Epic EHR

Identifying Study Populations and Managing Research Studies using Epic Analytics Tools

12/14/2023

The GW Medical Faculty Associates

Introductions

Jorge Sepulveda, MD, PhD

Chief Medical Analytics Officer
Professor of Pathology
Medical Director of Pathology and Laboratory
Informatics
George Washington University SMHS

jsepulveda@mfa.gwu.edu

Liz Owens

Manager of Business
Intelligence
The GW Medical Faculty
Associates

elowens@mfa.gwu.edu

The best way to reach us for questions and requests is to email
solutioncenter@mfa.gwu.edu

This will open a support ticket for Epic Reporting

Epic Reporting Tools Overview

- We have four primary Self-Service Analytics tools

- End users need an Epic account to use any of the tools

- We have a team of 5 Business Intelligence analysts to assist you with submitting requests and learning the tools or to assist with custom analytics development as needed.

- We also have a connection to a team of data scientists for more advanced analytical support.

- We have a SharePoint request site to coordinate review and approval of research requests through our Research Advisory Council
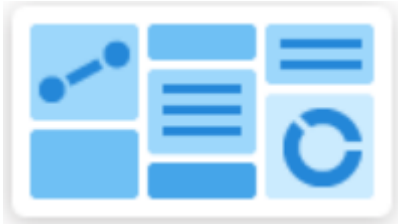
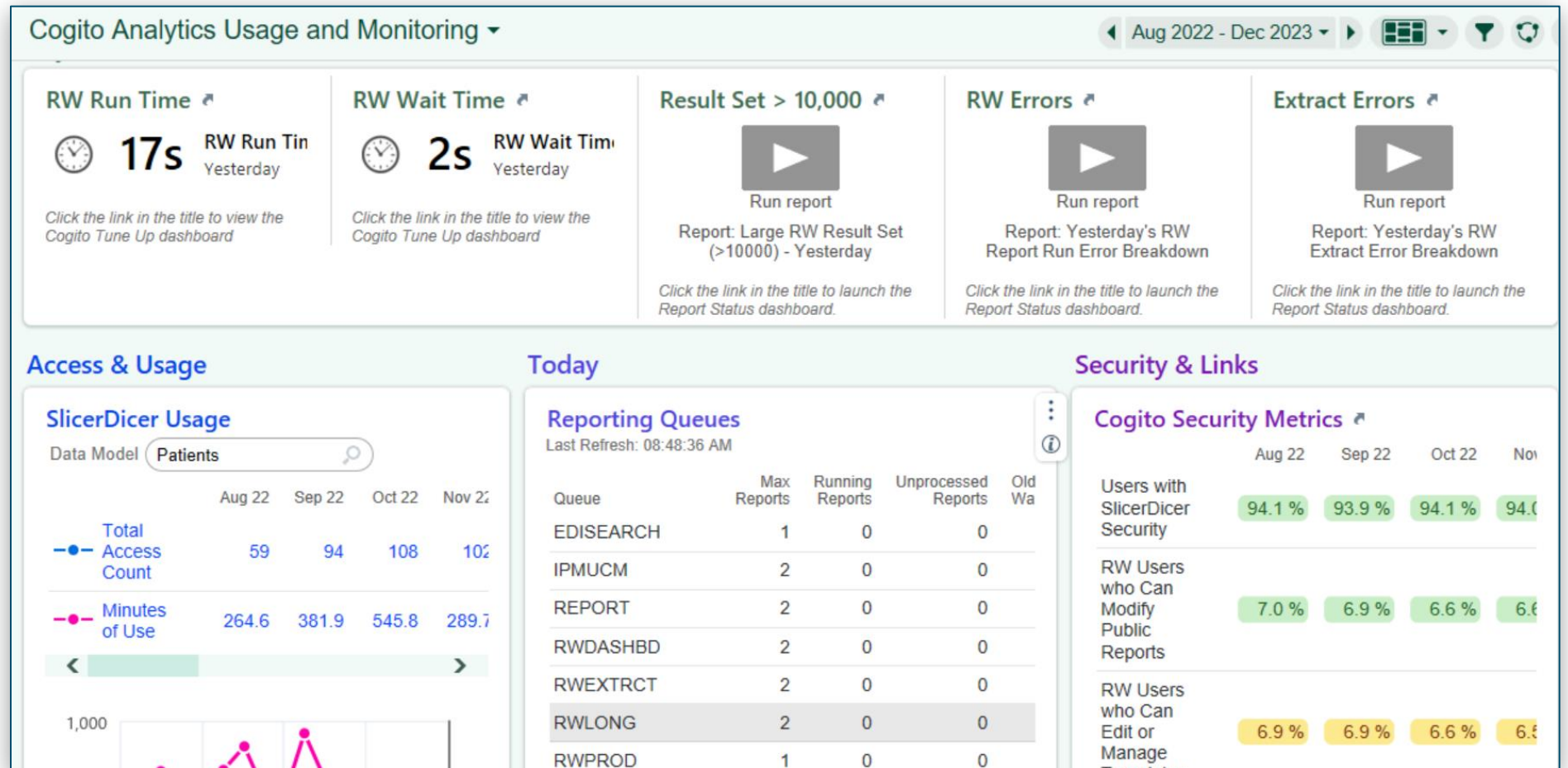DASHBOARDS   WORKBENCH   SLICERDICER   COSMOS

Epic Reporting Tools Overview

DASHBOARDS

- Used for visualizing and monitoring multiple data sources and reporting content in one place

- Organized around roles and workflows

- High level views with additional drill down details

Epic Reporting Tools Overview

WORKBENCH

- Real-time, actionable data

- Excel look and feel with limited summarization functionality

- Mostly geared towards finding "my" patients.

