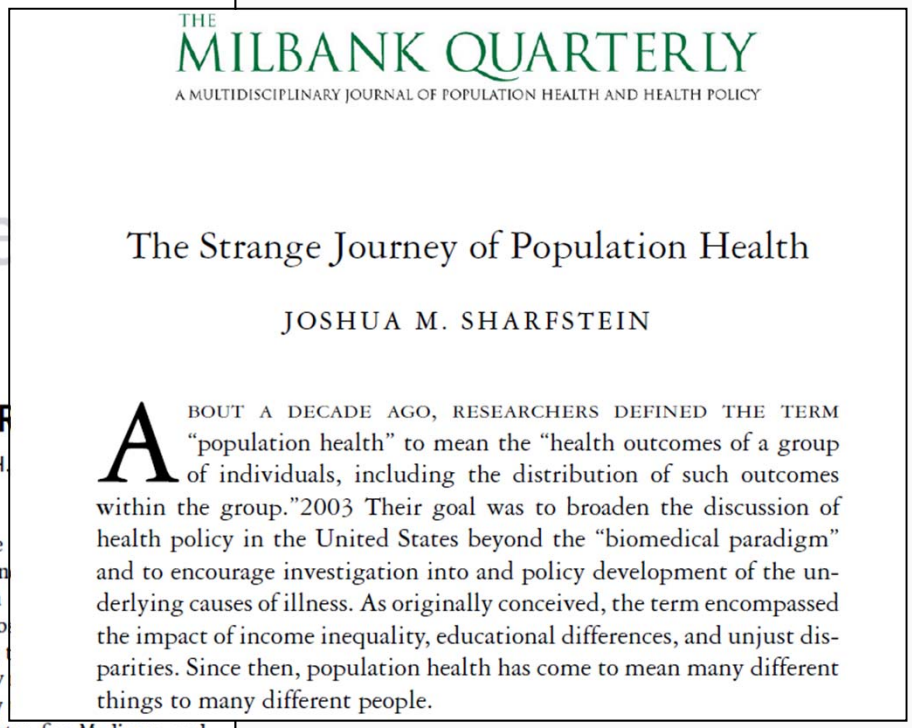
care delivery system. Although some aspects of the new approach may be unique to Maryland and not applicable elsewhere, both the principles of this model and the process that led to its development

in Medicare payment per hospital admission.

This system has eliminated cost shifting among payers, more equitably spread the costs of uncompensated care and medical

prospective
inpatient an
resulted in
hospital co
are among t

The new
possible by
to the Center for Medicare and Medicaid Innovation under the Affordable Care Act, will change the basis for Medicare's participation in Maryland's system. In place



