



University of California
San Francisco

De-identified Data: Research Data Browser and Flat Files

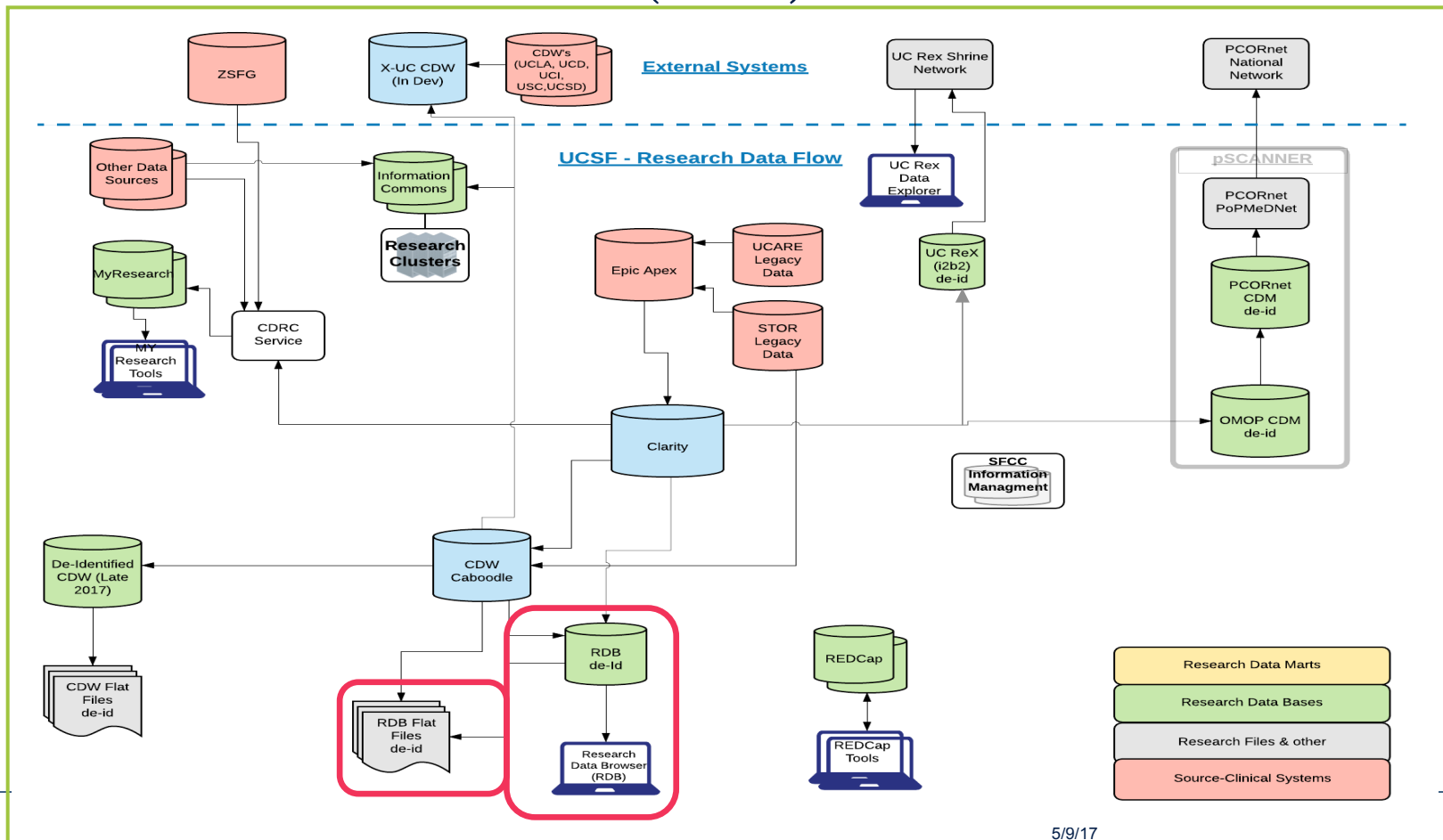
Dana Ludwig
Enterprise Information and Analytics

Research Data Browser (RDB)