Epic Reporting Tools Overview

SLICERDICER

- Best Epic Reporting tool for investigating and identifying MFA populations

- Multiple visualization styles and pre-defined measures

- See all aggregate data, but only see the details for your patients
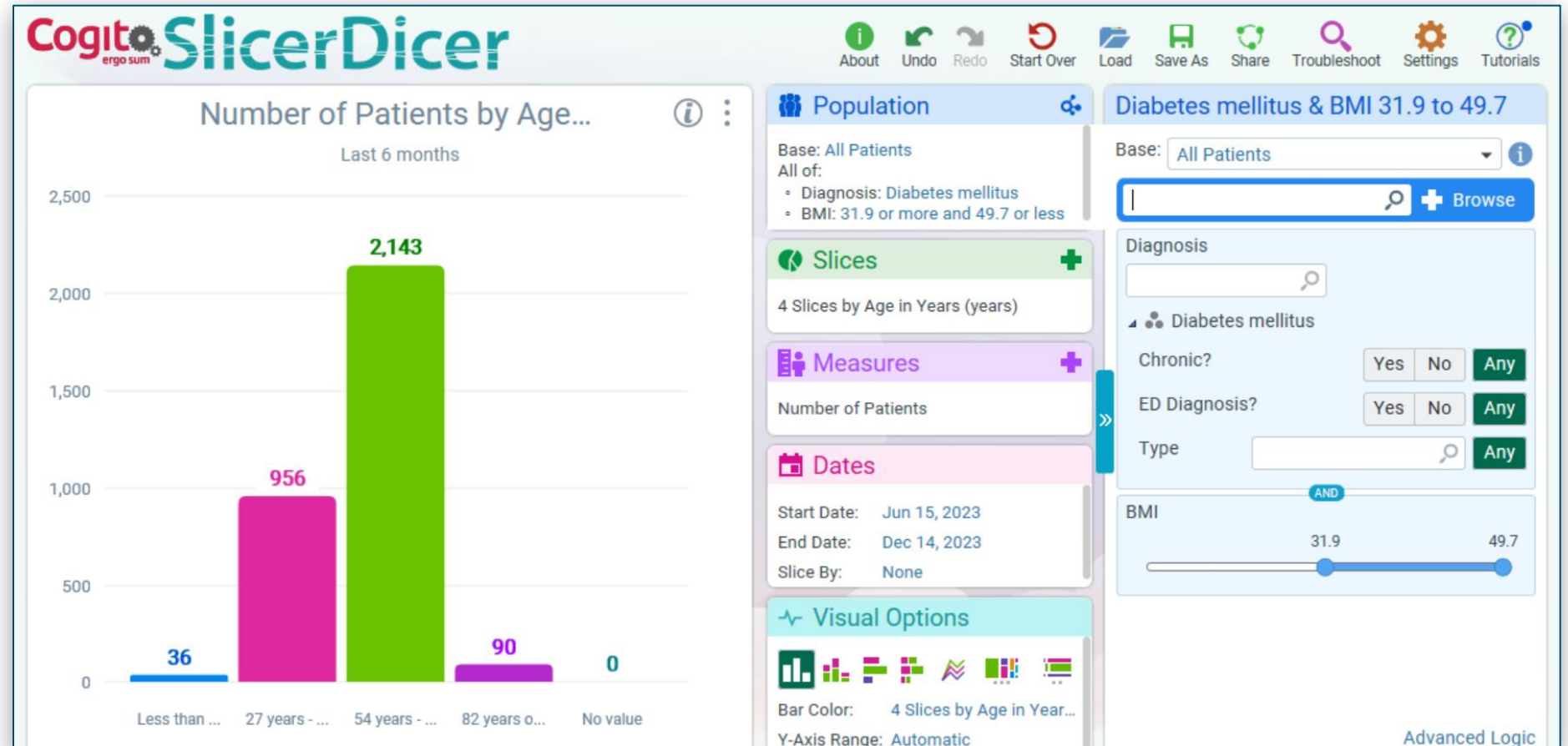
Epic Reporting Tools Overview

COSMOS

- Deidentified data from hundreds of participating organizations and hundreds of millions of data