THE MILBANK QUARTERLY

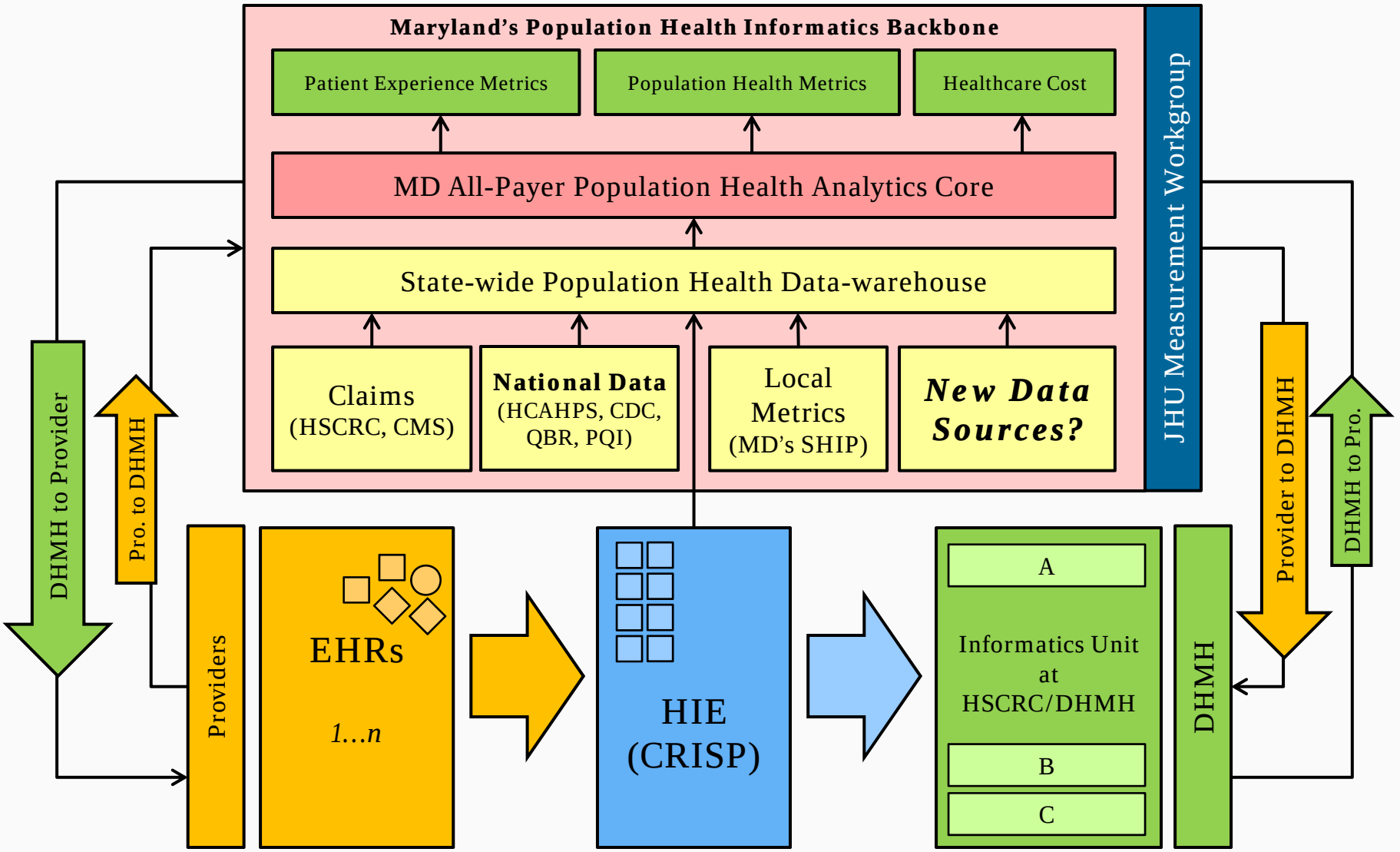
A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY

The Strange Journey of Population Health

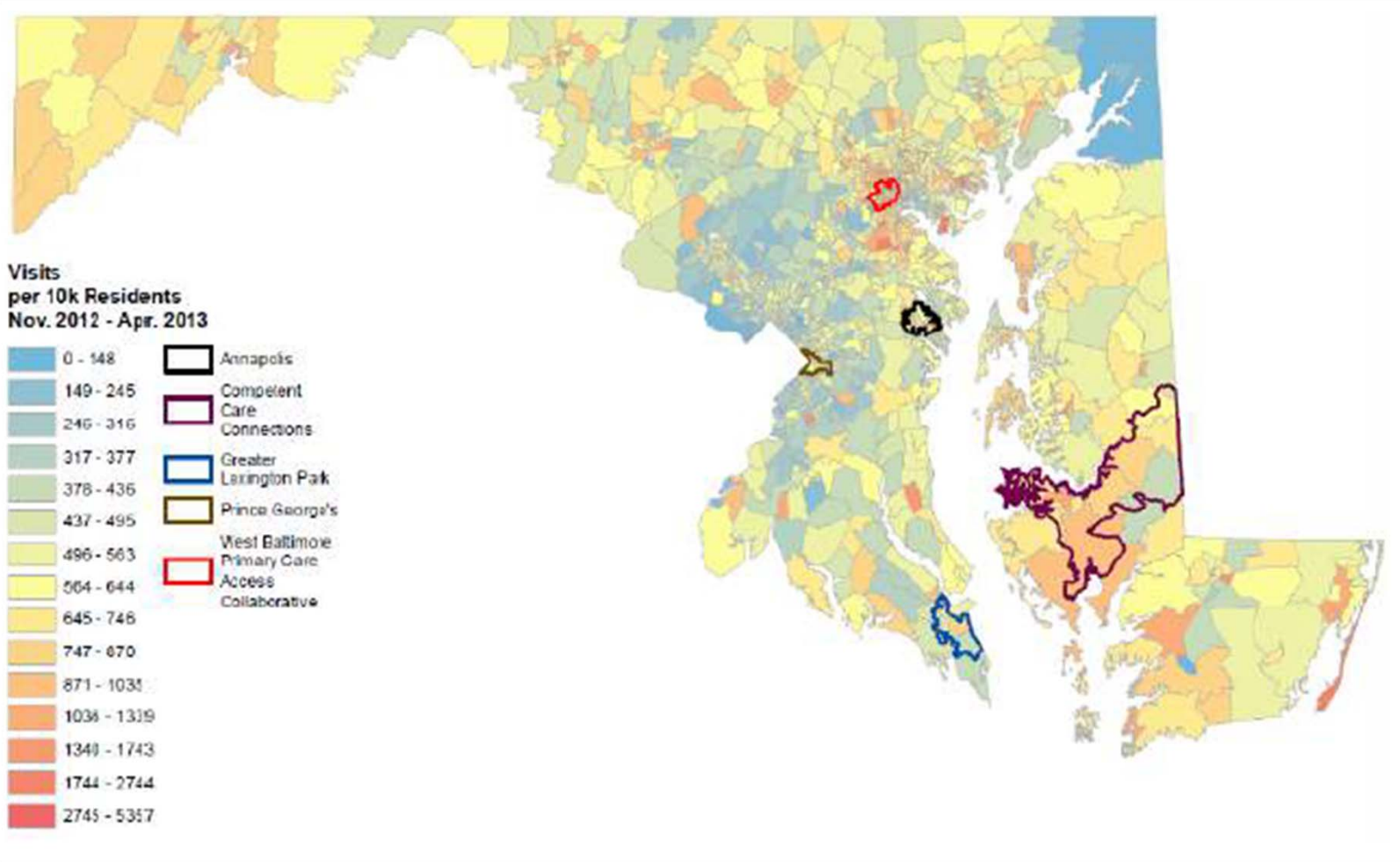
JOSHUA M. SHARFSTEIN

ABOUT A DECADE AGO, RESEARCHERS DEFINED THE TERM "population health" to mean the "health outcomes of a group of individuals, including the distribution of such outcomes within the group."2003 Their goal was to broaden the discussion of health policy in the United States beyond the "biomedical paradigm" and to encourage investigation into and policy development of the underlying causes of illness. As originally conceived, the term encompassed the impact of income inequality, educational differences, and unjust disparities. Since then, population health has come to mean many different things to many different people.