Scenario	Self-serve (Free)	Consult Required (May have recharge)	IRB needed	Requires MyResearch account or other secure environment	Includes clinical notes	UC Health data available in addition to UCSF data
Counts	Yes	No	No	No	No	Yes
De-ided data	Yes	No	No	No	No	Yes *
Limited data	No	Yes	Yes	Yes	No	Yes *
Id'ed data	No	Yes	Yes	Yes	Yes	Yes *
Recruit	No	Yes	Yes	Yes	No	Yes *

* Process to obtain data from UC Health sites can be very lengthy

Research Data Browser (RDB)



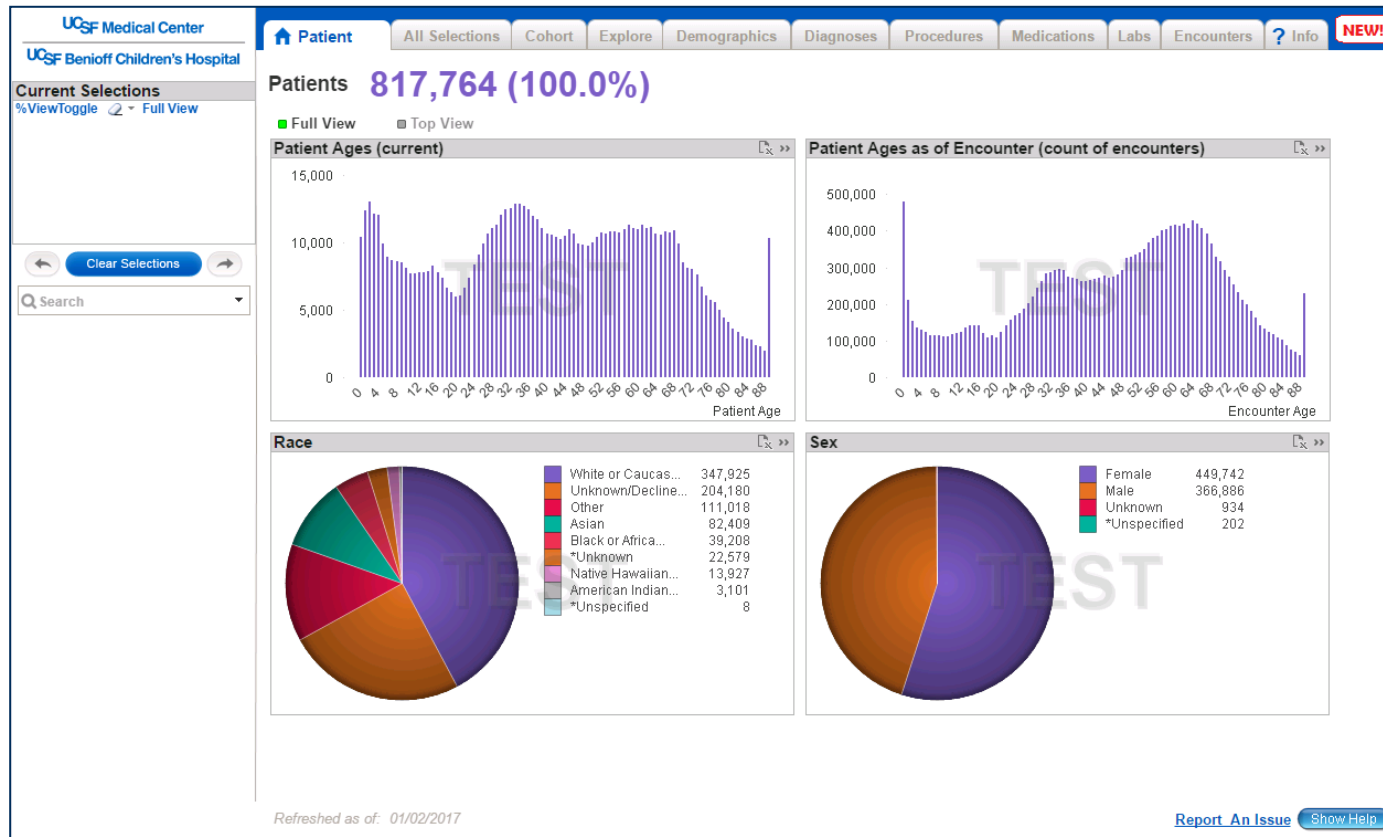
Research Data Browser

APeX data query and extraction

The goal is to accelerate the clinical scientific discovery process for UCSF investigators by enabling:

- Accelerated hypothesis generation to a wider audience authorized to use de-identified data
- Rapid Query – all data in main memory
- Visual Profiling of selected patients
- Export to Excel or Flat File

Research Data Browser



Research Data Browser - All Selections

UCSF Medical Center
UCSF Benioff Children's Hospital

Current Selections
%ViewToggle Full View

Clear Selections

Search

[Patient](#)
[All Selections](#)
[Cohort](#)
[Explore](#)
[Demographics](#)
[Diagnoses](#)
[Procedures](#)
[Medications](#)
[Labs](#)
[Encounters](#)
[Info](#)
NEW!