- Similar in look and feel to SlicerDicer

- Levels of access from regular end user to Data Scientist
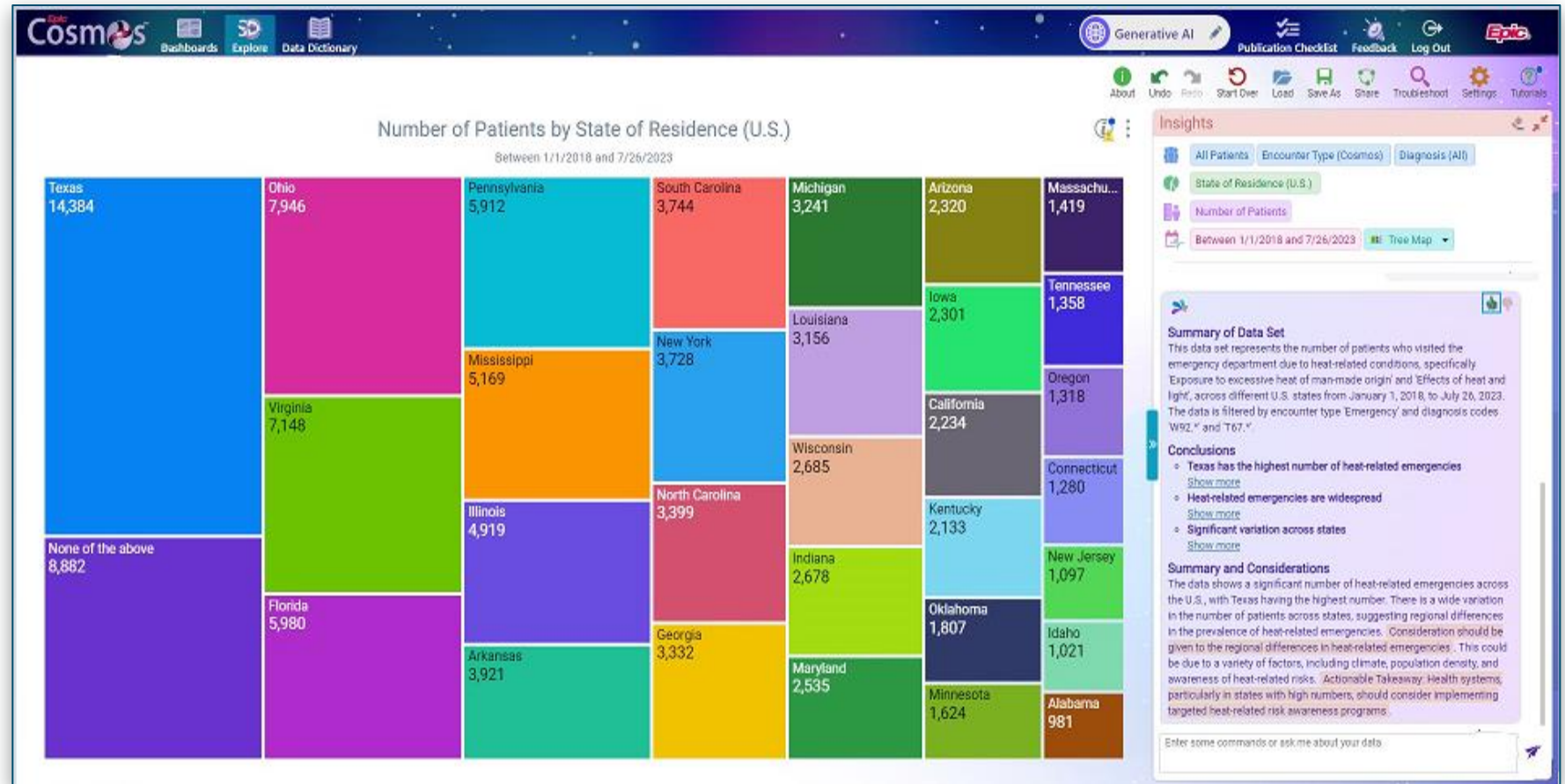
Questions?

The best way to reach us for questions and requests is to email
solutioncenter@mfa.gwu.edu

This will open a support ticket for Epic Reporting

GW Medical Faculty Associates

gwdocs.com

Identifying Research Populations Using the EHR

Hiroki Morizono, PhD

hiroki@gwu.edu

What is an Electronic Health Record?

- To a first approximation, it is an electronic version of a patient's paper chart
  - Administrative and Billing Data
  - Demographics
  - Vital signs
  - Diagnoses
  - Medications
  - Immunization Dates
  - Allergies
  - Lab and Test Results
  - Radiology images
  - Progress notes

