UNIVERSITY OF CALIFORNIA

Precisely practicing medicine from 700 trillion points of UC Health data

Atul Butte, MD, PhD

Chief Data Scientist, University of California Health (UC Health)

Historical context

- **BRAID**: Collaborative foundation
- **UC ReX**: Nationally recognized pioneering work, enabled counting patients across the system, harmonization of data elements
- **Other UC-wide projects**: PCORI pScanner, ATHENA, NIH Precision Medicine
- **Health systems**: Launched Leveraging Scale for Value (LSfV) Program
- **Nov 2014 retreat**: BRAID proposal & Precision Medicine: seeded the need to come together
- **Feb 2015 Oakland workshop**: Convened UC healthcare leaders
- **Mar 2015 UC Leadership retreat**: Big Health Data Initiative (BHDI) proposal to UC Health leadership to recognize the potential
- **Aug 2015-May 2016:** UCHDW Proof of Concept (Epic Caboodle)
  - $1.25 million funded through startup funds from UCSF
  - Resource contributions from all five sites
- **Nov 2016:** Funding approved for Pilot
- **Mar 2017 – December 2017:** Pilot UCHDW F2F kick-off and execution
- **Jan 2018:** Center for Data Driven Insights and Innovation (CDI2) created
- **Feb 2018:** UCHDW funding released to sites
- **Aug 2018:** UCHDW in production for operations

UC ReX

University of California Research eXchange

Home | Data Explorer | Our Data | Get Started ★

SEARCH 15 MILLION+ PATIENT RECORDS FROM THE UNIVERSITY OF CALIFORNIA WITH THE UC ReX DATA EXPLORER

ABOUT THE TOOL

ABOUT THE DATA

GET STARTED

UC Health, United Healthcare Form New ACO & Clinically Integrated Networks

by Staff Writer 10/03/2016 0 Comments

UNIVERSITY OF CALIFORNIA UC Health

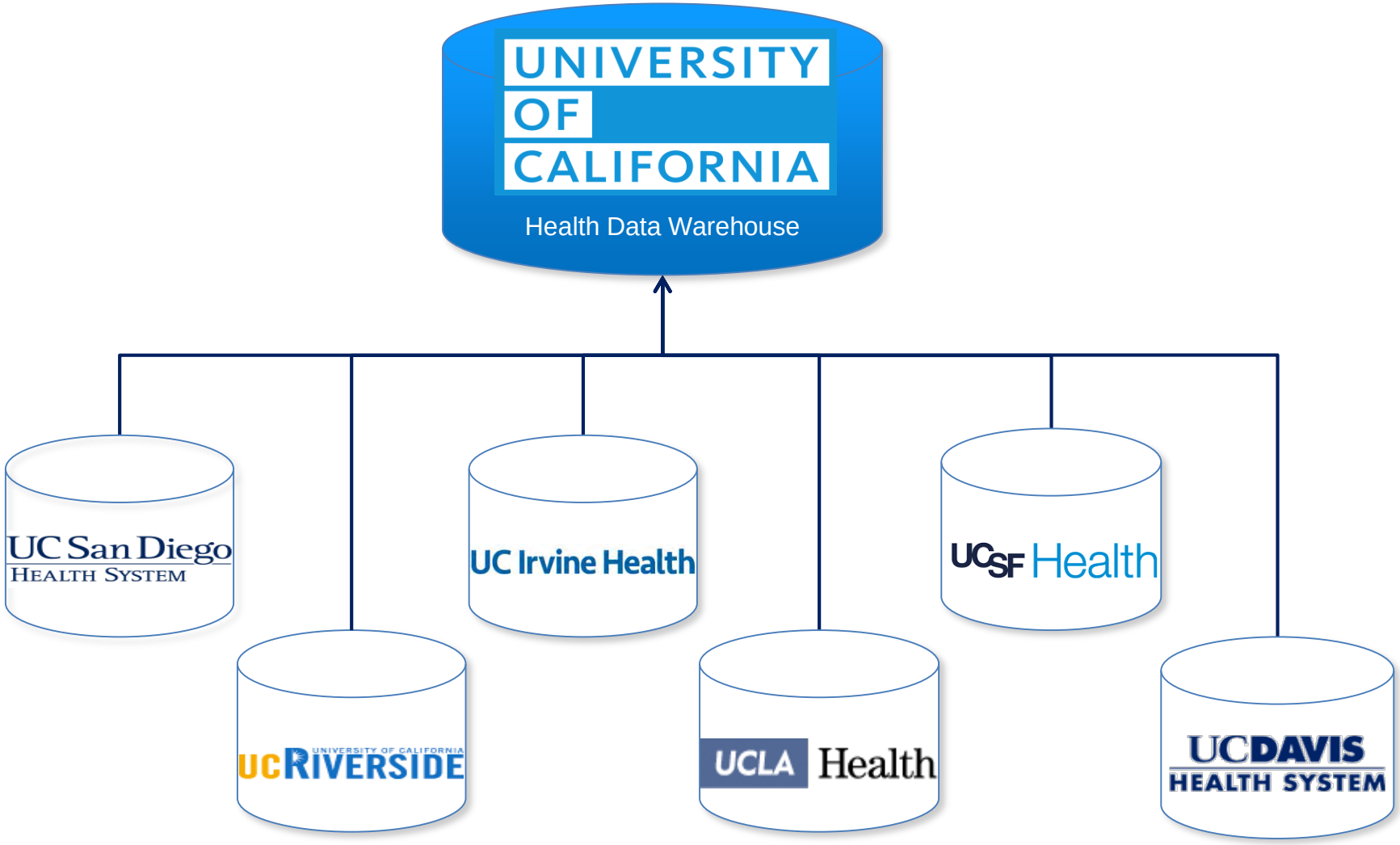
The University of California and UnitedHealth Group are teaming up to form a new accountable care organization (ACO) and clinically integrated network. As part of the 10-year strategic relationship, UC Health's five academic medical centers will expand use of Optum's clinically integrated network services and advanced data analytics services.

What is the UCHDW?

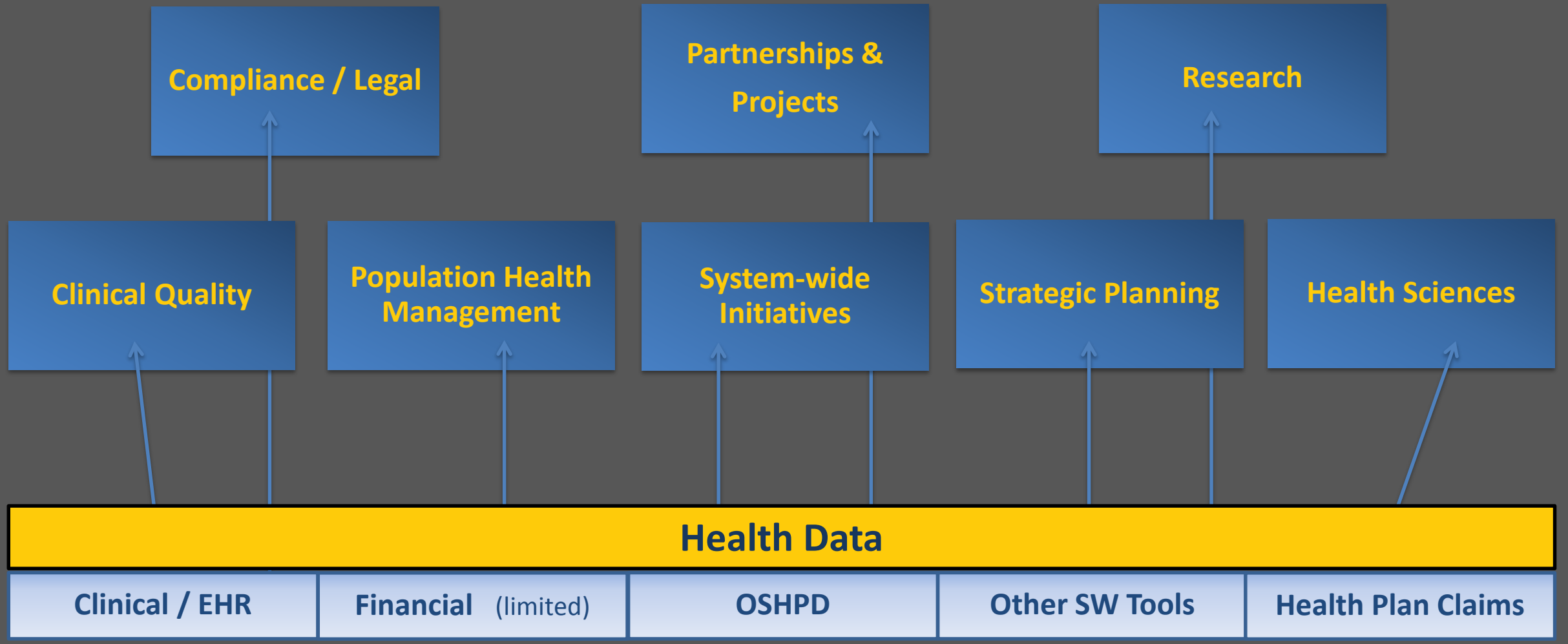
A centralized healthcare data warehouse and analytics platform that supports UC Health strategic data driven initiatives:

- Best practice and operations analysis
- Quality measure and regulatory reporting
- Clinical research

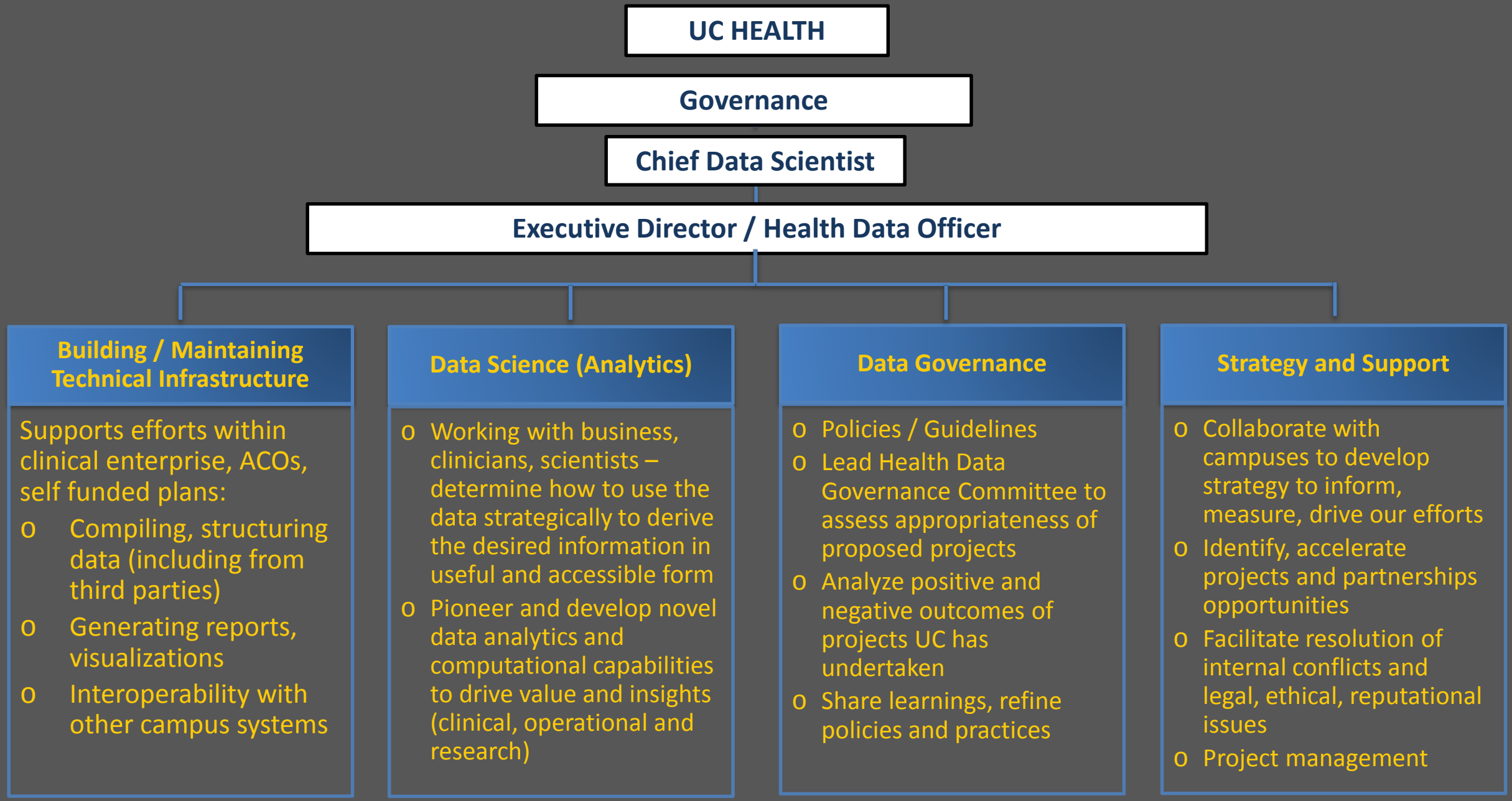
Combining healthcare data from across the six University of California medical schools and health systems