CPHIT Portfolio → Population Health Metrics *(cont.)*

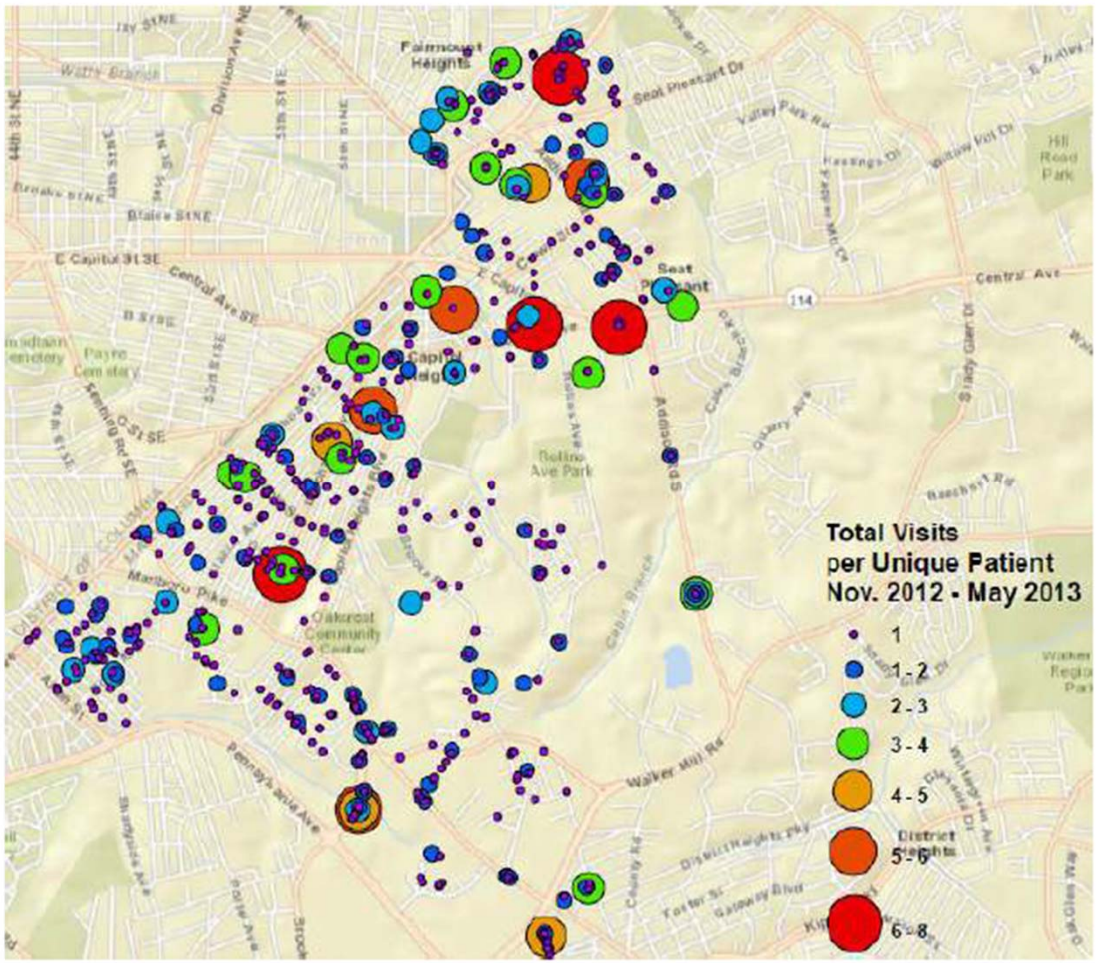


CPHIT Portfolio → Population Health Metrics *(cont.)*



Inpatient Utilization By Census Tract – State Level View

CPHIT Portfolio → Population Health Metrics *(cont.)*



Inpatient Utilization By Census Tract – Neighborhood View (Capitol Heights Area)

CPHIT Portfolio *(cont.)*

- **National Pop Health IT Curriculum and Workforce Training**: Developing a national curriculum for Population Health IT and training more than 7000 healthcare professional incumbents
- https://www.healthit.gov/providers-professionals/health-it-education-opportunities
- http://learnhit.com

Discussion

Discussion → Challenges and Opportunities *(cont.)*

- **Data sources/types:**
  - How to compare data types and their added value?
  - What are the limits of each data type? What are we missing?
  - What can be used from unstructured data?
- **Data quality:**
  - How much juice is left in this data type (e.g., claims)?
  - Do objective measures have data quality issues (e.g., BMI)?
  - How can we measure the quality of subjective data?
- **Denominator:**
  - Are we excluding noise or signal?
  - Is this a too big of a cut or too narrow – sample size issues?
  - Patient attribution issues…

Discussion → Challenges and Opportunities *(cont.)*

- **Comparing models:**
  - What are the guidelines?
  - How to compare models across different subpopulations?
- **Feature reduction:**
  - How can we mix multiple approaches such as expert opinion + automated approaches to reduce the feature space?
- **Temporal data:**
  - What window is appropriate?
  - How to deal with large zero fills in temporal data?
- **Privacy and Security:**
  - Is HIPAA helping or hurting HPM's data scientists?

Discussion → Challenges and Opportunities *(cont.)*

- **Big Data:**
  - Data volume/quantity (e.g., VHA data query)
  - Data veracity/quality (e.g., VHA BMI data)
  - Data variety/sources (e.g., EHR vs Claims)
  - Data velocity/timeliness (e.g., VHA vital signs)
  - Unstructured data → NLP and HIPAA De-identification issues

A proposed national research and development agenda for population health informatics: summary recommendations from a National Expert Workshop

RECEIVED 6 August 2015
REVISED 17 December 2015
ACCEPTED 21 December 2015

Hadi Kharrazi[1,2], Elyse C Lasser[1], William A Yasnoff[2,3], John Loonsk[1], Aneel Advani[1], Harold P Lehmann[2], David C Chin[1], Jonathan P Weiner[1,2]

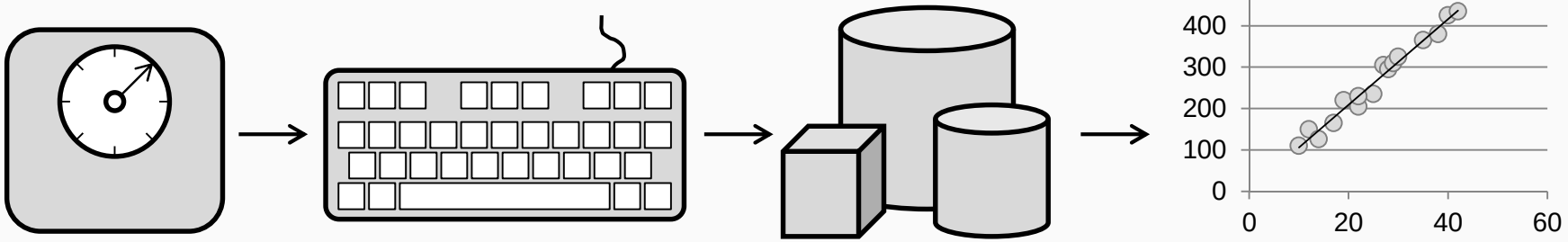
Discussion → Data Quality

Actual value:
200.6 lbs.