Capture evidence needed to make decisions about patient care

Provide workflow management so information can be shared across the organization

Types of Healthcare Data

- The Electronic Health Record
- Administrative data
  - Hospital discharge data
- Claims data
  - Billing data
- Disease registries
  - Specialized tracking of specific disease
- Clinical trial data
- Population Health
  - NHANES
  - Medicare
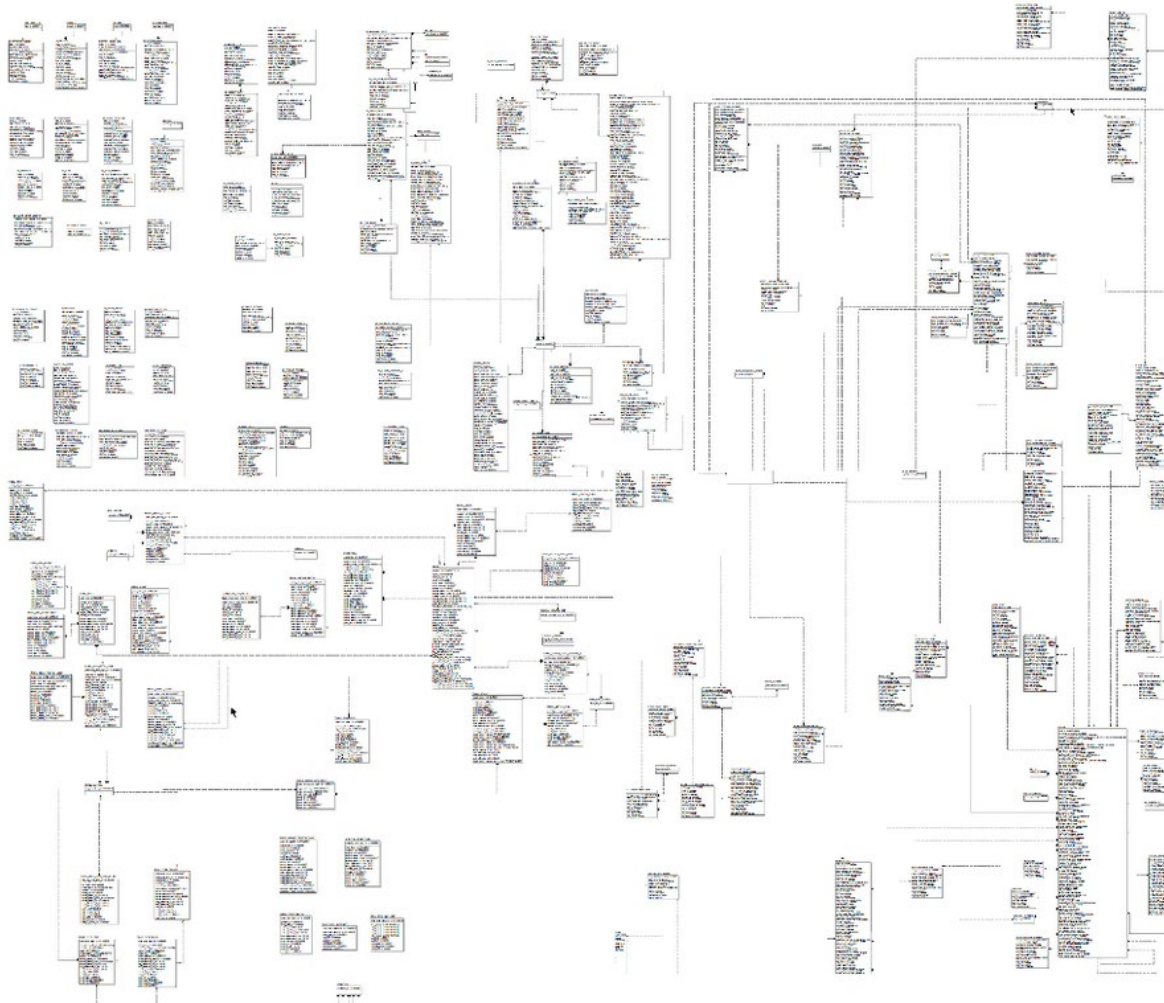- Genomic respositories
- Other

Huge Volume
Huge Variety

EHRs are complex

- They address a variety of needs
- As a result, the traditional relational database approaches become quite unwieldy very quickly

- Epic and Cerner (now Oracle Health) are the two largest EHR vendors
  - Epic is used at the MFA and has probably overtaken Cerner
  - Cerner used at the GW Hospital and Children's National

A fraction of the Cerner Millenium Database schema. Each table has thousands if not millions of rows of data

About 30,000 tables the last time I looked

Many of these tables need to be joined with others to get useful information

diagnosis – 0.4 Mb	
ENCNTR_KEY 1435	DIAGNOSIS_KEY 890

encounter_waitlist – 18 Mb	
PERSON_KEY 1858	WAITLIST_KEY 775

procedure – 65 Mb	
ENCNTR_KEY 1435	PROCEDURE_ID 546

encounter_radiology – 134 Mb	
ENCNTR_KEY 1435	CLINICAL_EVENT_KEY 188

encounter_location_history – 280 Mb	
ENCNTR_KEY 1435	ENCNTR_LOC_HIST_KEY 229

encounter_user_defined – 351 Mb	
ENCNTR_KEY 1435	ENCNTR_INFO_KEY 899

medications – 584 Mb	
ENCNTR_KEY 1435	ORDER_KEY 565

perioperative – 740 Mb	
ENCNTR_KEY 1435	SURG_PROCEDURE_KEY 115

ED_tracking – 1646 Mb	
ENCNTR_KEY 1435	TRACKING_EVENT_KEY 366

encounter_order – 1891 Mb	
ENCNTR_KEY 1435	ORDER_KEY 278

encounter_provider – 3191 Mb	
ENCNTR_KEY 1435	ENCNTR_PRSNL_RELTN_KEY 663

pathology_results – 3995 Mb	
ENCNTR_KEY 1435	CLINICAL_EVENT_KEY 192

scanned_doc - ECG pdf images – 14397 Mb	
ENCNTR_KEY 1435	CLINICAL_EVENT_KEY 78734

note_data – 19257 Mb	
ENCNTR_KEY 1435	CLINICAL_EVENT_KEY 458

form_data – 21453 Mb	
ENCNTR_KEY 1435	FORM_EVENT_KEY 789

encounter – 181 Mb	
PERSON_KEY 1858	ENCNTR_KEY 1435
	ENCNTR_KEY 2423
	ENCNTR_KEY 1544