Center for Data-Driven Insights and Innovation



Center for Data-Driven Insights and Innovation



UCHDW Timeline

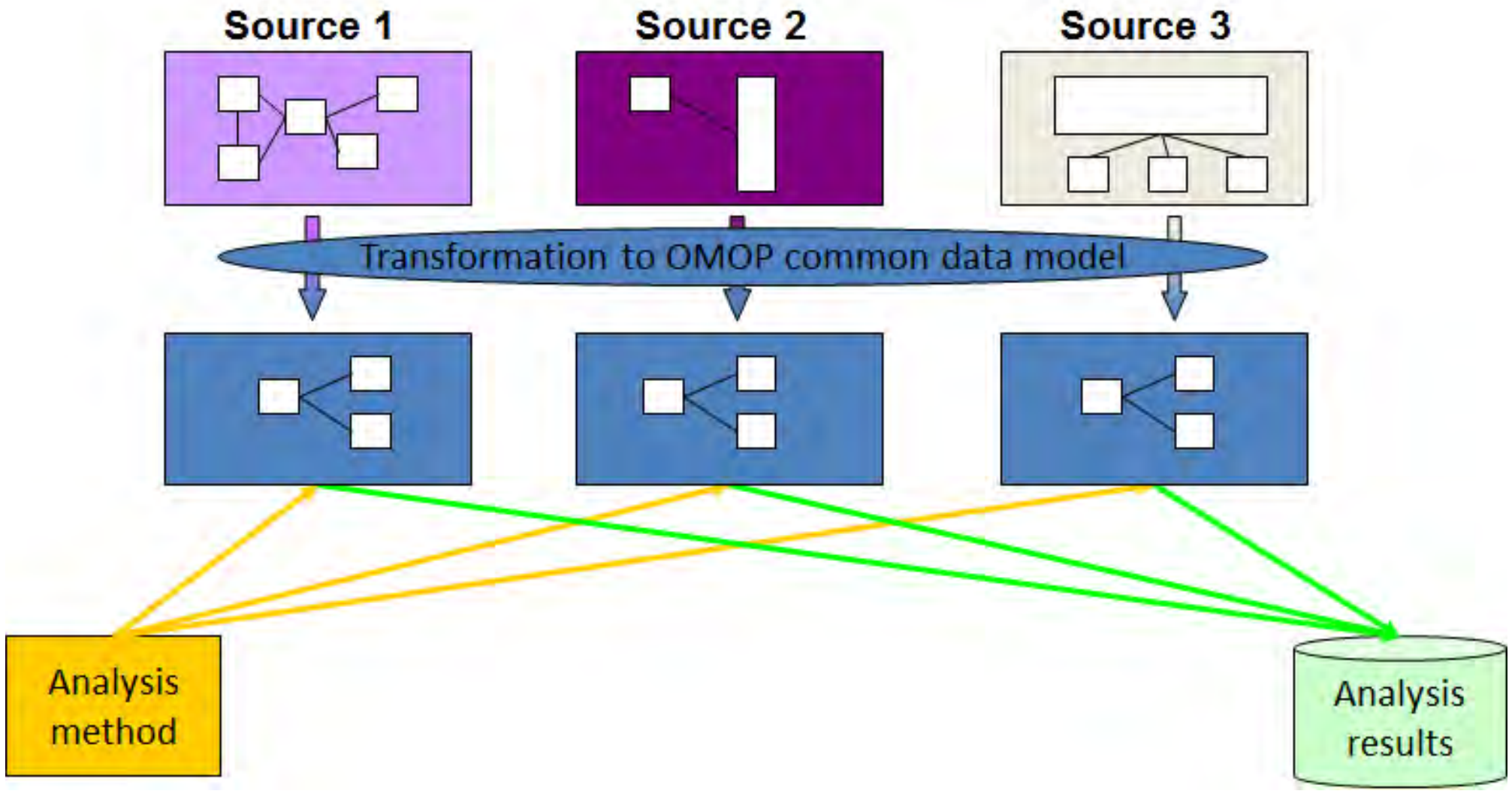
- **August 2015 to February 2017**: Proof of Concept
- **March to December 2017**: UCHDW Pilot
- **January 2018**: Center for Data Driven Insights (CDDI) created and initial budget and funding approved

UC Health Data Warehouse (UCHDW) Campus Leadership		
UC Health Campus	**Executive Sponsor**	**Business Intelligence & Analytics**
UC Davis	John Cook	Kent Anderson; Douglas Berman
UC Irvine	Chuck Podesta	Lisa Dahm; Ayan Patel; Charles Wilson
UC Los Angeles	Mike Pfeffer	Mohammed Mahbouba; Albert Duntugan
UC San Diego	Chris Longhurst	Jennifer Holland; Joshua Glandorf; Eugene Lee
UC San Francisco	Joe Bengfort	David Dobbs; Rick Larsen
UC Health	Atul Butte, Tom Andriola	Lisa Dahm; Ayan Patel

Data Platform

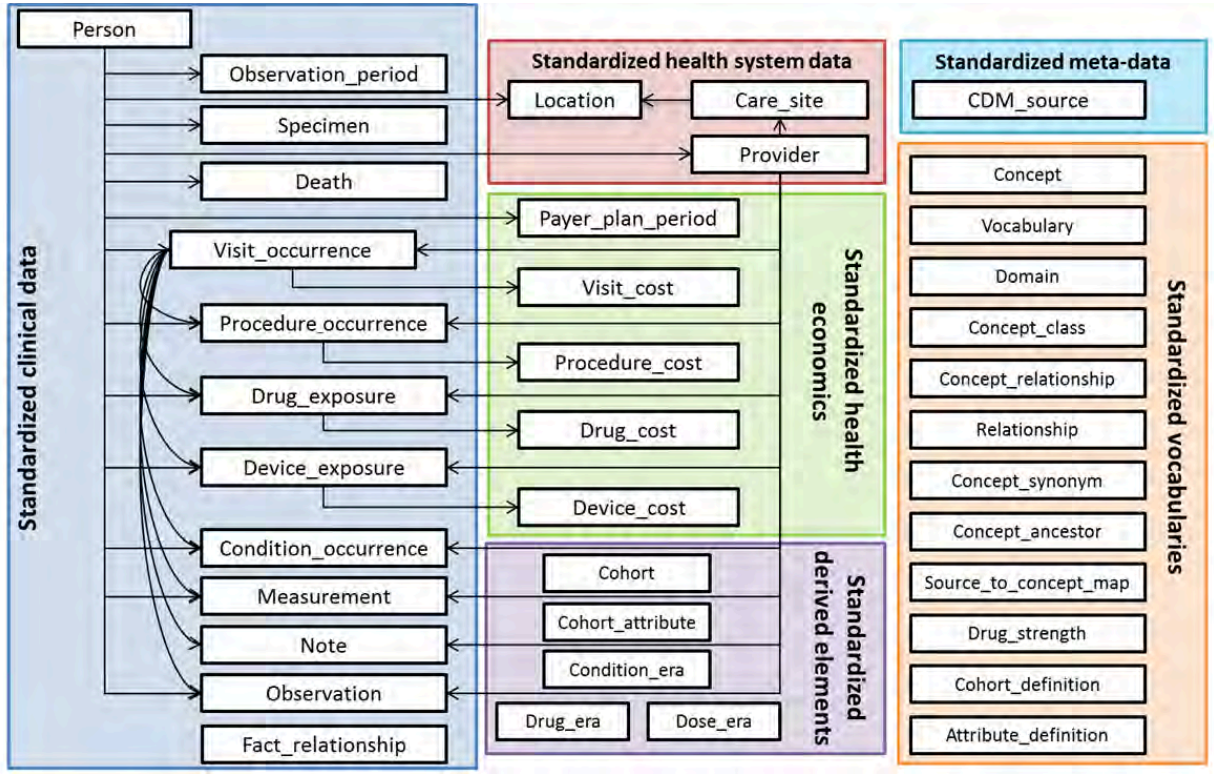
- Tableau – Data Visualization
  - Dashboards, Data Exploration, Trending
- Collibra – Data Governance
  - Data Definitions, Terminology Mapping
  - Value Set Management with Quality Measures
- OHDSI (Observational Health Data Science & Informatics)
  - OMOP Data Model
  - Standards-based model to harmonize data
  - Consortium of industry, government, and academia
- Microsoft SQL Server – Database
- Azure – Cloud-based user platform

Observational Medical Outcomes Partnership (OMOP)

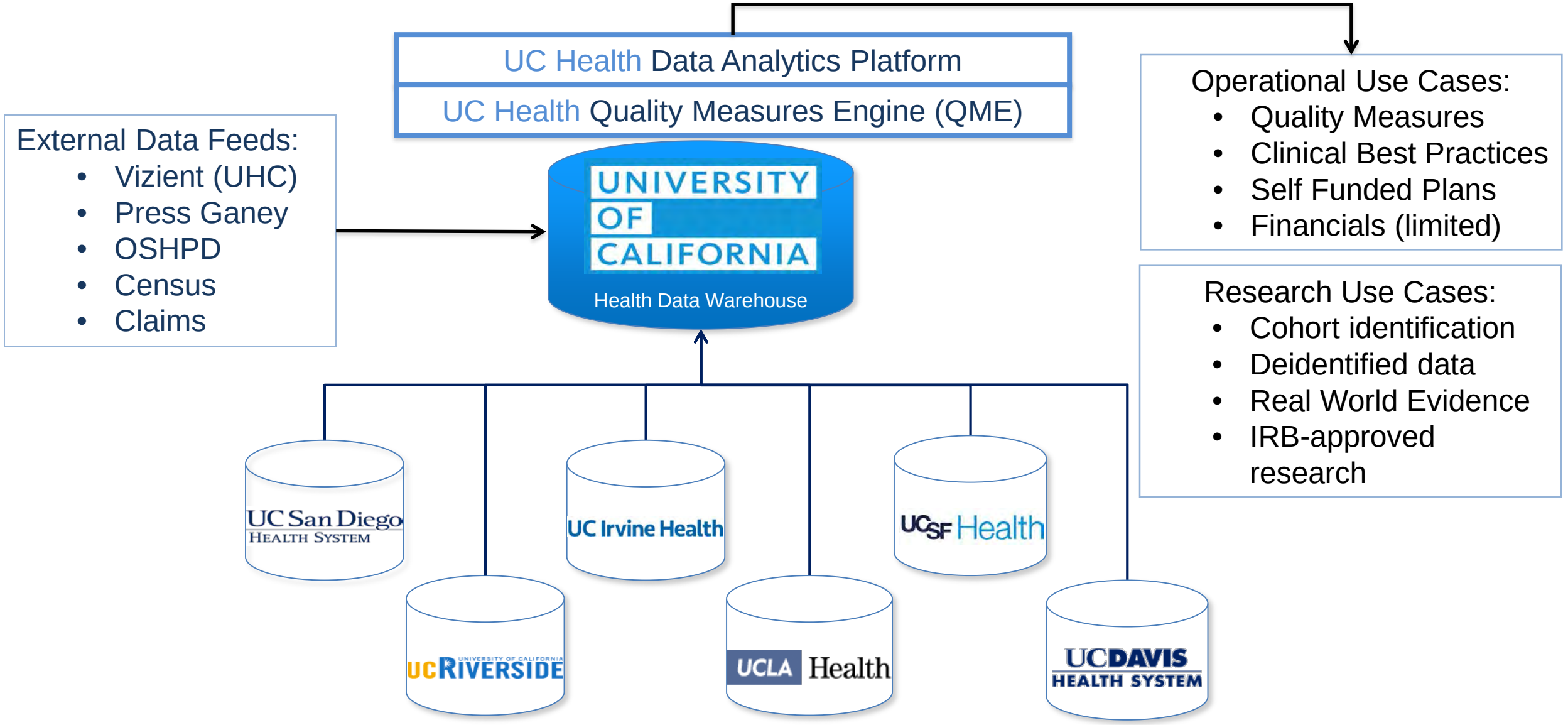


Why Use OMOP?

- Optimized for analytics rather than operations
- Data harmonized to national standards
  - LOINC, RxNORM, SNOMED-CT, ICD 10, CPT
- Scalable, extensible, backwards compatible data model
- Supports easy code sharing between sites
- Tool Set and Standard Queries available from OHDSI
  - Achilles: Data Quality Assessment Tool
- No further reliance on Epic
- The same OMOP-driven system is used for operational and research uses

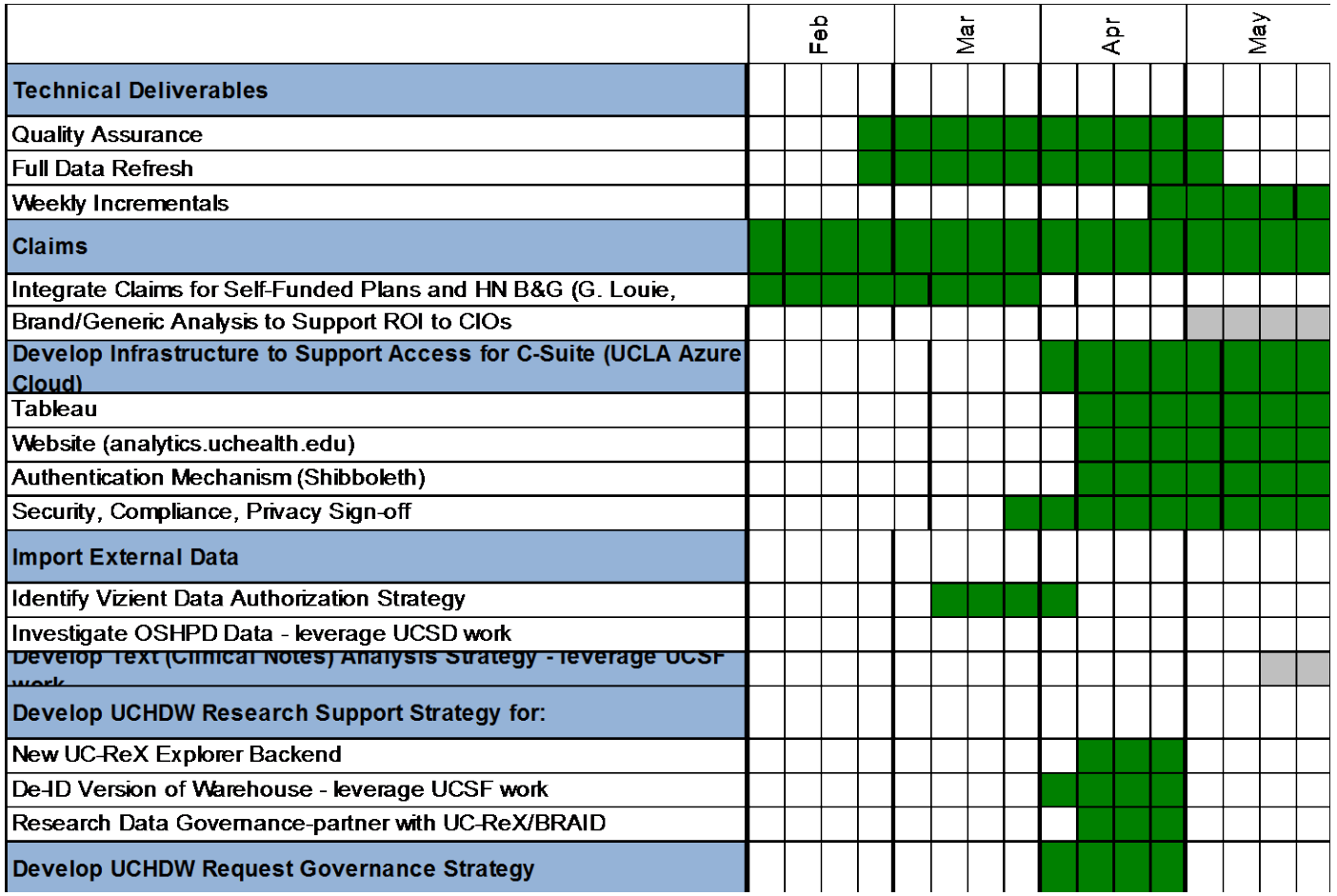


UCHDW Data Flow



Implementation Timeline

- Mar-Dec 2017
  - Infrastructure developed & implemented
- Since February 2018:
  - Integrated claims from self-funded & HealthNet B&G plans
  - Brand/Generic medication analytics
  - UCLA-based compliance and C-suite dashboard development
  - Data harmonization, first full data refresh in 6 months
  - Implemented monthly data refresh
  - Moved claims ingest to PCSSC from SDSC
  - Starting text results processing, i.e. radiology, pathology