Recorded value:
200 lbs.

Data warehouse value:
200 kg

Analytic value:
100 kg (mean 200 & 0)



Measured (same day)
• Validity challenge
198.9 | 198.9 | 198.9 lbs.
• Reliability challenge
200.6 | 198.9 | 202.2 lbs.

Measured (diff. days)
• Missing data challenge
NULL | 200.6 | NULL lbs.

User Typed (one entry)
• Typos
200.6 lbs. → 20.06, 2006
• Mismatching units
200.6 lbs. → 200.6 kg
• Assumptions/truncations
200.6 lbs. → 200 lbs.
NULL → 0
• Free-text additions
200.6 lbs. → 200.6 pounds

DB Operations (one entry)
• Truncations/Rounding
200.6 → 200.0
• Error conversions
200.6 pounds → NULL
200.6 lbs. → 200.6 kg
• Cleaning
200+ lbs. → 200.0

Analytics (data points)
• Aggregation of data points
200 | 0 → mean of 100
• Selecting a representative
190 | 200 | 210 → 210 (first)
190 | 200 | 210 → 200 (mean)
190 | 200 | 210 → 210 (last)
• Removing outliers
200 | 200 | 350 → 200 | 200 | NULL

Example of data quality issues across various data collection and use stages

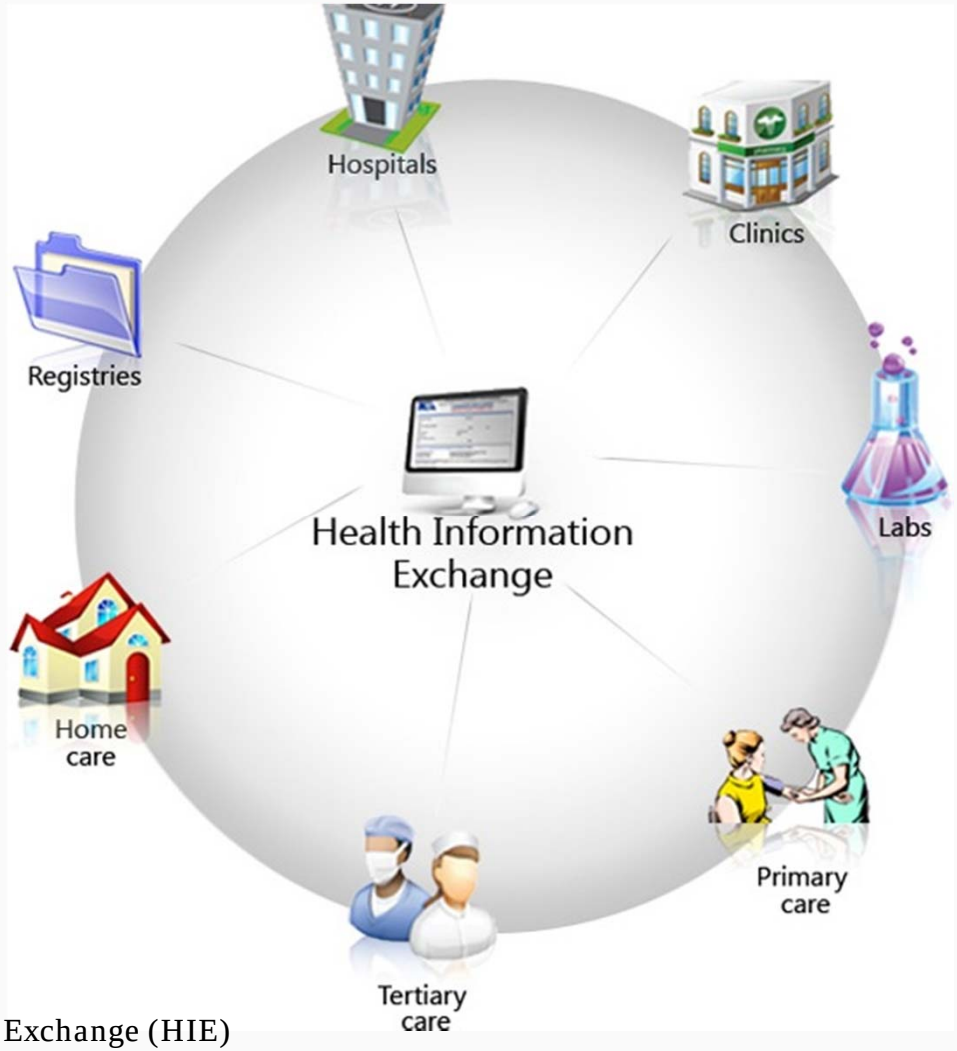
Thank you!