Patients 817,764 (100.0%)

Encounter Dates	Name	Ther Class
Encounter Information	Generic Name	*Unspecified
Encounter Location Information	Ther Class	ANALGESIC AND ANTIHISTAMINE COMBINATION
Encounter Hospital Information	Pharm Class	ANALGESICS
Patient (Latest Values)	Pharm Subclass	ANESTHETICS
Diagnoses (ICD-9)	Order Mode	ANTI-OBESITY DRUGS
Diagnoses (ICD-10)	Strength	ANTIALLERGY
Medications	Form	ANTIARTHRITICS
Procedures	Controlled	ANTIASTHMATICS
Procedures II	Pat Age as of Order Date	ANTIBIOTICS
Labs	Encounter IP/OP	ANTICOAGULANTS
Vitals	Ordered Date	ANTIDOTES
Immunizations	Ordered Time	ANTIFUNGALS
Coverages	Start Date	ANTIHISTAMINE AND DECONGESTANT COMBINATION
Billing Codes	Start Time	ANTIHISTAMINES
Providers		ANTIHYPERTENSIVES
		ANTIINFECTIVES
		ANTIINFECTIVES/MISCELLANEOUS
		ANTINEOPLASTICS
		ANTIPARKINSON DRUGS
		ANTIPLATELET DRUGS
		ANTIVIRALS
		AUTONOMIC DRUGS
		BIOLOGICALS
		BLOOD
		CARDIAC DRUGS
		CARDIOVASCULAR
		CNS DRUGS
		COLONY STIMULATING FACTORS
		CONTRACEPTIVES
		COUGH/COLD PREPARATIONS
		DIAGNOSTIC
		DIURETICS
		EENT PREPS
		ELECT/CALORIC/H2O
		GASTROINTESTINAL
		HEPATIC

Refreshed as of: 01/02/2017

[Report An Issue](#) [Show Help](#)

Research Data Browser - All Selections

UCSF Medical Center

[Patient](#)
[All Selections](#)
[Cohort](#)
[Explore](#)
[Demographics](#)
[Diagnoses](#)
[Procedures](#)
[Medications](#)
[Labs](#)
[Encounters](#)
[? Info](#)
NEW!

UCSF Benioff Children's Hospital

Current Selections

Medication ... [ANTHYPERGL...](#)

%ViewToggle [Full View](#)

[Clear Selections](#)

Search

Patients 26,519 (3.2%)

		Ther Class
Encounter Dates	Name	ANTHYPERGLYCEMICS
Encounter Information	Generic Name	*Unspecified
Encounter Location Information	Ther Class	ANALGESIC AND ANTIHISTAMINE COMBINATION
Encounter Hospital Information	Pharm Class	ANALGESICS
Patient (Latest Values)	Pharm Subclass	ANESTHETICS
Diagnoses (ICD-9)	Order Mode	ANTI-OBESITY DRUGS
Diagnoses (ICD-10)	Strength	ANTIALLERGY
Medications	Form	ANTIARTHRITIC
Procedures	Controlled	ANTIASTHMATIC
Procedures II	Pat Age as of Order Date	ANTIBIOTICS
Labs	Encounter IP/OP	ANTICOAGULANTS
Vitals	Ordered Date	ANTIDOTES
Immunizations	Ordered Time	ANTIFUNGALS
Coverages	Start Date	ANTIHISTAMINE AND DECONGESTANT COMBINATION
Billing Codes	Start Time	ANTIHISTAMINES
Providers		ANTIINFECTIVES
		ANTIINFECTIVES/MISCELLANEOUS
		ANTI NEOPLASTICS
		ANTIPARKINSON DRUGS
		ANTIPLATELET DRUGS
		ANTIVIRALS
		AUTONOMIC DRUGS
		BIOLOGICALS
		BLOOD
		CARDIAC DRUGS
		CARDIOVASCULAR
		CNS DRUGS
		COLONY STIMULATING FACTORS
		CONTRACEPTIVES
		COUGH/COLD PREPARATIONS
		DIAGNOSTIC
		DIURETICS
		EENT PREPS
		ELECT/CALORIC/H2O
		GASTROINTESTINAL
		HERBALS