UC has an unprecedented view of the medical system

- Combined data from UCSF, UCLA, UC Irvine, UC Davis, UC San Diego, and UC Riverside
- Central database built using OMOP (not Epic) as a data backend
  - Structured data from 2012 to the present day: 4.7 million patients with "modern" data, **total 15 million with an MRN**
  - 111M encounters, 310M procedures, 365M med orders, 676M vital signs, 433M lab test results, 443M diagnosis codes
  - Claims data from our self-funded plans now included
  - Continually harmonizing elements
- Quality and performance dashboards

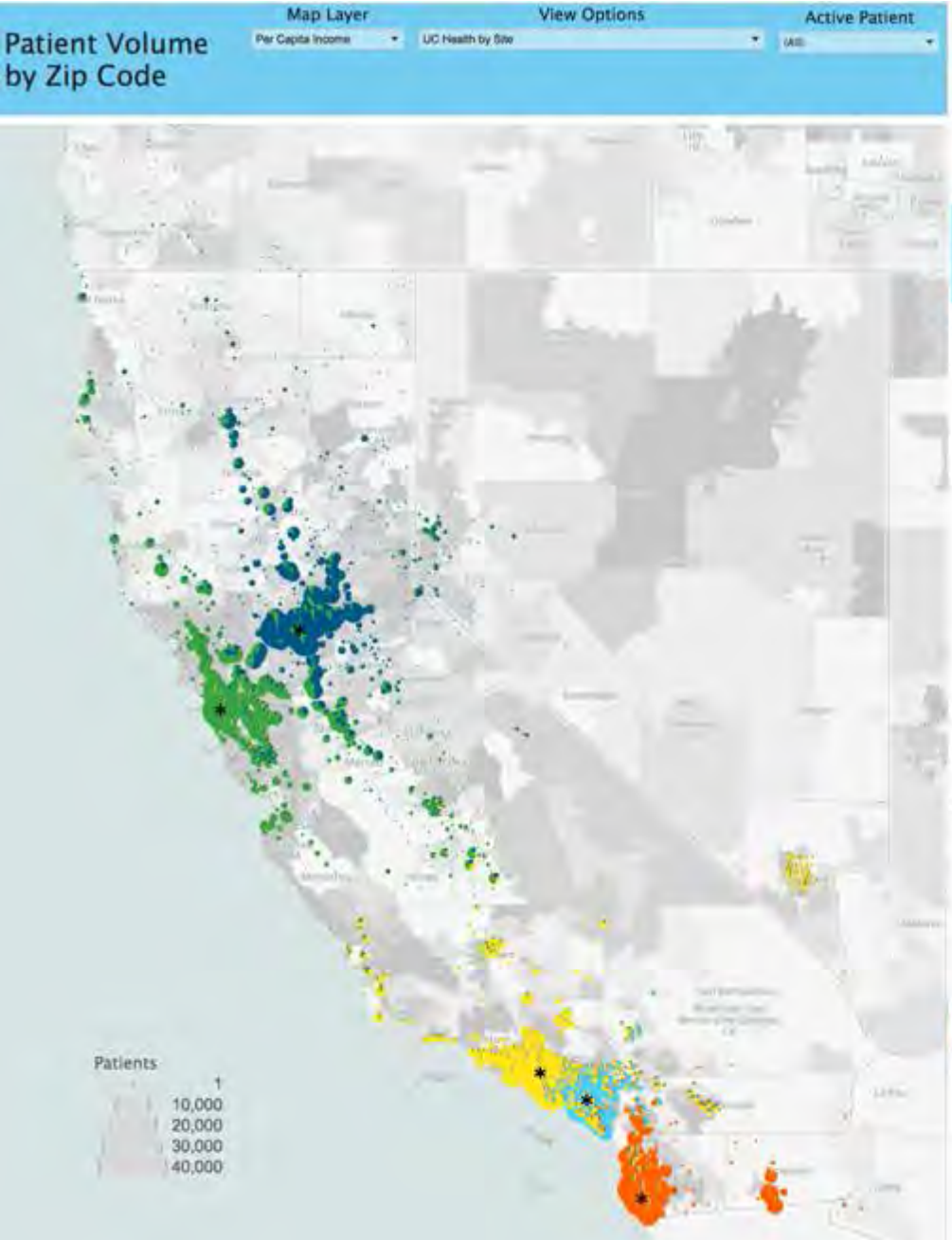
UC HEALTH

Data Overview

Batch ID
32

	UCD	UCI	UCLA	UCR	UCSD	UCSF	Grand Total
Condition Count	18,697,125 (4.2%)	46,089,557 (10.4%)	114,385,872 (25.8%)	187,770 (.0%)	94,857,444 (21.4%)	168,720,099 (38.1%)	**442,937,867** (100.0%)
Patient Count	820,922 (15.2%)	649,119 (12.0%)	1,700,128 (31.5%)	408 (.0%)	1,146,381 (21.2%)	1,085,434 (20.1%)	**5,402,392** (100.0%)
Visit Count	18,730,371 (16.9%)	6,028,245 (5.4%)	47,823,123 (43.2%)	39,569 (.0%)	17,878,602 (16.1%)	20,239,588 (18.3%)	**110,739,498** (100.0%)
Procedure Count	36,459,227 (11.7%)	61,816,741 (19.9%)	69,521,876 (22.4%)	159,189 (.1%)	93,027,882 (30.0%)	49,427,152 (15.9%)	**310,412,067** (100.0%)
Lab Count	73,918,920 (17.1%)	64,840,579 (15.0%)	118,729,747 (27.4%)	108,231 (.0%)	83,243,227 (19.2%)	92,516,644 (21.3%)	**433,357,348** (100.0%)
Vital Count	132,635,695 (19.6%)	54,820,106 (8.1%)	248,093,526 (36.7%)	171,000 (.0%)	110,944,950 (16.4%)	129,531,048 (19.2%)	**676,196,325** (100.0%)
Drug Exposure Count	85,203,771 (23.4%)	36,304,724 (9.9%)	84,103,493 (23.0%)	64,833 (.0%)	70,839,694 (19.4%)	88,369,808 (24.2%)	**364,886,323** (100.0%)

UC Health Patients (since January 2012)



Totals

4,370,110 Total UC Health Patients

963,838 (22.1%) | 672,289 (15.4%) | 1,492,010 (34.1%) | 522,691 (12.0%) | 719,282 (16.5%)

Most Recent Visit Date Range

January 2012 — May 2018

Age Group

18–35	96,373	89,540	150,368	65,482	84,560
65+	87,422	67,623	139,959	50,336	56,085
36–50	91,251	77,867	115,207	50,527	54,930
51–64	75,794	65,650	109,840	48,615	55,572
2–17	92,893	57,515			59,544
<2					

Race

White	175,606	152,074	248,492	157,073	
Unknown	139,819	186,240			
Other Race					
Asian					
No matching concept					
Black or African American					
MultiRace					
Native Hawaiian or Other Pacific Islan..					
American Indian or Alaska Native					

Gender

Female	526,682	367,372	810,460	293,651	378,900	2,377,065
Male	437,156	304,917	681,476	229,040	340,382	1,992,971
Other	74	74				

Legends/Filters

Locations: UCD, UCI, UCLA

Gender: (Multiple values)

Age Group: (All)

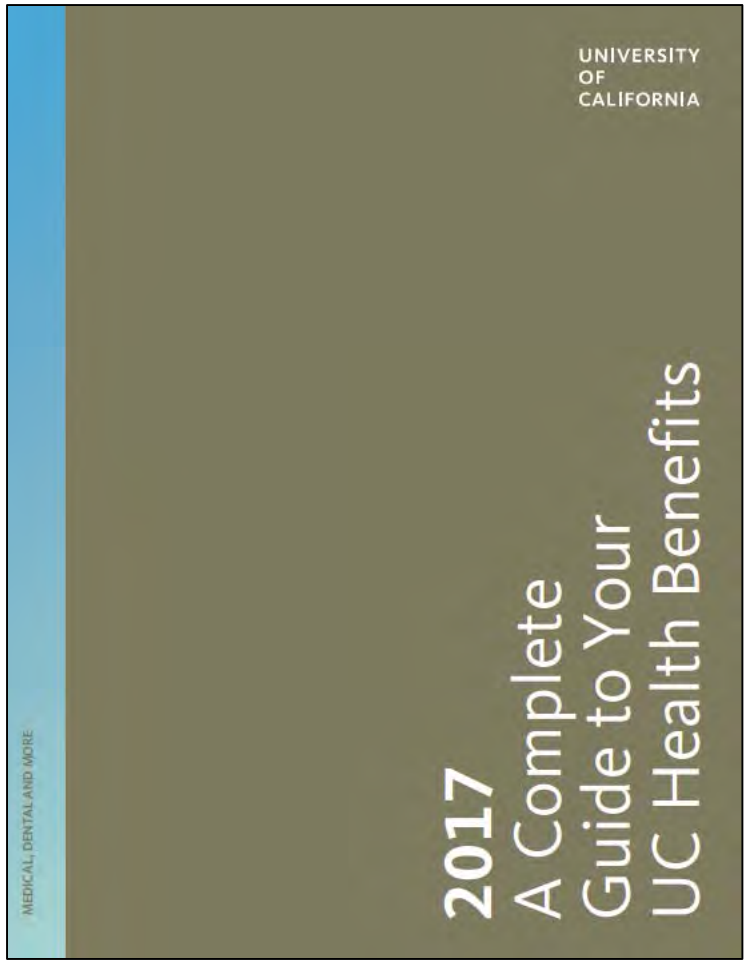
Ethnicity: (All)

2017 Population: 0 to 771; 771 to 2,170; 2,170 to 6,250; 6,250 to 20,100; 20,100 to 124,000

2017 Per Capita Income: 0 to 27,000; 27,000 to 31,900; 31,900 to 36,800; 36,800 to 44,800; 44,800 to 194,000

Initial Collaborative Groups

- **Self Funded Plans and Blue & Gold**
- **Population Health and ACO leaders**
- **UC Cancer Consortium**
- CEOs and strategy
- Pharmacy Chiefs
- UC Employee Diabetes Prevention Program
- CQOs and Quality Consortium
- Legal and Risk Management
- Leveraging Scale for Value: Supply Chain
- **Clinical Research and BRAID**



Medical Plans

HEALTH MAINTENANCE ORGANIZATIONS (HMO)

HMOs require you to choose a primary care physician (PCP) from their network of providers to coordinate your care. To see a specialist, you must have a referral from your PCP. The HMO covers your expenses only if your PCP has authorized the services, unless it's an emergency. You pay a copayment for some products and services, and there is no annual deductible.

You must live (or work, depending on the plan's rules) in the plan's service area to be eligible. Service areas are established by ZIP codes; you cannot use a P.O. box to establish eligibility. If you want to know whether your ZIP code is in a plan's service area, check the plan's website or call the plan directly. You can also use the Medical Plan Chooser (uc.chooser.pbgh.org) to see if your ZIP code is in a plan's service area.

UC's HMOs are available to employees living and working in certain counties in California only.

UC'S HMO

Health Net Blue & Gold HMO	Offers a tailored network of medical groups, doctors and hospitals, and includes all of UC's medical centers and medical groups. For more information, see healthnet.com/uc
Kaiser Permanente—CA	Offers a closed network, meaning you must use only Kaiser doctors and hospitals. For more information, see kp.org/universityofcalifornia
Western Health Advantage	Provides a regional network of medical groups, doctors and hospitals in the following areas: Davis/Sacramento (including UC Davis Health System), Marin, Napa and Sonoma counties. For more information, see westernhealth.com/mywha/welcome-to-wha/university-of-california

PREFERRED PROVIDER ORGANIZATIONS (PPO)

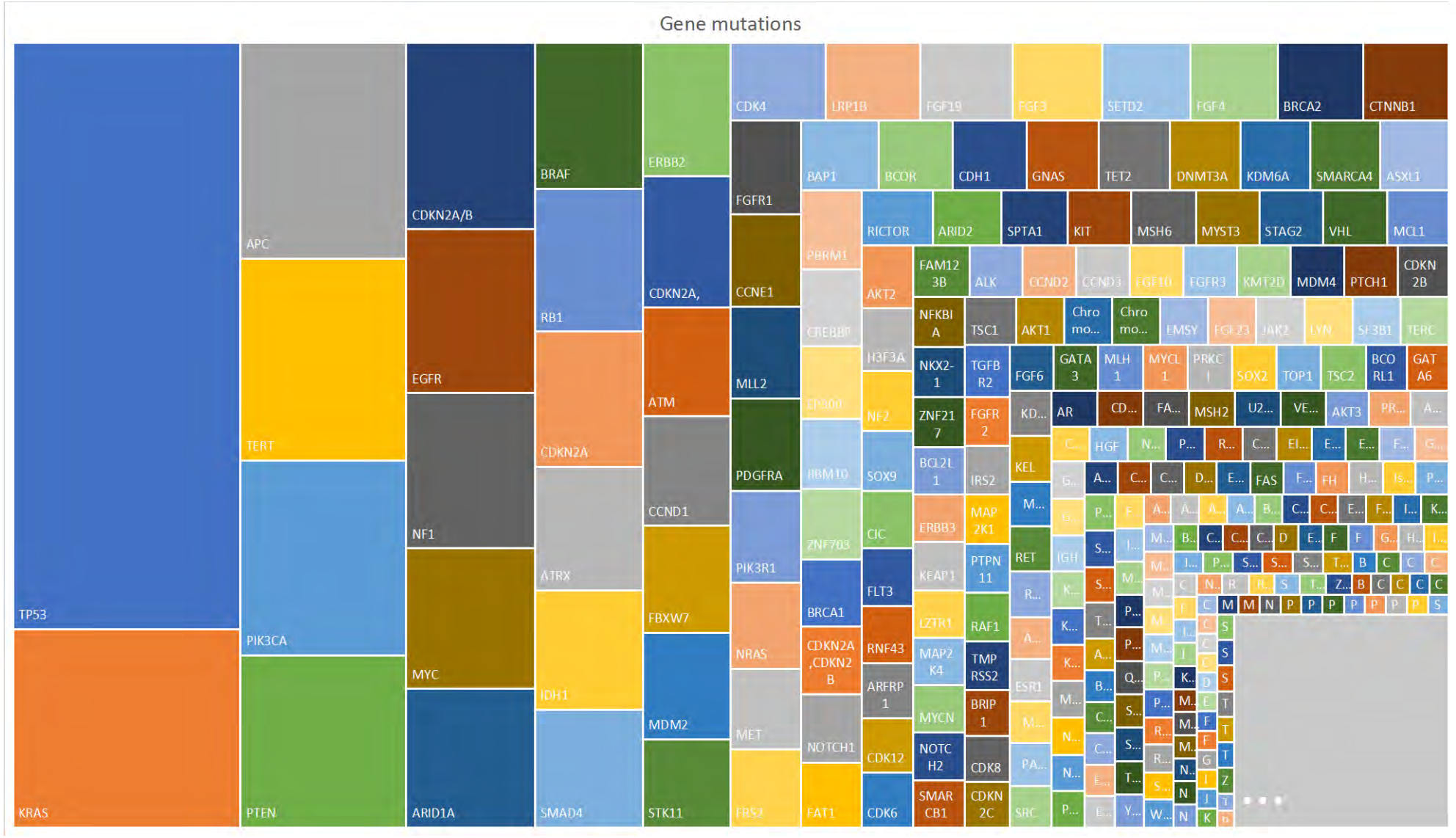
PPOs offer a broad network of providers and allow you the flexibility to see non-network providers if you wish. You don't need a referral to see your primary care doctor or specialists. Usually, you must meet the plan's deductible and then you pay coinsurance, which is a percentage of the cost of services. You pay a smaller percentage for in-network providers.