Q & A

kharrazi@jhu.edu
www.jhsph.edu/cphit

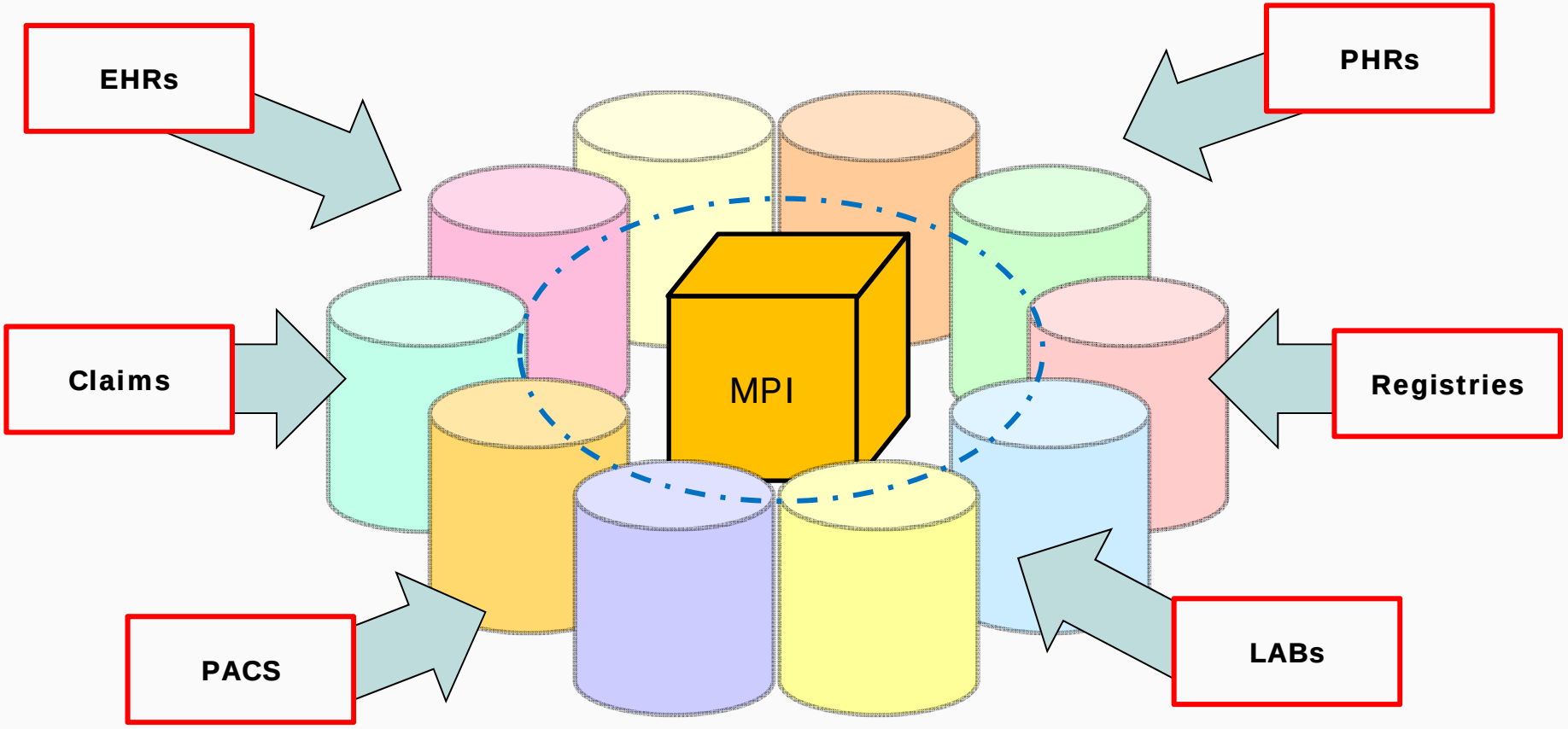
Appendix

CPHIT Portfolio → HIE-based Readmission Prediction



Health Information Exchange (HIE)