Refreshed as of: 01/02/2017

[Report An Issue](#)
[Show Help](#)

Research Data Browser - All Selections

UCSF Medical Center
UCSF Benioff Children's Hospital

Current Selections

Medication ... 10 of 9915
Medication ... ANTIHYPERGL...
%ViewToggle Full View

Clear Selections

Search

[Patient](#)
[All Selections](#)
[Cohort](#)
[Explore](#)
[Demographics](#)
[Diagnoses](#)
[Procedures](#)
[Medications](#)
[Labs](#)
[Encounters](#)
[Info](#)
NEW!

Patients 7,975 (1.0%)

Encounter Dates	Name	Generic Name
Encounter Information	Generic Name	metFORMIN 1,000 mg tablet
Encounter Location Information	Ther Class	metFORMIN 500 mg tablet
Encounter Hospital Information	Pharm Class	metFORMIN 500 mg/5 mL oral solution
Patient (Latest Values)	Pharm Subclass	metFORMIN 850 mg tablet
Diagnoses (ICD-9)	Order Mode	metFORMIN ER 1,000 mg 24 hr tablet,extended release
Diagnoses (ICD-10)	Strength	metFORMIN ER 1,000 mg tablet,extended release 24hr
Medications	Form	metFORMIN ER 500 mg 24 hr tablet,extended release
Procedures	Controlled	metFORMIN ER 500 mg tablet,extended release 24 hr
Procedures II	Pat Age as of Order Date	metFORMIN ER 500 mg tablet,extended release 24hr
Labs	Encounter IP/OP	metFORMIN ER 500 mg tablet,extended release 24hr
Vitals	Ordered Date	metFORMIN ER 750 mg tablet,extended release 24 hr
Immunizations	Ordered Time	acarbose 25 mg tablet
Coverages	Start Date	acarbose 50 mg tablet
Billing Codes	Start Time	acarbose 100 mg tablet
Providers		albiglutide 30 mg/0.5 mL subcutaneous pen injector
		albiglutide 50 mg/0.5 mL subcutaneous pen injector
		alogliptin 12.5 mg tablet
		alogliptin 25 mg tablet
		bromocriptine 0.8 mg tablet
		canagliflozin 50 mg-metformin 1,000 mg tablet
		canagliflozin 100 mg tablet
		canagliflozin 150 mg-metformin ER 1,000 mg tablet,extend release 24 hr
		canagliflozin 300 mg tablet
		dapagliflozin 5 mg-metformin ER 1,000 mg tablet,extended release 24hr
		dapagliflozin 5 mg tablet
		dapagliflozin 10 mg-metformin ER 1,000 mg tablet,extended release 24hr
		dapagliflozin 10 mg tablet
		dulaglutide 0.75 mg/0.5 mL subcutaneous pen injector
		dulaglutide 1.5 mg/0.5 mL subcutaneous pen injector
		empagliflozin 10 mg-linagliptin 5 mg tablet
		empagliflozin 10 mg tablet
		empagliflozin 25 mg-linagliptin 5 mg tablet
		empagliflozin 25 mg tablet
		exenatide 5 mcg/dose (250 mcg/mL)1.2 mL subcutaneous pen injector
		exenatide 10 mcg/dose(250 mcg/mL)2.4 mL subcutaneous pen injector

Refreshed as of: 01/02/2017

[Report An Issue](#)
[Show Help](#)

Research Data Browser - Explore / Export

The screenshot displays the Research Data Browser interface. The top navigation bar includes 'Patient', 'All Selections', 'Cohort', 'Explore', 'Demographics', 'Diagnoses', 'Procedures', 'Medications', 'Labs', 'Encounters', '? Info', and 'NEW!'. The left sidebar shows 'UCSF Medical Center' and 'UCSF Benioff Children's Hospital' with 'Current Selections' for Medication (10 of 9915), ANTIHYPERGL..., %AdHoc Medication, Dimension L... Generic Name, and %ViewToggle Full View. The main content area shows 'Patients 7,975 (1.0%)' and an 'Adhoc Report' table. The table has columns for Medication Generic Name, # Encounters, and # Patients. The 'Table' view is selected. On the right, there are buttons for 'Estimate Export', 'Generate Export', and 'Download Export' along with estimated size and duration information. The bottom of the interface includes a refresh timestamp 'Refreshed as of: 01/02/2017' and links for 'Report An Issue' and 'Show Help'.