Anthem Blue Cross is the administrator of medical benefits for UC's PPO plans and OptumRx is the administrator of prescription drug benefits. The administrator of your plan processes claims, creates a network of health care providers or pharmacies and sets clinical policies and guidelines.

UC'S PPO

UC Health Savings Plan	This is a high-deductible PPO plan with a Health Savings Account (HSA), which you can use to pay your eligible medical expenses. UC contributes to the HSA and you can, too. You pay the cost of medical services until you meet the deductible, then you pay a percentage of the cost of services, with lower costs when you use in-network providers. You own the HSA, so the money goes with you if you leave UC. You can continue to contribute to it as long as you are enrolled in a qualifying high deductible health plan. For more information, see anthem.com/ca/uc and healthequity.com/ed/uc
UC Care	This is a PPO plan with three tiers. If you use providers in the UC Select Network, which includes UC medical center doctors, hospitals and other facilities as well as select providers near other UC locations, you pay copayments for services. If you use other providers in the Anthem Preferred network, you pay 20 percent coinsurance once you've met the deductible. You pay a higher deductible and a greater percent of the coinsurance if you use a provider outside the network. For more information, see uc-care.org

All Mutations



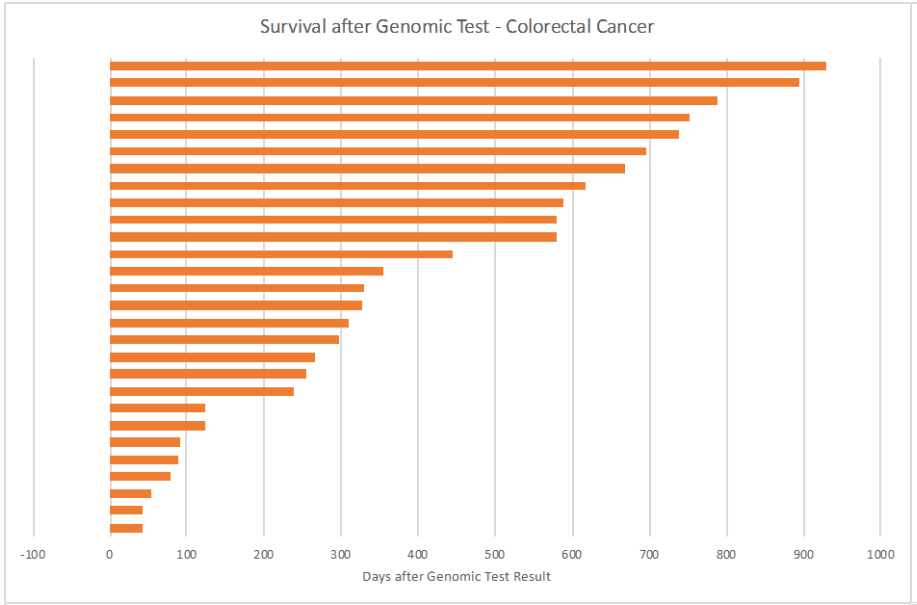
Mutations found in all 1851 patients with UCSF500 or Foundation 1 testing.

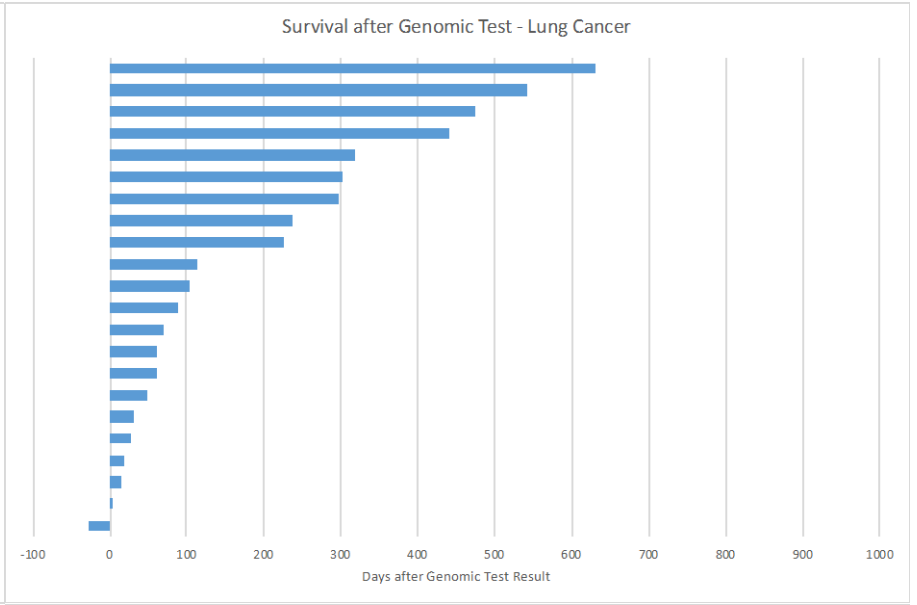
557 total genes

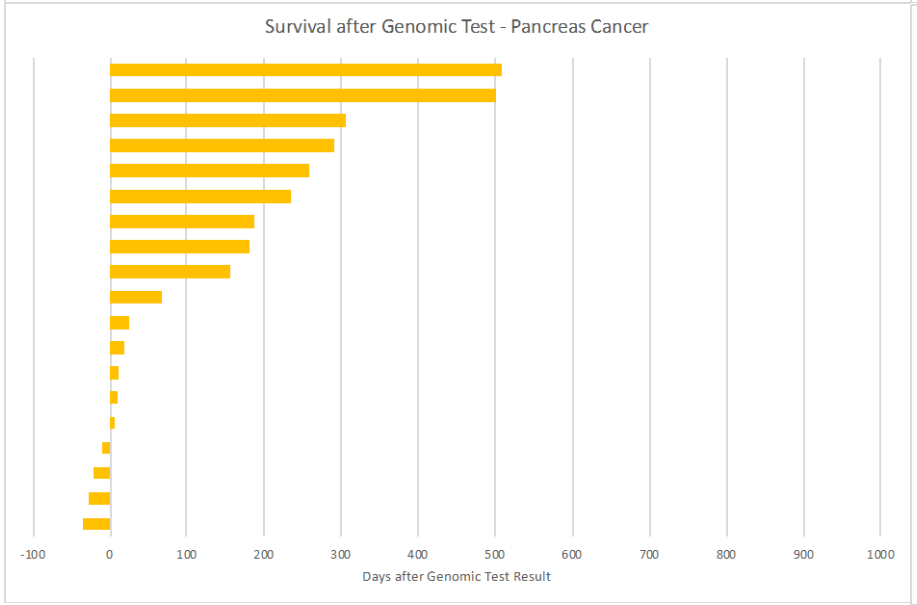
153 genes mutated in 10 or more samples.

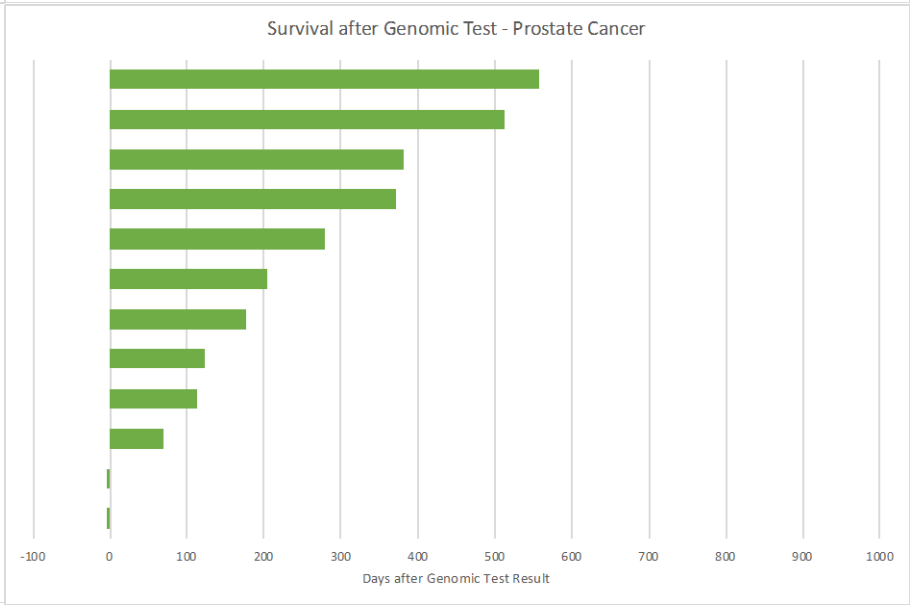
182 genes only mutated in one patient.

Survival after return of the Genomic Test Result









Colorectal – 28 patients

Lung – 22 patients

Pancreas – 19 patients

Prostate – 12 patients

Initial Collaborative Groups

- **Self Funded Plans and Blue & Gold**
- **Population Health and ACO leaders**
- **UC Cancer Consortium**
- CEOs and strategy
- Pharmacy Chiefs: inpatient drugs
- UC Employee Diabetes Prevention Program
- CQOs and Quality Consortium
- Legal and Risk Management
- Leveraging Scale for Value: Supply Chain
- **Clinical Research and BRAID**

Location of Next Appointment	Next Appointment When	Appointment Scheduled With	Measure Name
(All)	Today	(All)	(All)

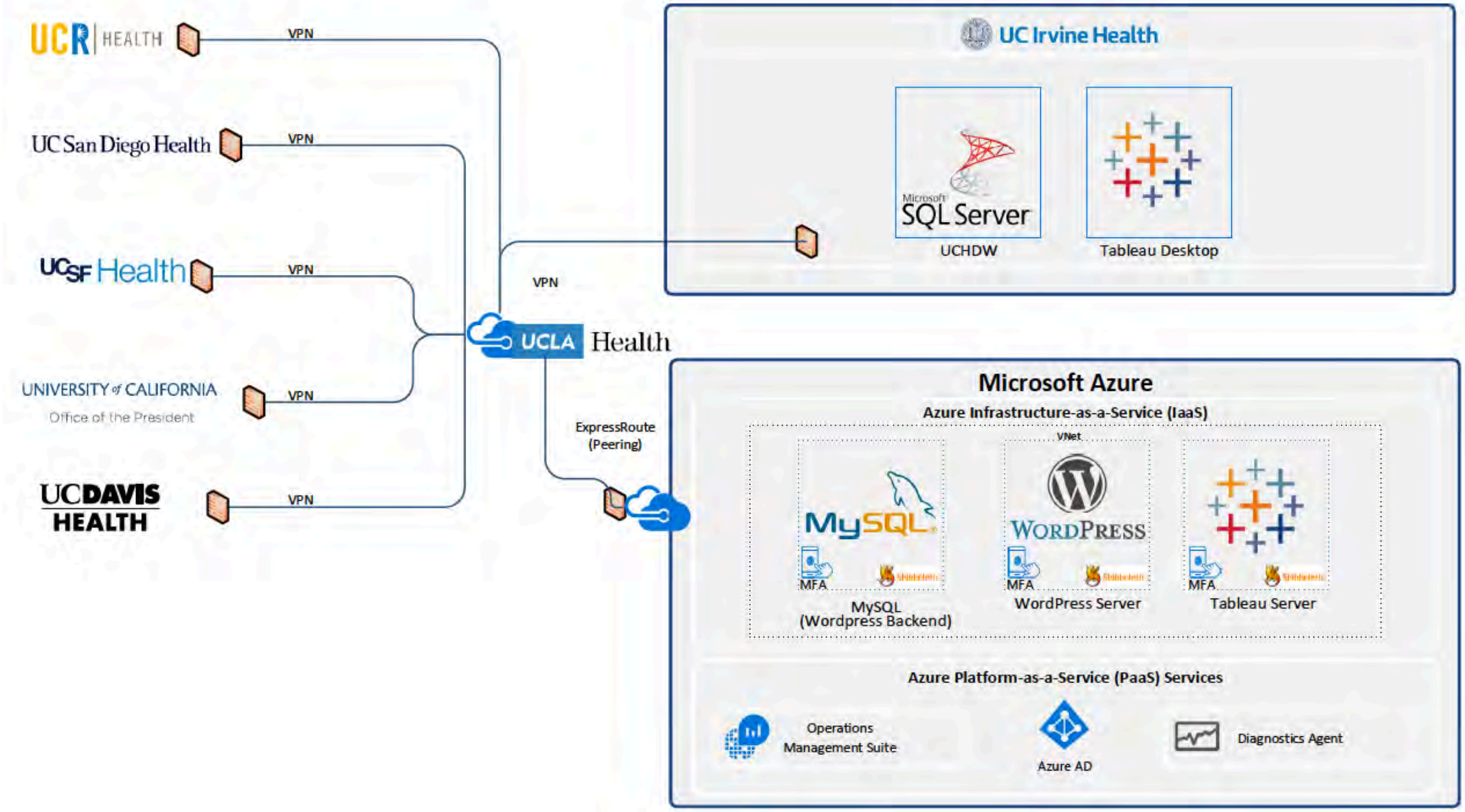
Improvement Patient List

Location of Next Scheduled Appointment	*Appointment Scheduled With*	*Date/Time of Next Scheduled Appointment*	*MRN*	*Patient Name*	***Number of Measures***
UCI BLOOD DONOR CENTER	APHERESIS	2018-06-	17	Hu	1
			18	M	1
UCI BREAST CTR DEXA	UCI PV 1 DEXA	2018-06-	19	Dc	4
UCI BREAST CTR MAMMO	UCI PV 1 MA1 DIAGNOSTIC	2018-06-	16	El	3
	UCI PV 1 MA TOMO	2018-06-	17	Lo	4
		2018-06-	23	He	3
		2018-06-	10	Ng	2
		2018-06-	25	Ve	2
		2018-06-	23	Sh	5
		2018-06-	13	Va	6
		2018-06-	23	Fo	2
UCI BREAST CTR US	UCI PV 1 US	2018-06-	25	Ca	6

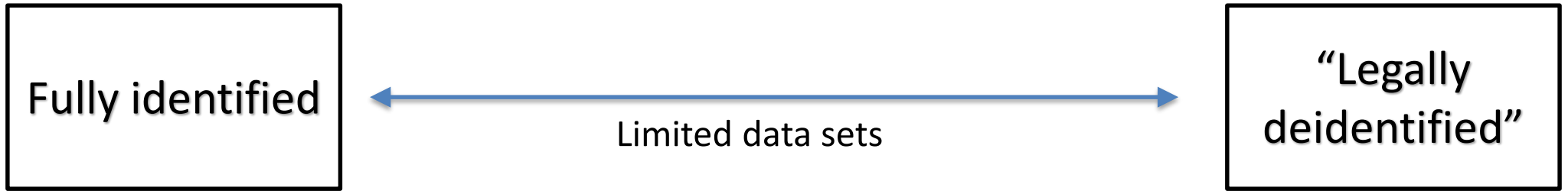
Initial Collaborative Groups

- **Self Funded Plans and Blue & Gold**
- **Population Health and ACO leaders**
- **UC Cancer Consortium**
- CEOs and strategy
- Pharmacy Chiefs: inpatient drugs
- UC Employee Diabetes Prevention Program
- CQOs and Quality Consortium
- Legal and Risk Management
- Leveraging Scale for Value: Supply Chain
- **Clinical Research and BRAID**

UCHDW: Access to Data



How will research access work?