CPHIT Portfolio → HIE-based Readmission Prediction *(cont.)*



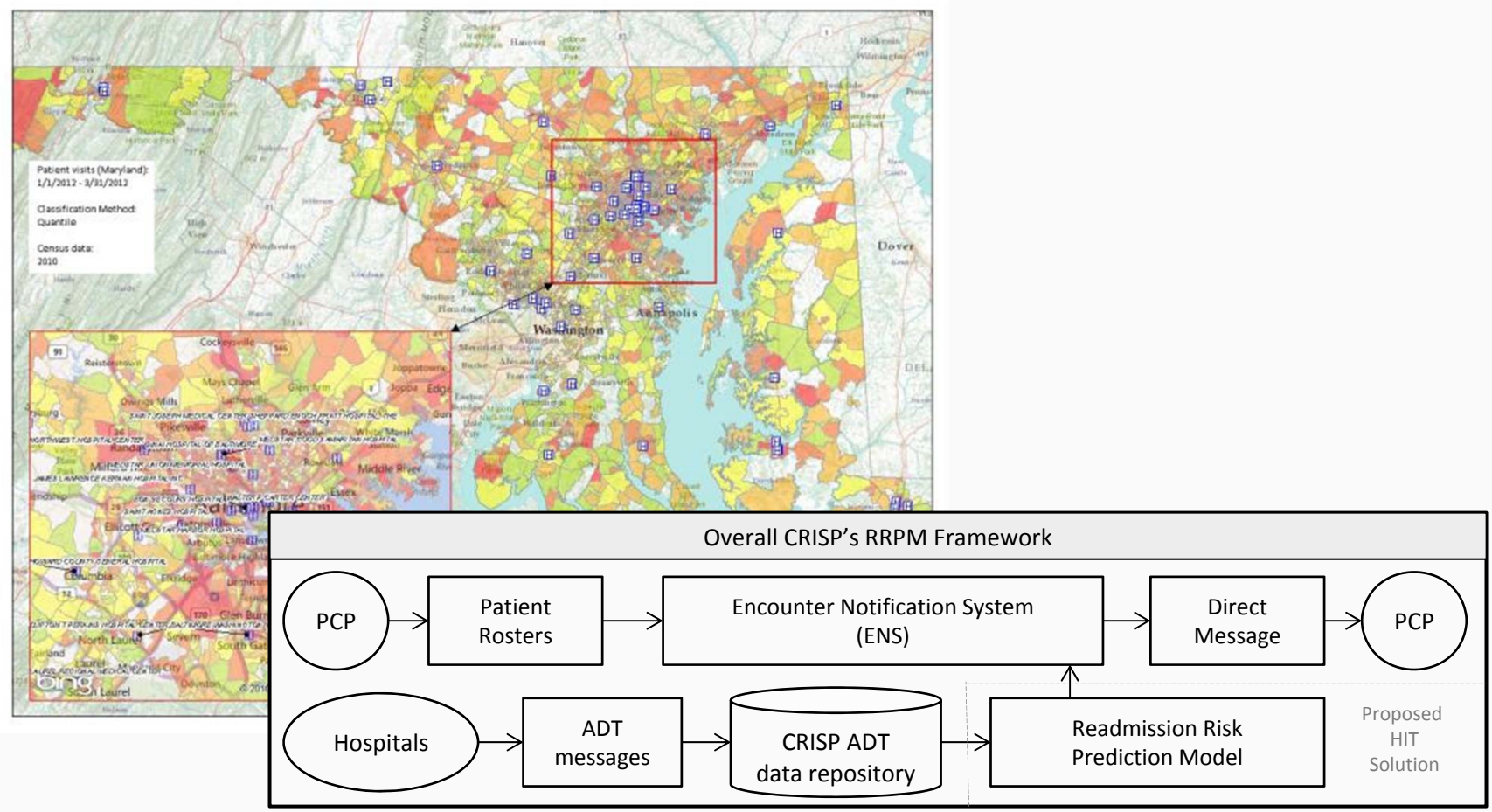
Federated Consistent HIE

CPHIT Portfolio → HIE-based Readmission Prediction *(cont.)*

- **Focus Areas**:
  - Query Portal Growth
  - Direct Secure Messaging
  - Encounter Notification System (ENS)
  - Encounter Reporting System (ERS)
  - Health Benefits Exchange integration