Adhoc Report

Medication Generic Name	# Encounters	# Patients
	1432758	9980
metFORMIN 500 mg tablet	750791	5600
metFORMIN 1,000 mg tablet	348659	2307
metFORMIN ER 500 mg tablet...	106881	672
metFORMIN 850 mg tablet	128592	763
metFORMIN ER 750 mg tablet...	32639	207
metFORMIN ER 1,000 mg tabl...	22023	142
metFORMIN ER 500 mg 24 hr t...	13311	85
metFORMIN ER 1,000 mg 24 h...	17920	117
metFORMIN ER 500 mg tablet...	7576	57
metFORMIN 500 mg/5 mL oral ...	4376	30

Research Data Browser - All Selections

UCSF Medical Center
UCSF Benioff Children's Hospital

Current Selections
%ViewToggle Full View

Clear Selections

Search

Patient All Selections Cohort Explore Demographics Diagnoses Procedures Medications Labs Encounters Info **NEW!**

Patients **862,357 (100.0%)**

Encounter Dates	Name
Encounter Information	Proc Name
Encounter Location Information	Proc Category
Encounter Hospital Information	Abnormal
Patient (Latest Values)	Unit
Diagnoses (ICD-9)	Value
Diagnoses (ICD-10)	LOINC
Medications	LOINC Name
Procedures	LOINC Level 1
Procedures II	LOINC Level 2
Labs	LOINC Level 3
Vitals	Pat Age as of Result Date
Immunizations	Encounter IP/OP
Coverages	Result Date
Billing Codes	Result Time
Providers	Collection Date
	Collection Time

creatinine

Name

- AMYLASE/CREATININE RATIO
- BUN/CREATININE RATIO (EXTERNAL LAB)
- CALCIUM/CREATININE RATIO (EXTERNAL LAB)
- CREATININE
- CREATININE (NI)
- CREATININE (SELRU)
- CREATININE (SK),UR
- CREATININE CLEARANCE - EXTERNAL
- CREATININE CLEARANCE (EXTERNAL LAB)
- CREATININE RANDOM UR
- CREATININE WHOLE BLOOD
- CREATININE(CRT),U
- CREATININE, 24 (OR TIMED) URINE - EXTERNAL
- CREATININE, BODY FLUID
- CREATININE, RANDOM U
- CREATININE, RANDOM UR
- CREATININE, RANDOM URINE
- CREATININE, SERUM
- CREATININE, SERUM - EXTERNAL
- CREATININE, UR 24HR (EXTERNAL LAB)
- CREATININE, UR 24HR (EXTERNAL LAB)
- CREATININE, URINE
- OXALIC ACID, 24HR UR (MG/G CREATININE) (EXTERNAL LAB)
- PROTEIN CREATININE RATIO - EXTERNAL
- SODIUM/CREATININE, URINE (EXTERNAL LAB)
- URINE MICROALBUMIN/CREATININE RATIO - EXTERNAL

Research Data Browser - All Selections

UCSF Medical Center

UCSF Benioff Children's Hospital

Current Selections

Lab Name ↗
 CREATININE, CREATININE, WHOLE BLOOD, CREATININE, SERUM, CREATININE, SERUM - EXTERNAL

%ViewToggle ↗
 Full View

← Clear Selections →

🔍 Search

🏠 Patient
All Selections
Cohort
Explore
Demographics
Diagnoses
Procedures
Medications
Labs
Encounters
? Info
NEW!