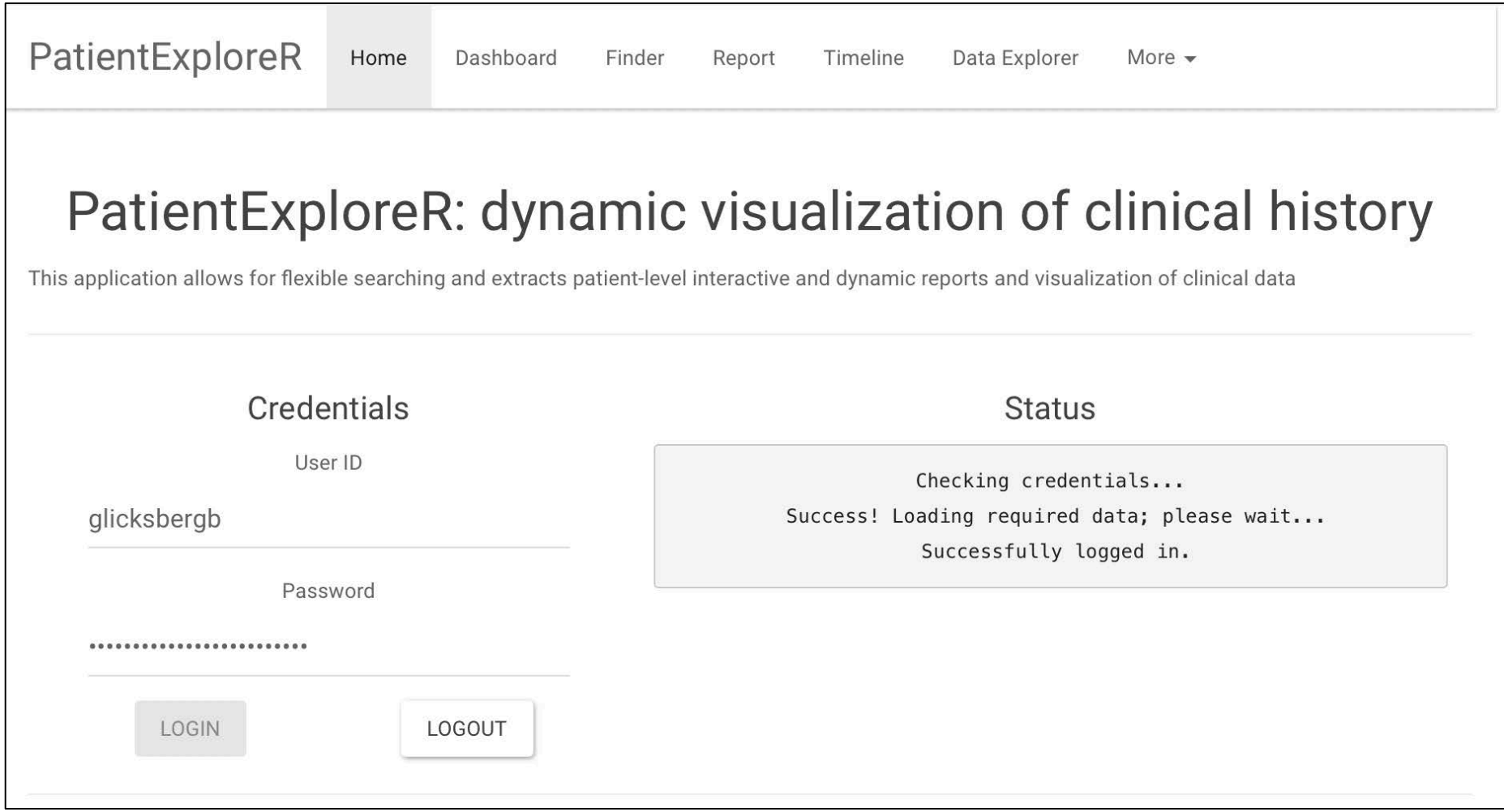
- Researchers should first write and optimize OMOP SQL queries locally, on their own campuses
- When ready to scale (and authorized), we spin up a virtual machine for the researcher, populated with common tools
  - Docusign a UC Health data use agreement
  - R, Tableau, Jupyter Notebooks, Julia, SQL, Windows or Linux
- Upload your scripts and run, but cannot download data

Campus OMOP Resources

- UCSF
- UC Davis
- UCLA
- UC Irvine
- UCSD
- Others

A new patient cohort browser

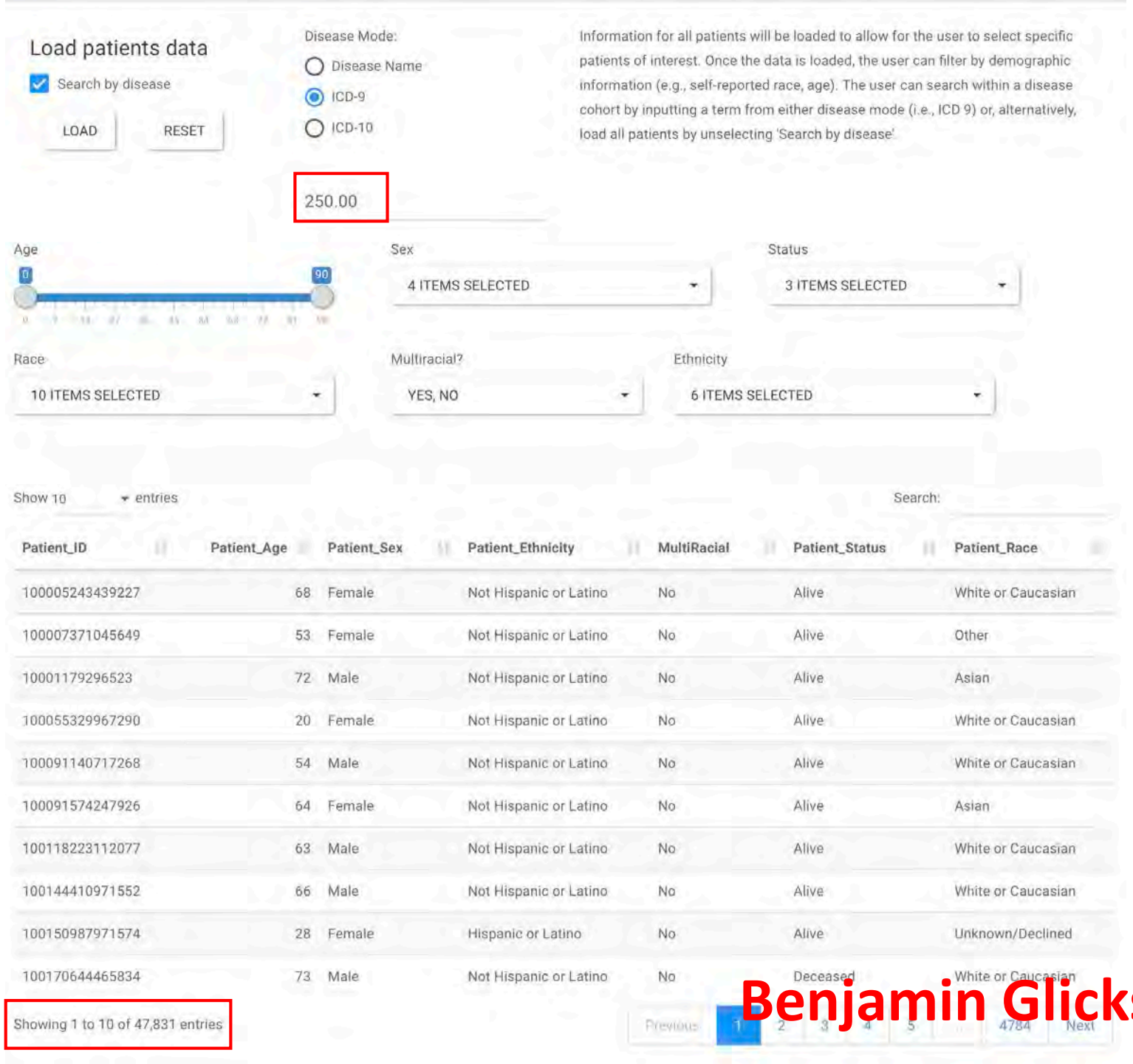
Search and build virtual patient cohorts



code at GitHub
https://github.com/BenGlicksberg/PatientExploreR

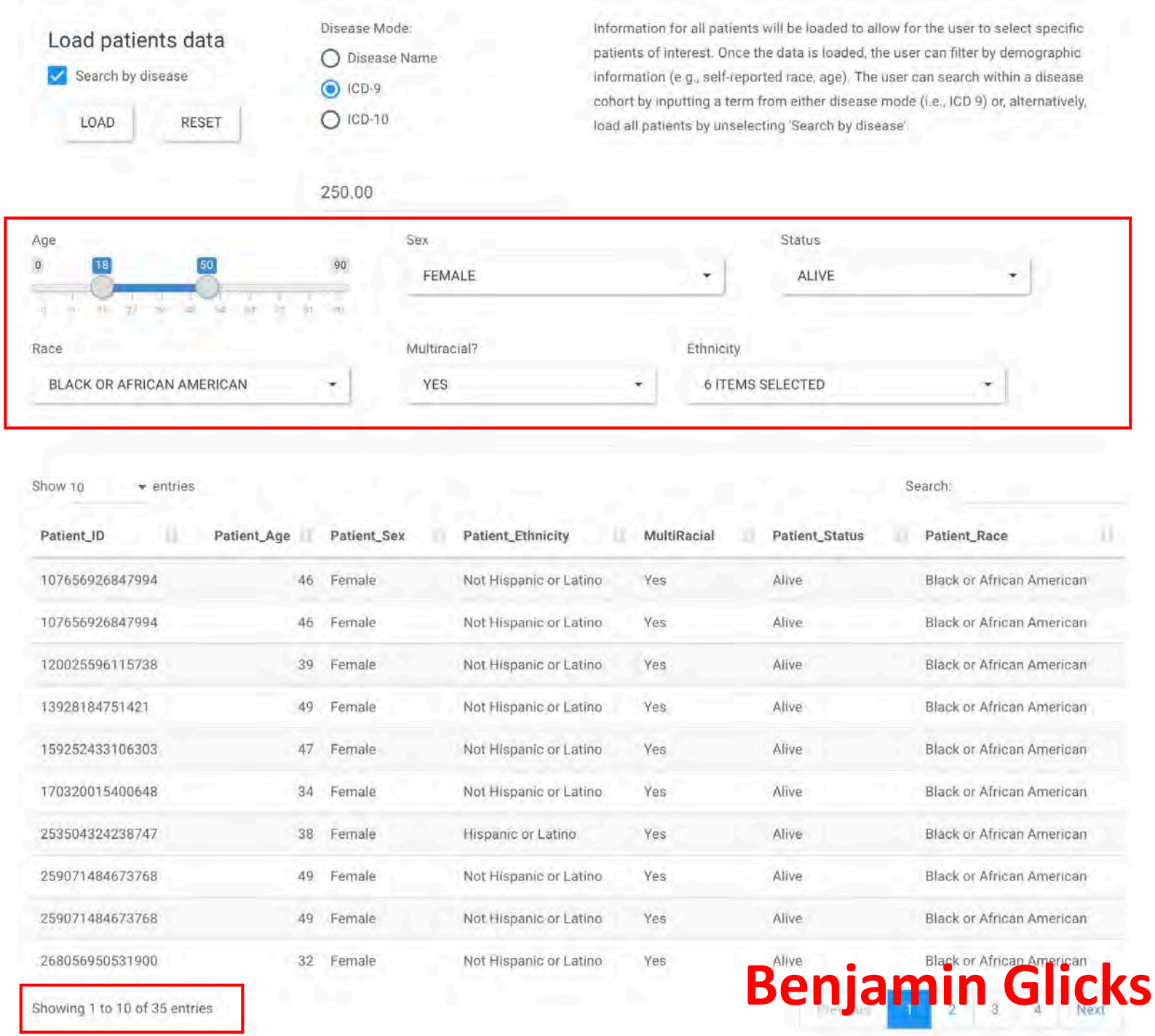
Benjamin Glicksberg

Search for patients



Benjamin Glicksberg

Filter patient list



Benjamin Glicksberg

Automatically generated clinical history

Show 25 entries Search:

Date	Type	Event	Value
2011-10-30	Flowsheet	WEIGHT/SCALE	3248 (ounces)
2011-10-30	Flowsheet	ZZZ R RT (MALE ADULT) PREDICTED BODY WEIGHT	66.1 (kg)
2011-10-30	Flowsheet	HEIGHT	67 (inches)
2011-10-30	Flowsheet	BLOOD PRESSURE	101/58 (NA)
2011-10-30	Flowsheet	R BSA	2.09 (sq meters)
2011-10-30	Flowsheet	ZZZ R RT (FEMALE ADULT) PREDICATED BODY WEIGHT	61.6 (kg)
2011-10-30	Flowsheet	R BSA DUBOIS_IP_CD_UCSF	2.03 (sq meters)
2011-11-18	Diseases	Malignant neoplasm of prostate	185 | C61
2011-11-21	Diseases	Malignant neoplasm of prostate	185 | C61
2011-11-21	Labs	Alanine transaminase	20 (U/L)
2011-11-21	Labs	Aspartate transaminase	23 (U/L)
2011-11-21	Labs	Alkaline Phosphatase	54 (U/L)
2011-11-21	Procedures	ALANINE TRANSAMINASE	NA
2011-11-21	Procedures	ALKALINE PHOSPHATASE	NA
2011-11-21	Procedures	ASPARTATE TRANSAMINASE	NA
2011-12-04	Diseases	Mild nonproliferative diabetic retinopathy(362.04)	362.04 | E11.3299
2011-12-04	Diseases	Unspecified glaucoma(365.9)	365.9 | H40.9
2011-12-04	Diseases	Open angle with borderline findings, low risk	365.01 | IMO0002
2011-12-04	Diseases	Open angle with borderline findings, low risk	365.01 | H40.019
2011-12-04	Diseases	Senile nuclear sclerosis	366.16 | H25.10
2011-12-04	Procedures	COLOR STEREO PHOTOGRAPHY	NA
2011-12-04	Flowsheet	PAIN SCORE (CATEGORY LIST)	0 (NA)
2011-12-04	Flowsheet	PAIN LOCATION (CATEGORY LIST)	Eye (NA)
2011-12-05	Diseases	Other specified pre-operative examination	V72.83 | Z01.818
2011-12-05	Diseases	Malignant neoplasm of prostate	185 | C61