patient – 9.9 Mb	
PERSON_KEY 1858	GIVEN_NAMEaaaa
	MIDDLE_NAMEbbbb
	SURNAME cccc

patient_family_history – 0.06 Mb	
PERSON_KEY 1858	ACTIVITY_KEY 12

patient_past_medical_history – 0.08 Mb	
PERSON_KEY 1858	PROBLEM_KEY 5467

patient_allergy – 1.3 Mb	
PERSON_KEY 1858	ALLERGY_KEY 1728

alert – 5 Mb	
ENCNTR_KEY 1435	PROBLEM_KEY 111

patient_social_history – 20 Mb	
PERSON_KEY 1858	SOCIAL_HISTORY_KEY 7104

Benefits of Mining EHR and other Healthcare Data

Systematic analysis of patterns

Developing a computational "phenotype" of a disease or a population
Can provide a wealth of information if you have the right access permissions and knowledge

Learning how to use an EHR at a site doesn't necessarily mean you know how to use the EHR from that vendor at another hospital.

The EHRs are significantly customized

Challenges in Mining EHR and other Healthcare Data

- Ensuring privacy
- Access is restricted for some data elements
- Data is scattered across systems and locked in some instances
- Data cannot be analyzed without extensive data wrangling
- Data from different parts of an organization are in different units
- Data variety
- Important data is kept in unstructured fields

- **EHRs are actively in use by clinical staff to enter data; extensive queries can degrade performance for data entry**

Cerner Millennium PowerChart

- Cerner was acquired by Oracle and their products and services are undergoing rebranding and renaming
- Millennium PowerChart is the name of the system that includes the Electronic Health Record (EHR)
- Bear TRACKS, Power Chart, Millennium are used interchangeably
- The EHR includes a reporting tool called Discern
  - To get the most out of Discern you will likely need to learn Cerner Command Language (a variant of Structured Query Language)

- Millennium is good for viewing an individual patient record but not well suited for querying across a cohort

Data from these systems feed into Cerner Millennium
From there, data is pulled into HealtheIntent
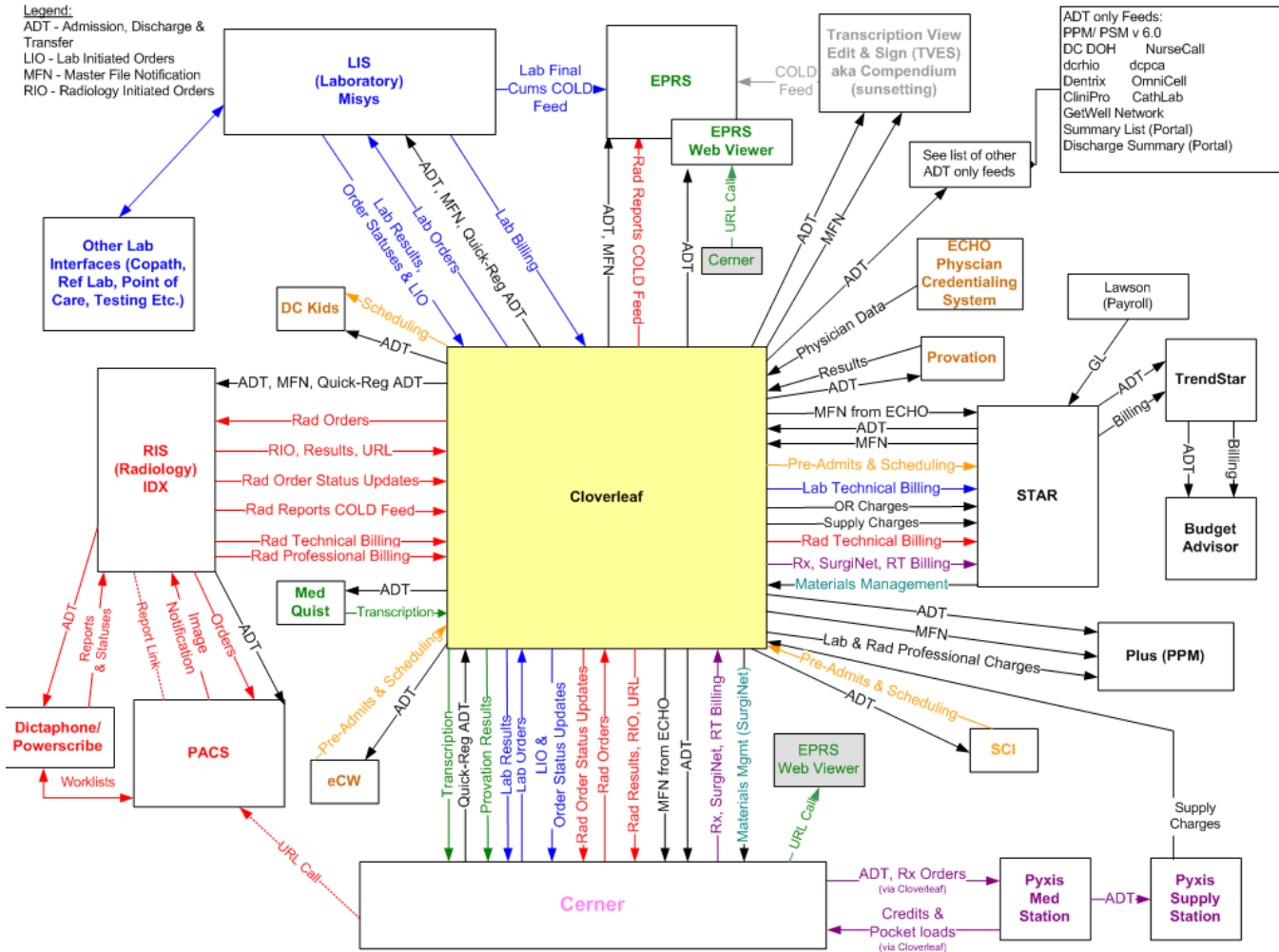
HealtheIntent