Patients 201,660 (23.4%)

Encounter Dates	Name	>0<3.0
Encounter Information	Proc Name	0.02
Encounter Location Information	Proc Category	0.074
Encounter Hospital Information	Abnormal	0.08
Patient (Latest Values)	Unit	0.085
Diagnoses (ICD-9)	Value	0.10
Diagnoses (ICD-10)	LOINC	0.11
Medications	LOINC Name	0.12
Procedures	LOINC Level 1	0.13
Procedures II	LOINC Level 2	0.14
Labs	LOINC Level 3	0.15
Vitals	Pat Age as of Result Date	0.16
Immunizations	Encounter IP/OP	0.17
Coverages	Result Date	0.18
Billing Codes	Result Time	0.187
Providers	Collection Date	0.19
	Collection Time	0.20
		0.21
		0.22
		0.23
		0.24
		0.25
		0.26
		0.27
		0.28
		0.29
		0.3
		0.31
		0.32
		0.33
		0.34
		0.35
		0.36
		0.37
		0.38
		0.39

Refreshed as of: 04/18/2017
[Report An Issue](#) [Show Help](#)

Research Data Browser - All Selections

UCSF Medical Center
UCSF Benioff Children's Hospital

[Patient](#)
[All Selections](#)
[Cohort](#)
[Explore](#)
[Demographics](#)
[Diagnoses](#)
[Procedures](#)
[Medications](#)
[Labs](#)
[Encounters](#)
[Info](#)
NEW!

Patients 198,354 (23.0%)

Current Selections

Lab Value >0<3,0
CREATININE,
CREATININE
WHOLE BLOOD

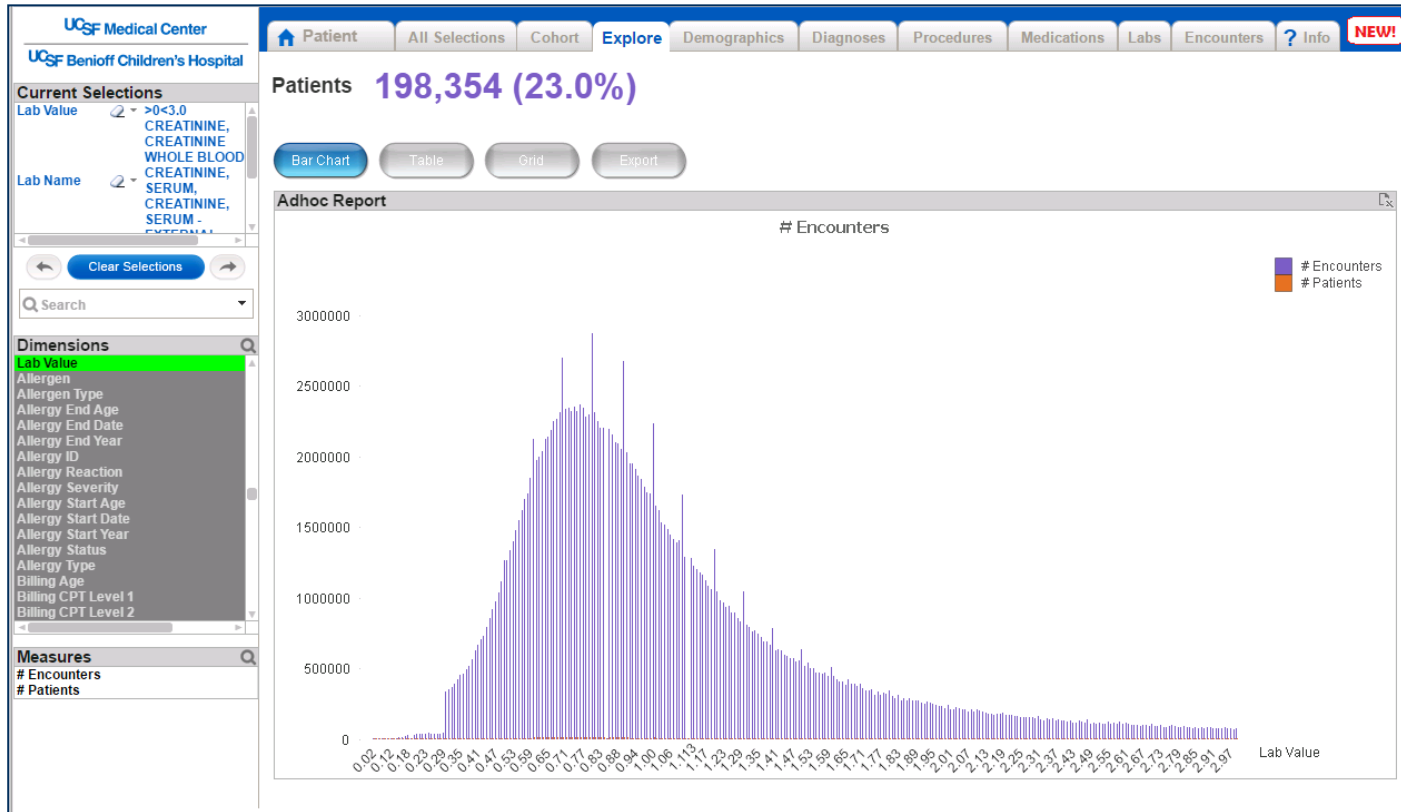
Lab Name CREATININE,
SERUM,
CREATININE,
SERUM -
EXTERNAL