Showing 101 to 125 of 7,443 entries

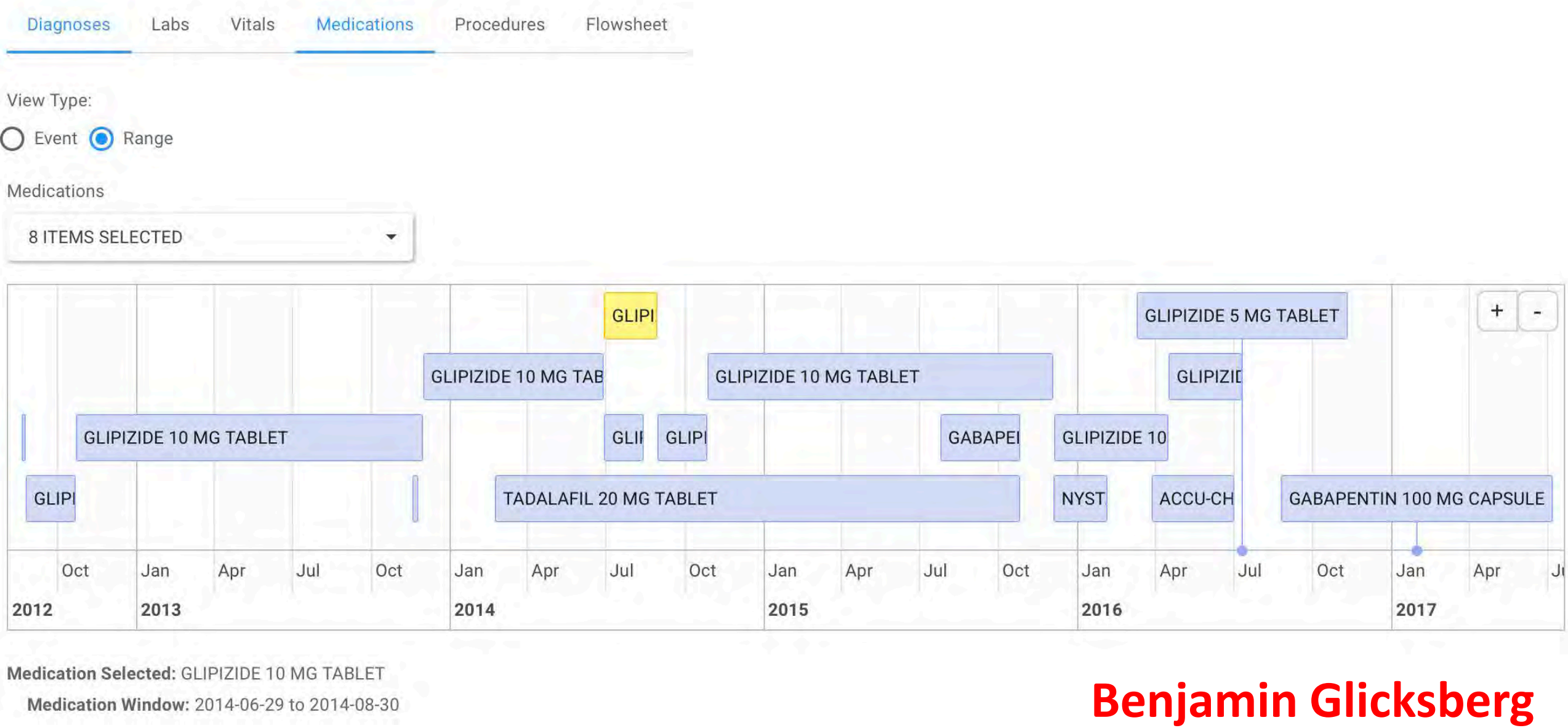
Previous 1 . . 4 5 6 . . 298 Next

Benjamin Glicksberg

Explore Trends in Data/ Outcomes (targeted)

Benjamin Glicksberg

Explore Trends in Data/Outcomes (targeted)

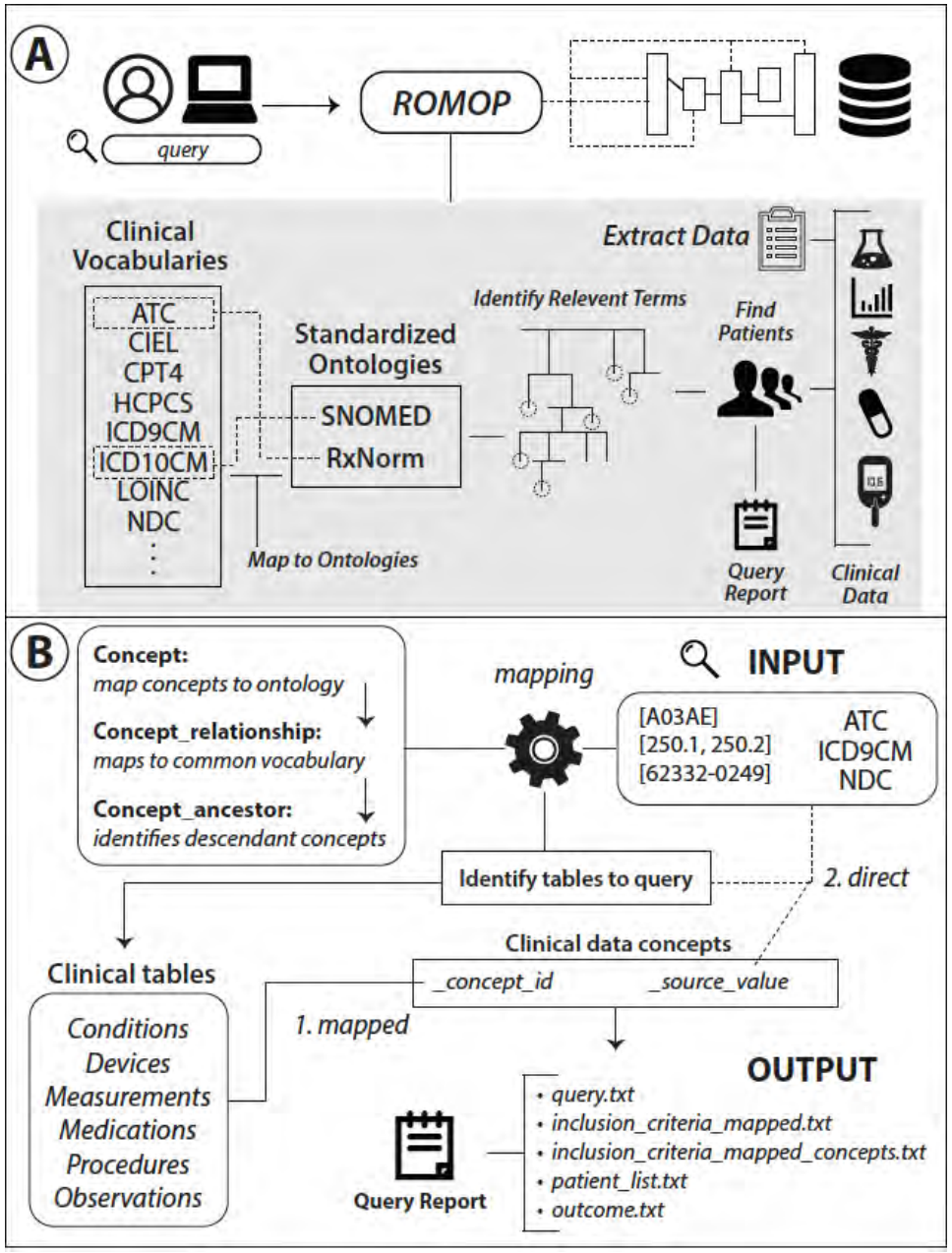


Benjamin Glicksberg

Explore Trends in Data/ Outcomes (numeric; targeted)

Benjamin Glicksberg

ROMOP

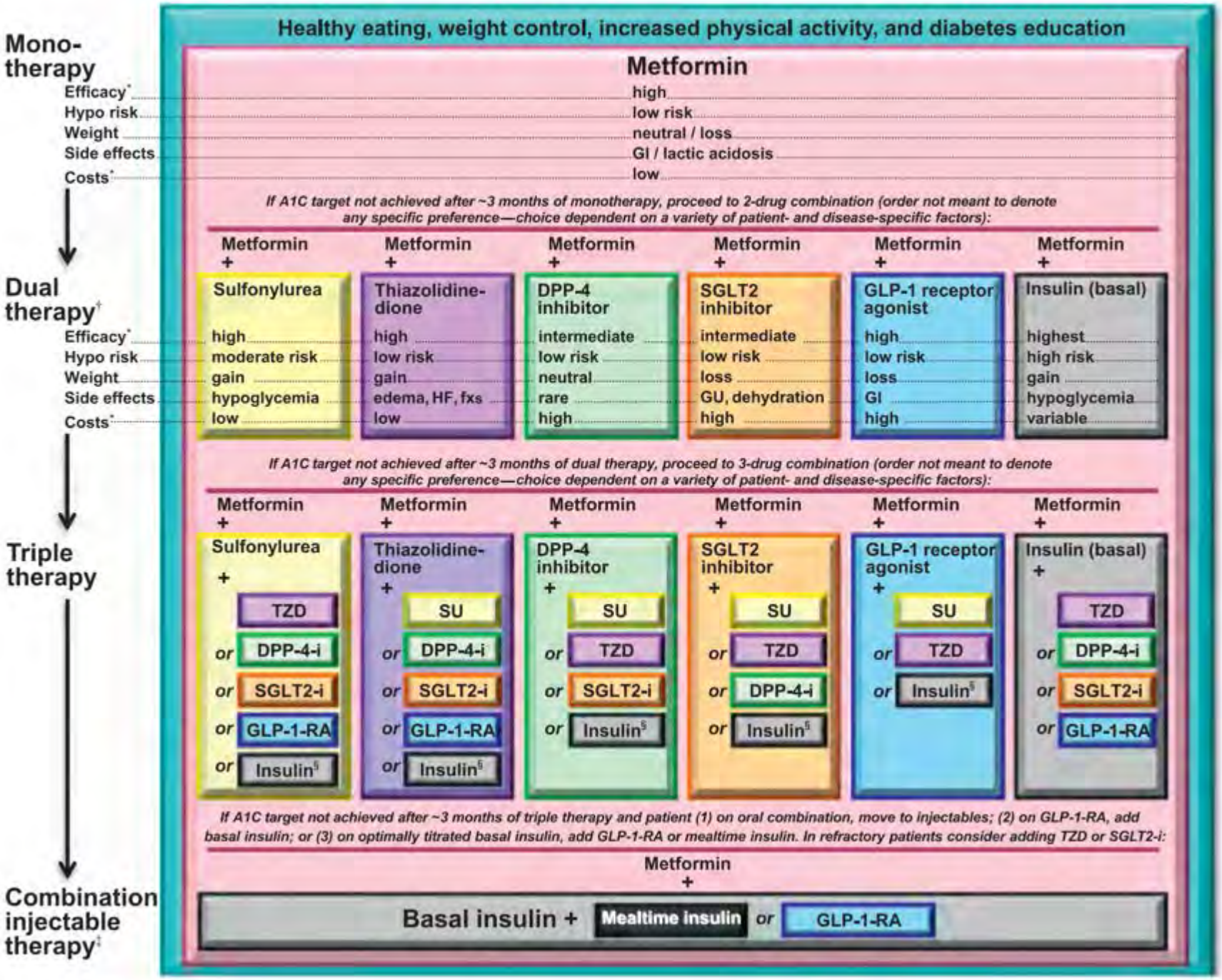


code at GitHub
https://github.com/BenGlicksberg/ROMOP

Benjamin Glicksberg

What's Next?