- HealtheIntent is our Enterprise Data Warehouse (EDW)
- HealtheAnalytics is an SQL based workbench to query the EDW
- Find the HealtheAnalytics New User Access Request Job Aid in the Intranet (they keep changing the location of things in there)
  - A search for Discern found GRP_IS_Learning_Clinical Informatics which led to the correct location *under Discern Analytics and HealtheAnalytics Materials*
  - Instructions in there to have your Network ID added to the correct access groups

- https://childrensnational.analytics.healtheintent.com/

What is HealtheIntent?

- Originally Cerner's population health platform, aggregates information from the disparate systems shown earlier.

- Unstructured data (notes, images, pdfs) are not included

- Identifiers remain in the data (not deidentified)

- Data are normalized, for example the lab value units are harmonized

- Custom datasets and reports can be created and secured

- Reporting tools are included

HealtheIntent/HealthAnalytics

Users with the appropriate access can create their own reports using:

- Business Objects - drag-and-drop reporting to create data tables or patient lists
- Tableau - dashboarding, visual analytics and ad-hoc exploration
- SQL – for more complex data analyses and extracts
- Data Syndication to automatically download raw data files

- Most users will need support to perform queries
  - https://cnmc.sharepoint.com/sites/GRP_AnalyticsChampions_InformationResources/SitePages/analytics-champions-home.aspx
  - Request Analytic/Reporting/Data Extract at https://cnmc.sharepoint.com/sites/MyIT

- Training also provided at CNH

Cerner's counterparts to Epic Slicer Dicer and Cosmos

- HealtheDataLab (nearly ready for use at CNH)
  - Jupyter notebook front end to a Cerner Data Lake incorporating HealtheIntent
        Both Children's data and Cerner Real World Data
  - Cerner Real World Data
    - Float a query across data provided from Cerner HealthIntent sites
    - >150million patients
- Cerner Learning Health Network
  - Similar in concept to PEDSnet
  - Also provides access to CRWD
  - A front door request is needed to access, and seats are limited

- These will require familiarity with Python or R as well as a variety of packages for handling large data sets

TriNetX

- Self service cohort discovery tool
  - Can use as part of preparatory to research
- Cloud based health research platform
  - Can query Children's National, US sites, international sites
  - Clean easy to use interface with crosswalks between common coding systems

Can query

- Demographics
- Diagnosis (ICD-9-CM and ICD-10-CM)
- Medications
- Procedures
- Labs
- Visit Types

Ability to link back to HealtheIntent and with an approved IRB protocol Obtain additional data from the EDW

Create a new study

Obtain Summary counts and demographics

- Enter inclusion and exclusion criteria

Run

View History

Demographics

Grouped | Stacked   ☑ F   ☑ M   ☑ U   📷
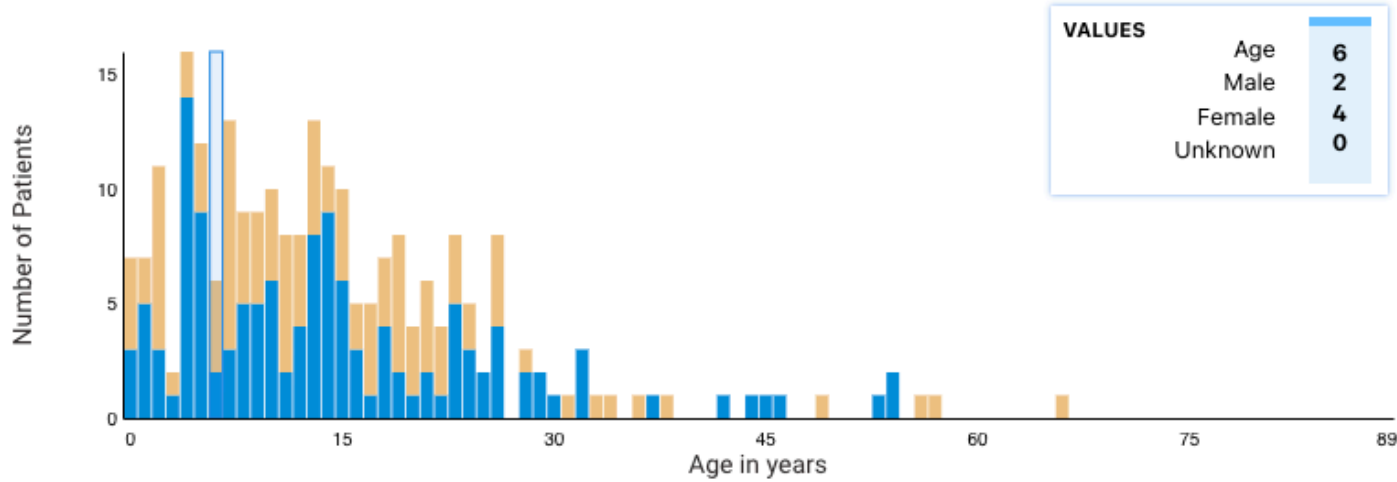
VALUES