Clear Selections

Search

Encounter Dates	Name	Value
Encounter Information	Proc Name	0.02
Encounter Location Information	Proc Category	0.074
Encounter Hospital Information	Abnormal	0.08
Patient (Latest Values)	Unit	0.085
Diagnoses (ICD-9)	Value	0.10
Diagnoses (ICD-10)	LOINC	0.11
Medications	LOINC Name	0.12
Procedures	LOINC Level 1	0.13
Procedures II	LOINC Level 2	0.14
Labs	LOINC Level 3	0.15
Vitals	Pat Age as of Result Date	0.16
Immunizations	Encounter IP/OP	0.17
Coverages	Result Date	0.18
Billing Codes	Result Time	0.187
Providers	Collection Date	0.19
	Collection Time	0.20
		0.21
		0.22
		0.23
		0.24
		0.25
		0.26
		0.27
		0.28
		0.29
		0.3
		0.31
		0.32
		0.33
		0.34
		0.35
		0.36
		0.37
		0.38
		0.39

Research Data Browser - Explore



Research Data Browser - Cohort

UCSF Medical Center
UCSF Benioff Children's Hospital

Current Selections
PatientID2 2938 of 817764
%ViewToggle Full View

Clear Selections

Search

[Patient](#)
[All Selections](#)
[Cohort](#)
[Explore](#)
[Demographics](#)
[Diagnoses](#)
[Procedures](#)
[Medications](#)
[Labs](#)
[Encounters](#)
[Info](#)
NEW!

Patients 2,938 (0.4%)

Group	Count	Medications
Gp01	7,975 Pts	10 of 9915 ANTIHYPERTENSIVES Medication Generic Name
Gp02	20,043 Pts	ANTIHYPERTENSIVES, ACE INHIBITORS
Gp03	817,764 Pts	
Gp04	817,764 Pts	
Gp05	817,764 Pts	
Gp06	817,764 Pts	

Coh01 = Gp01*Gp02 Coh02 = Gp02 Coh03 = Gp03

Cohort	Count	Medications
Coh01	2,938 Pts	2938 of 817764
Coh02	817,764 Pts	
Coh03	817,764 Pts	

Compare Cohort Demographics

Patient Age

Refreshed as of: 01/02/2017

[Report An Issue](#) [Show Help](#)

Deidentified APeX Data Flat Files

- Use Cases:
 - “For people who code”
 - Use your own analytic tools
 - Machine learning; Hypothesis Discovery
 - Example: Readmission Rates – What are the predictors?
- Sources of Data:
 - 94 GB - Research Data Browser extract (since 6/1/2012)
 - 115 GB - All APeX Flowsheets (since 6/1/2012)
- Consistent surrogate Patient Identifiers
- Consistent Dates – Same random date offset per patient
- We can re-identify cases later with CHR Approval

Deidentified APeX Data Flat Files

as of 4/28/2017

Table	Rows
PATIENTS	865,693
PATIENT RACE	877,872
ALLERGIES	1,041,279
BILLING	42,545,196
COVERAGE	11,261,772
DEPARTMENTS	28,886,762
DIAGNOSES	48,528,046
ENCOUNTERS	23,458,078

Table	Rows
FLWSHEETROWDIM	46,825
FLWSHEEVALUEFACT	1,507,075,050
IMMUNIZATIONS	320,442
LAB	73,914,519
MEDICATION ORDERS	21,096,655
PROCEDURE ORDERS	33,644,888
PROVIDERS	31,046,836
VITALS	2,573,564