Progress Metric	Result
Organizations Live	
Hospitals (Total 48)	48
Hospital Clinical Data Feeds (Total 143 - Lab, Radiology, Clinical Docs)	86
National Labs	2
Radiology Centers (Non-Hospital)	5
Identities and Queries	
Master Patient Index (MPI) Identities	~4M
Opt-Outs	~1500
Queries (Past 30 Days)	~3500
Data Feeds Available	
Lab Results	~16M
Radiology Reports	~5M
Clinical Documents	~2M

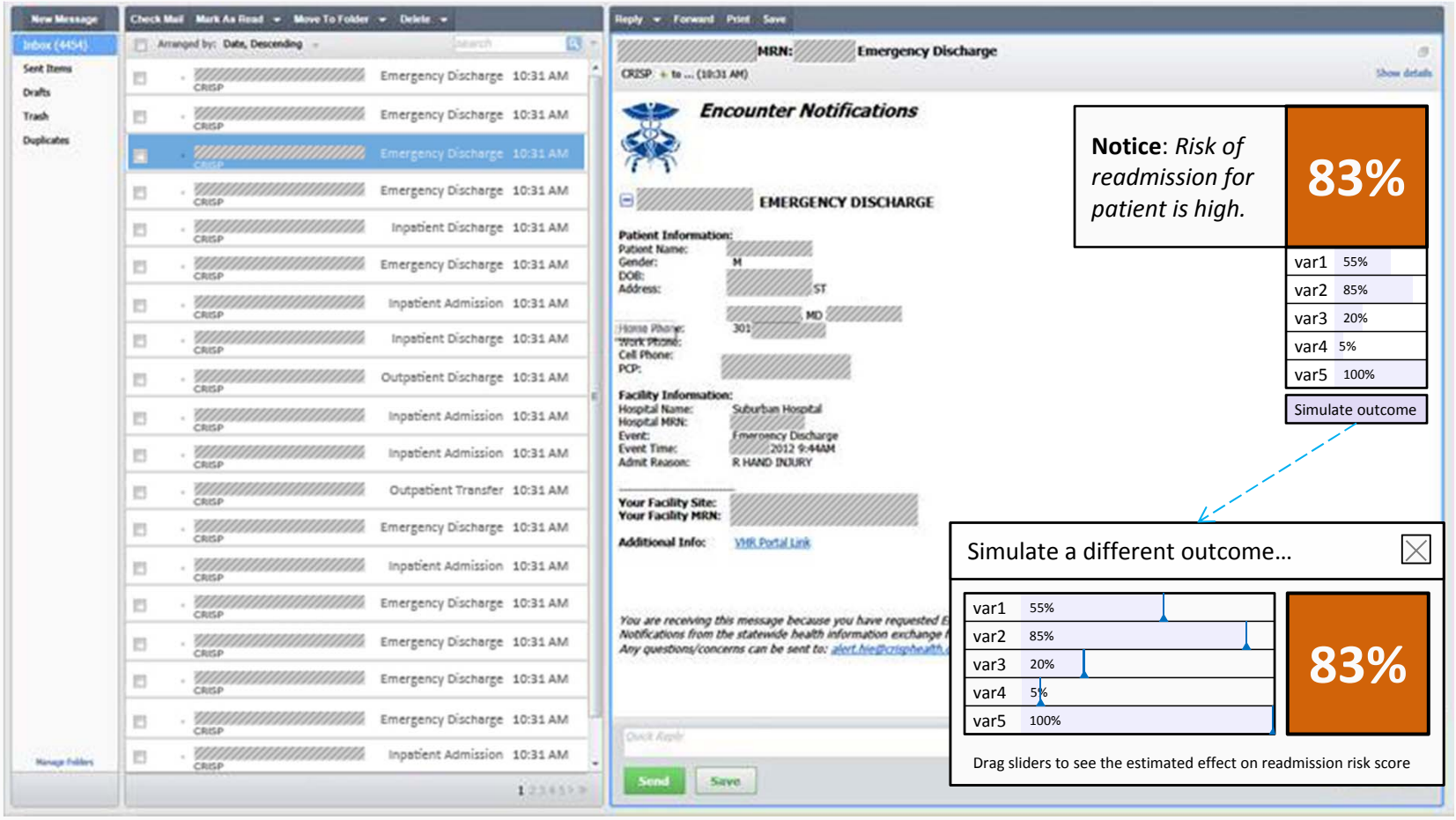
CPHIT Portfolio → HIE-based Readmission Prediction *(cont.)*



CRISP Encounter Notification System + Readmission Risk Prediction Model (RRPM) integration

Copyright CRISP

CPHIT Portfolio → HIE-based Readmission Prediction *(cont.)*



Advance RRPM integration in CRISP Encounter Notification System