Age	6
Male	2
Female	4
Unknown	0

Patients 90 and Older: 0

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
240	**0**	**66**	**15**	**12**

Sex

Male	**54%**	████
Female	**46%**	███
Unknown	**0%**	

Race

White	**33%**	███
Other Race	**33%**	███
Black or Africa...	**20%**	██
Asian	**12%**	█
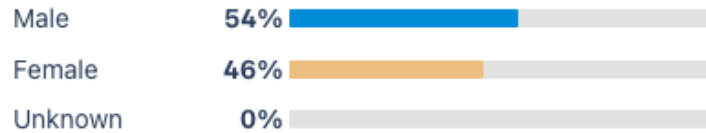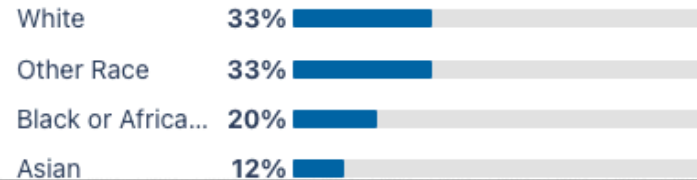
Diagnoses distribution

Diagnoses   Diagnoses within | 3M | 6M | 12M | 24M | Anytime | ❓ | All | Acute | Chronic | ❓ | 📷

Search...

Diagnoses			Patients	% of Cohort
>	ICD-10-CM E00-E89	Endocrine, nutritional and metabolic diseases	240	100%
>	ICD-10-CM R00-R99	Symptoms, signs and abnormal clinical and laboratory ...	230	96%
>	ICD-10-CM Z00-Z99	Factors influencing health status and contact with healt...	190	79%
>	ICD-10-CM A00-B99	Certain infectious and parasitic diseases	170	71%
>	ICD-10-CM J00-J99	Diseases of the respiratory system	150	63%
>	ICD-10-CM M00-M99	Diseases of the musculoskeletal system and connectiv...	150	63%
>	ICD-10-CM K00-K95	Diseases of the digestive system	140	58%
>	ICD-10-CM G00-G99	Diseases of the nervous system	130	54%
>	ICD-10-CM N00-N99	Diseases of the genitourinary system	130	54%
>	ICD-10-CM F01-F99	Mental, Behavioral and Neurodevelopmental disorders	110	46%
>	ICD-10-CM D50-D89	Diseases of the blood and blood-forming organs and c...	100	42%
>	ICD-10-CM S00-T88	Injury, poisoning and certain other consequences of ex...	100	42%
>	ICD-10-CM H00-H59	Diseases of the eye and adnexa	90	38%
>	ICD-10-CM I00-I99	Diseases of the circulatory system	90	38%
>	ICD-10-CM L00-L99	Diseases of the skin and subcutaneous tissue	80	33%

321   7   Data Flow

Search... ∧ ∨ ✕

Procedure			Patients	% of Cohort
› SNOMED	128927009	Procedure by method	180	75%
› CPT	1013625	Evaluation and Management	170	71%
› SNOMED	243120004	Regimes and therapies	170	71%
› SNOMED	362958002	Procedure by site	170	71%
› CPT	1012569	Medicine Services and Procedures	140	58%
› CPT	1003143	Surgery	110	46%
› CPT	1011136	Pathology and Laboratory Procedures	110	46%
› SNOMED	363691001	Procedure categorized by device involved	90	38%
› CPT	1010251	Radiology Procedures	80	33%
› CPT	1002796	Anesthesia	50	21%
› HCPCS	J	J: Drugs Administered Other Than Oral Method, Chemo...	50	21%
› ICD-10-PCS	0	Medical and Surgical	50	21%
› SNOMED	362961001	Procedure by intent	50	21%
› ICD-10-PCS	3	Administration	40	17%
› ICD-10-PCS	4	Measurement and Monitoring	40	17%
› TNX Curat...	10021	Chemotherapy Lines of Treatment (VA Class)	30	13%

321   7

Medications

Medications within [3M] [6M] [12M] [24M] [Anytime] ❓ [VA] [ATC] 📷

Search... [⌃] [⌄] [✕]

Medication				Patients	% of Cohort	
›	VA	CN000	Central nervous system medications	190	79%	▰▰▰▰▱
›	VA	DE000	Dermatological agents	190	79%	▰▰▰▰▱
›	VA	TN000	Therapeutic nutrients/minerals/electrolytes	190	79%	▰▰▰▰▱
›	VA	GA000	Gastrointestinal medications	180	75%	▰▰▰▰▱
›	VA	OP000	Ophthalmic agents	180	75%	▰▰▰▰▱
›	VA	AM000	Antimicrobials	170	71%	▰▰▰▰▱
›	VA	CV000	Cardiovascular medications	170	71%	▰▰▰▰▱
›	VA	HS000	Hormones/synthetics/modifiers	170	71%	▰▰▰▰▱
›	VA	NT000	Nasal and throat agents,topical	170	71%	▰▰▰▰▱
›	VA	PH000	Pharmaceutical aids/reagents	170	71%	▰▰▰▰▱
›	VA	RE000	Respiratory tract medications	170	71%	▰▰▰▰▱
›	VA	Other	Other medications (va)	170	71%	▰▰▰▰▱
›	VA	GU000	Genitourinary medications	160	67%	▰▰▰▰▱