- Continue work with partners
  - **BRAID**
  - Self-funded plans
  - Pharmacy Project
  - ACOs
  - Pre-Diabetes Project
  - Leverage Scale for Value: Supply Chain
  - Cancer
- Integrate with Vizient benchmark data
- Core Infrastructure
  - Complete Phase 1 Azure Implementation (Tableau, UCHDW Website)
  - Start Phase 2: Move UCHDW database to Azure
  - Finalize UCHDW data validation
  - **More data**: ~~Scheduling~~, ~~unit history,~~ radiology, pathology results, other notes, deidentified notes
- **Education**
- Data de-identification and access mechanisms
  - IRB has given us the go ahead
  - Finalize login mechanisms
  - Final checks with compliance, security, privacy, legal
- Reference Data Management - Standardized sources
  - Monthly Updates of drug terminology mappings
  - Yearly Updates of ICD, CPT, HCPCS
  - Quarterly Update of OMOP Vocabulary
  - Process to determine new reference data



Source: American Diabetes Association Standards of Medical Care in Diabetes

Medication Strategies for First-Time Type 2 Diabetes Patients

Monotherapies:
- Metformin (Met)
- Insulin
- Sulfonylurea / Glinide (Su/Gln)
- DPP-4 Inhibitor (DPP-4)

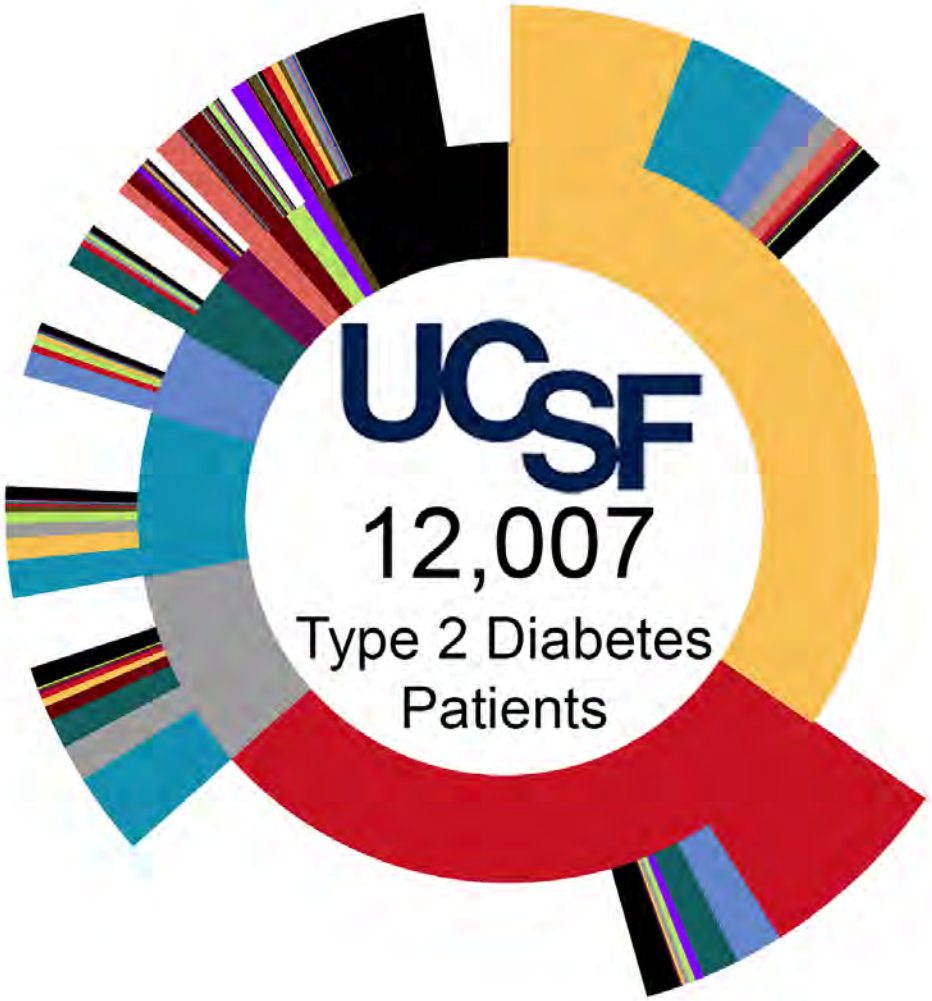
Dual therapies:
- Met & Su/Gln
- Met & DPP-4
- Met & Insulin
- Su/Gln & Insulin
- Su/Gln & DPP-4
- DPP-4 & Insulin

Triple therapies:
- Met, Su/Gln, & DPP-4
- Met, Su/Gln, & Insulin

- Other Less Frequent Combinations

Tom Peterson

Medication Strategies for First-Time Type 2 Diabetes Patients



Monotherapies:
- Metformin (Met)
- Insulin
- Sulfonylurea / Glinide (Su/Gln)
- DPP-4 Inhibitor (DPP-4)

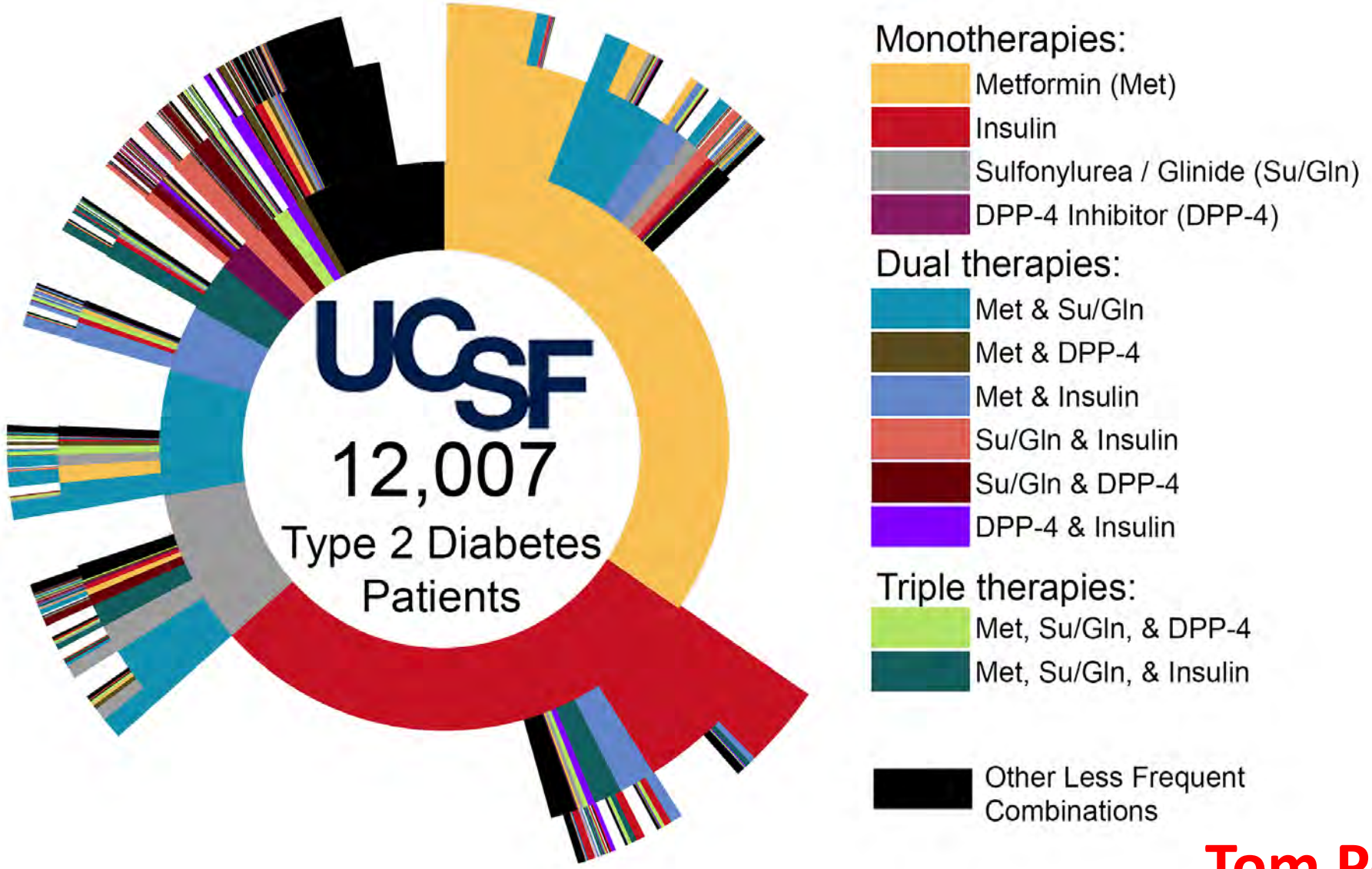
Dual therapies:
- Met & Su/Gln
- Met & DPP-4
- Met & Insulin
- Su/Gln & Insulin
- Su/Gln & DPP-4
- DPP-4 & Insulin

Triple therapies:
- Met, Su/Gln, & DPP-4
- Met, Su/Gln, & Insulin

- Other Less Frequent Combinations

Tom Peterson

Medication Strategies for First-Time Type 2 Diabetes Patients



Tom Peterson

1,640 Unique Medication Trajectories for Treating T2D at UCSF

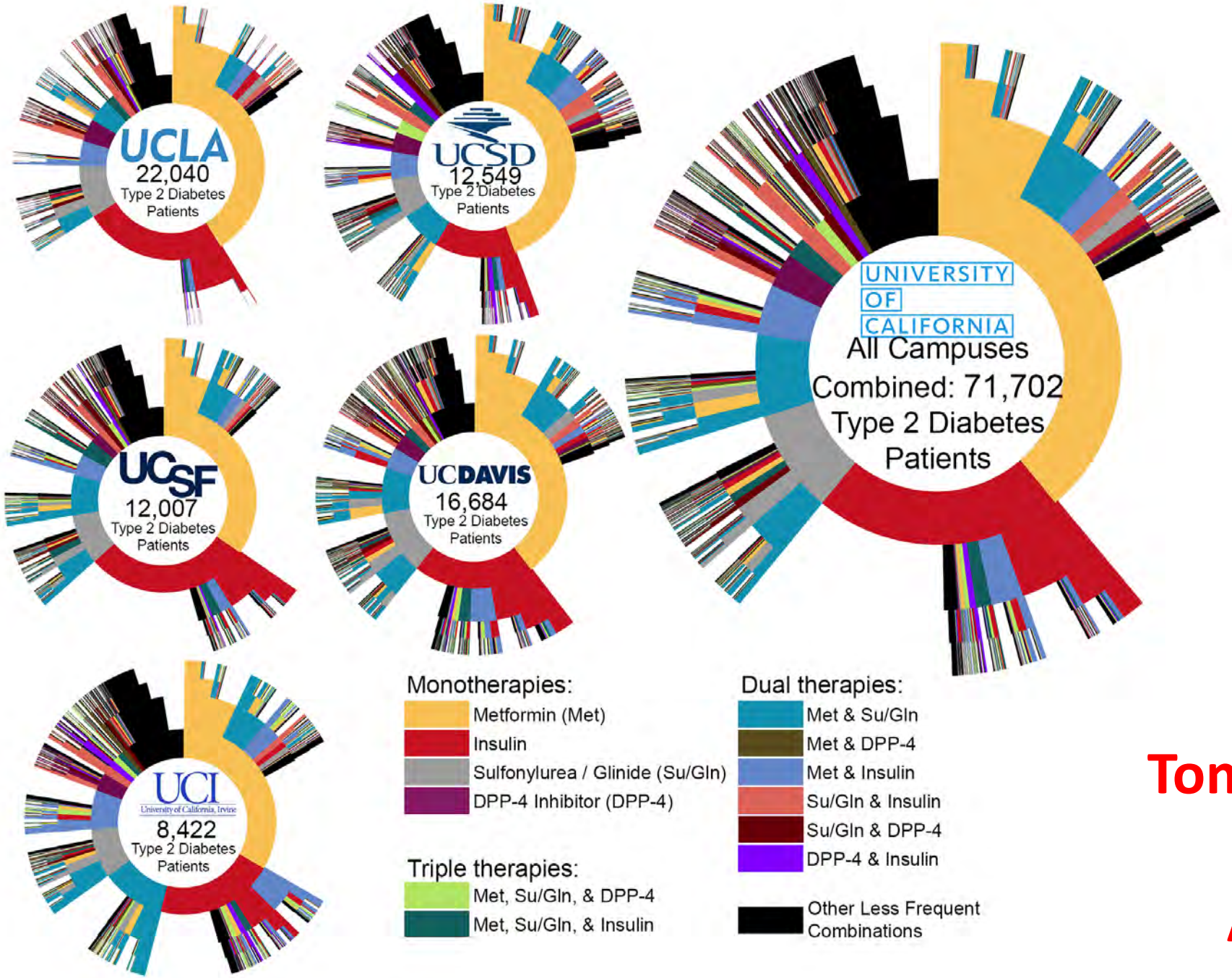
Monotherapies:
- Metformin (Met)
- Insulin
- Sulfonylurea / Glinide (Su/Gln)
- DPP-4 Inhibitor (DPP-4)

Dual therapies:
- Met & Su/Gln
- Met & DPP-4
- Met & Insulin
- Su/Gln & Insulin
- Su/Gln & DPP-4
- DPP-4 & Insulin

Triple therapies:
- Met, Su/Gln, & DPP-4
- Met, Su/Gln, & Insulin

- Other Less Frequent Combinations

Tom Peterson



Tom Peterson
Lisa Dahm
Ayan Patel

6,543 Unique Medication Trajectories for Treating T2D UC-Wide

Clinical Data Overview

- Only patients that have had an encounter after 1/1/2012 are included
- Primary data source is Epic's Caboodle (EDW)
  - Epic's Clarity is still needed for a subset of data
- Each site has a local implementation, which is transferred monthly to central database
- OMOP Common Data Model (CDM) v5.1
  - Plan to upgrade to v5.3 by end of CY18

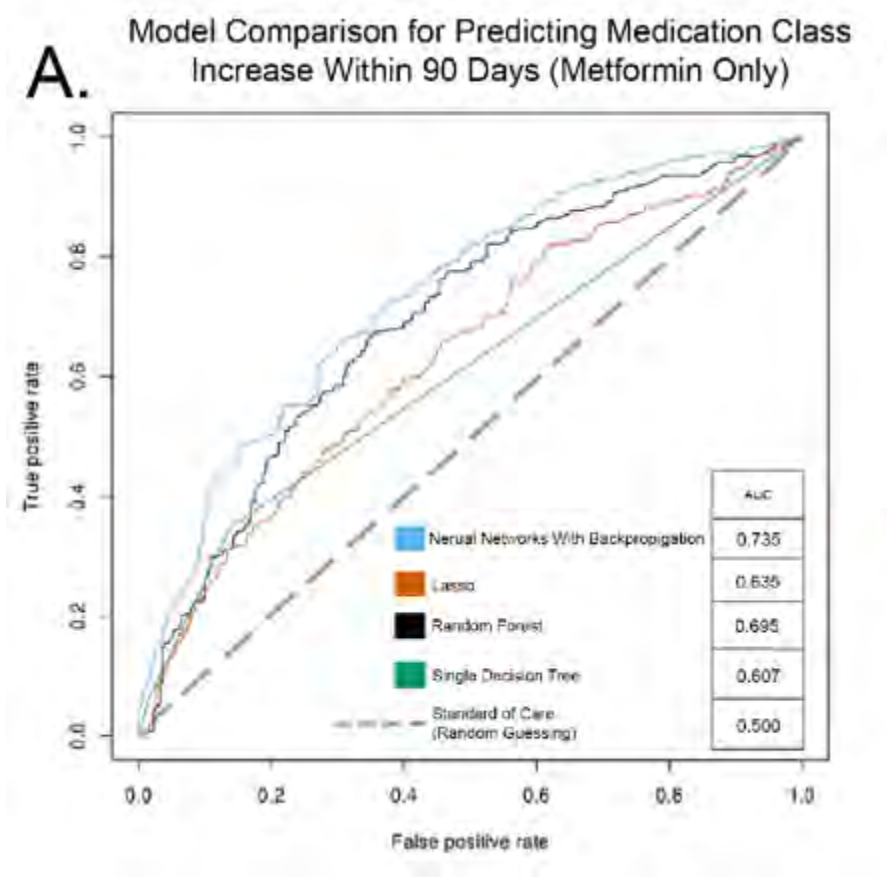
Data Mapping Details

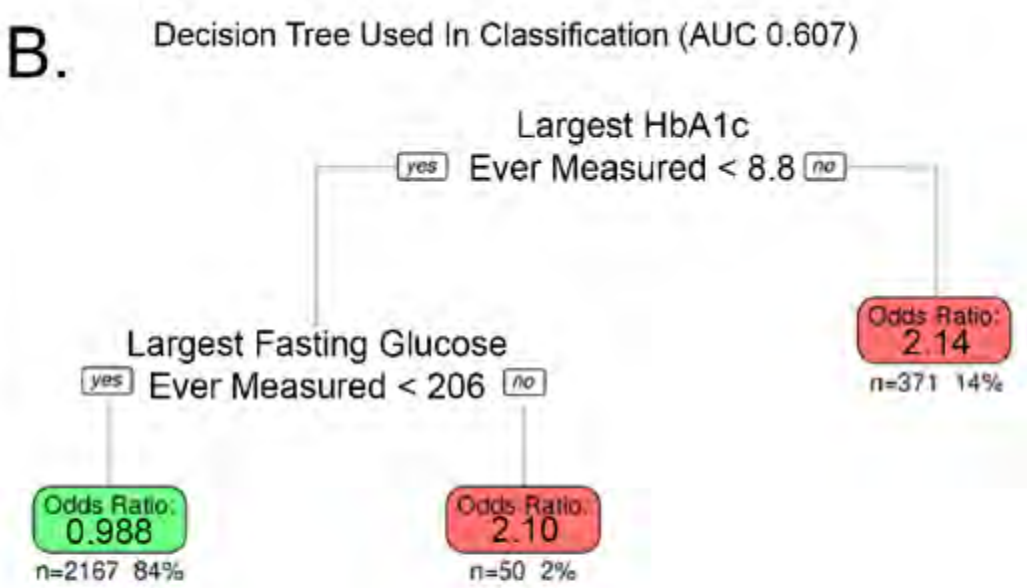
- **Conditions**: ICD9/10-CM
- **Procedures**: ICD9/10-PCS; CPTII/CPT4/HCPCS; Health Maintenance, Orders, and Referrals mapped to SNOMED
- **Drug exposures**: Medications mapped to RxNorm
  - Semantic Clinical Drug (SCD) term type (TTY)
- **Measurements**: Labs results and Vital signs from flowsheets mapped to LOINC
- **Observations**: Mapped to SNOMED, otherwise LOINC

Quality Measures Engine

- Reuse quality measure logic for multiple programs and different populations (externally defined or internally defined)
- Ability to choose value sets per measure per program, mix and match for programs that are flexible to maximize performance
- Allows us to handle multiple value-based quality measure programs with minimal resources

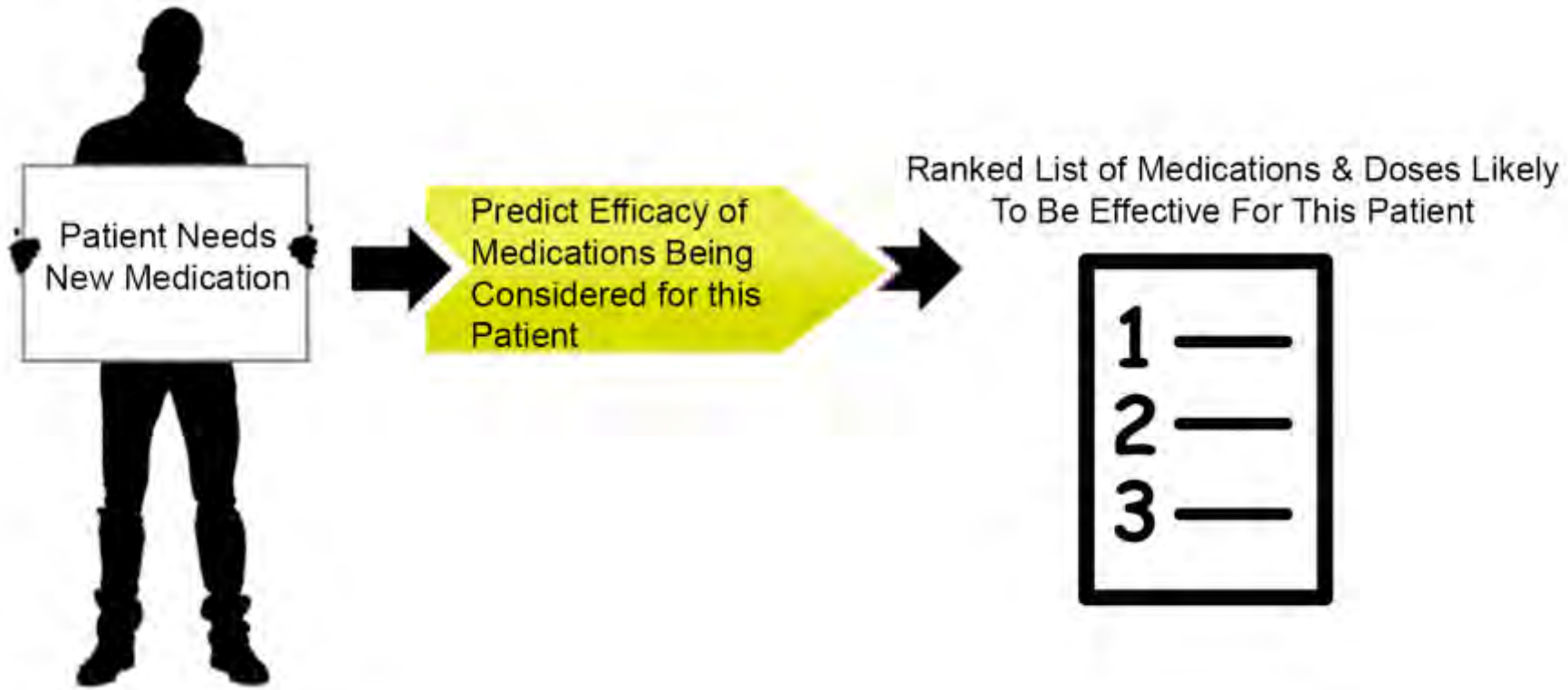
Performance for Predicting Medication Class Increase Within 90 Days from Metformin





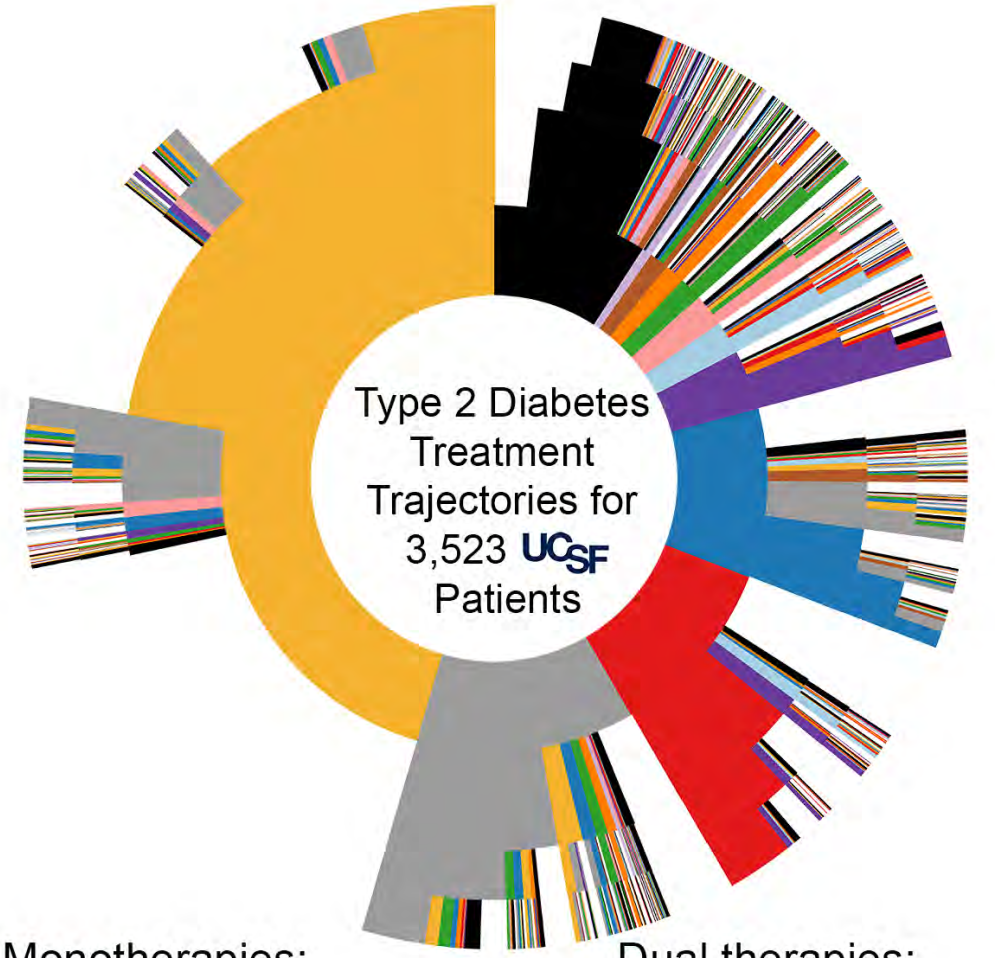
Tom Peterson

Precision Medicine In Practice



Tom Peterson

Tracking Patients Over Time



Monotherapies:
- Metformin (Met)
- Insulin
- Sulfonylurea / Glinide (Su/Gln)

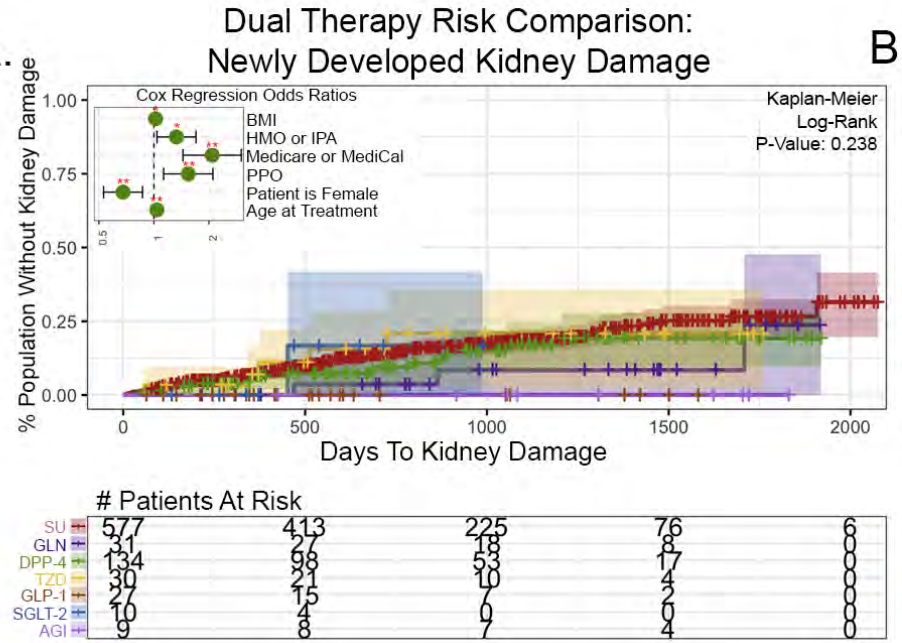
Triple therapies:
- Met, Su/Gln, & DPP-4
- Met, Su/Gln, & Insulin
- Met, Su/Gln, & Thiazoilidinedione

Dual therapies:
- Met & Su/Gln
- Met & DPP-4
- Met & Insulin
- Su/Gln & Insulin
- Su/Gln & DPP-4

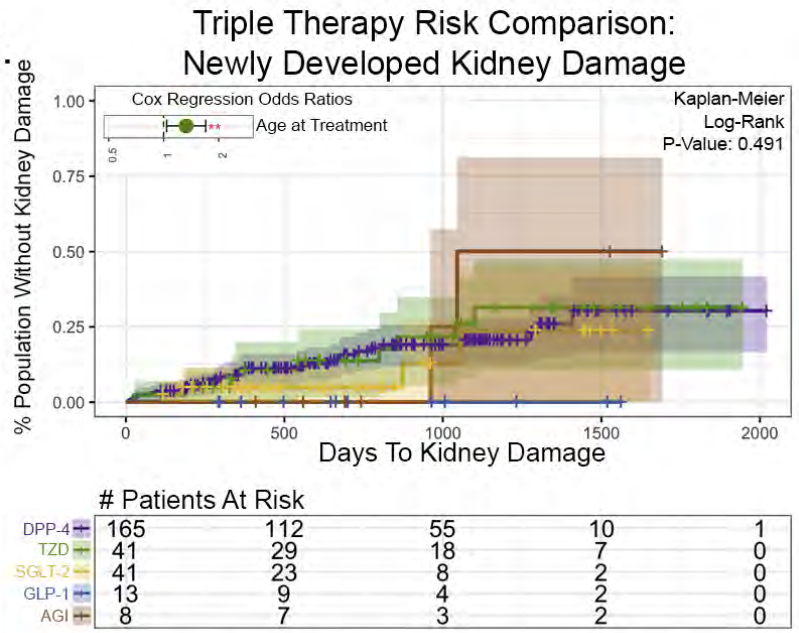
Other Less Frequent Combinations

Comparing Safety (Medication Classes)

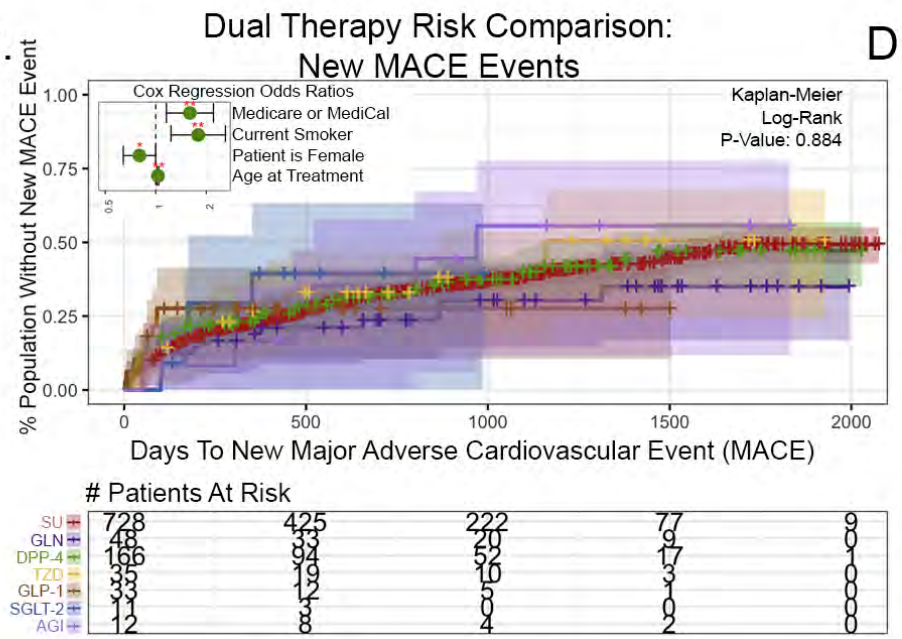
A.



B.



C.



D.