Patient Arrival Rate ❓

🔘 Show graph 📷

Chart [Line] [Area]
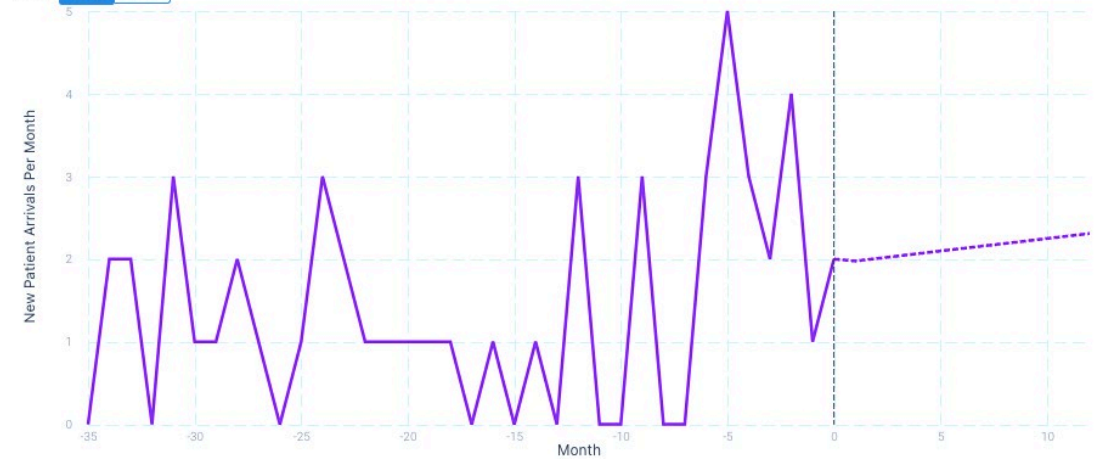
	Healthcare Organization (HCO)	Historic Arrivals (Monthly Avg Over Past 3 Yrs.)	Predicted Arrivals (Monthly Avg Over Next 1 Yr.)	Trend
☑ All				
☑ ▪	Children's National Medical C...	1.4	2.1	〰〰〰

Analyses

Analyze Outcomes

How do patients in a cohort experience outcomes?

Index Event

Explore Characteristics | Review Outcomes

● Single Cohort

Compare Outcomes

How do outcomes compare between cohorts?

Index Event

Compare Characteristics | Review Outcomes
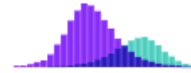
● ● Two Cohorts

Compare Cohorts

How do patient characteristics compare between cohorts?

Compare Characteristics
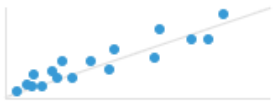
● ● Two Cohorts
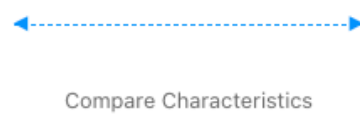
Treatment Pathways

In what order do patients receive treatments following a diagnosis?

Index Event

Review Characteristics | Review Treatment Pathways

● Single Cohort

Incidence and Prevalence

What are the incidence and prevalence of events of interest in a cohort?

Review Incidence and Prevalence

● Single Cohort
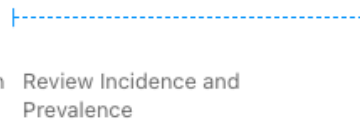
Other Analyses

Contact TriNetX or your organization's administrator to access these analyses.

Competing Risks

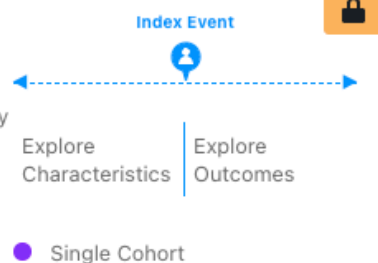**Index Event**

What is the probability for a patient to have competing risks?

Review Characteristics | Review Competing Risks

● Single Cohort
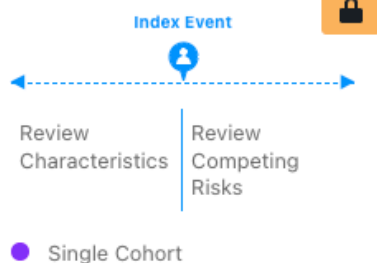
Advanced Explore Cohort

Index Event

What are the characteristics of my cohort in different time periods?

Explore Characteristics | Explore Outcomes

● Single Cohort

In The Lab

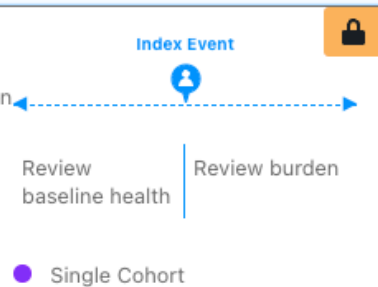New analyses under development. Contact TriNetX or your organization's administrator to get early access.

Patient Clustering

Index Event

How do outcomes compare between subtypes of like patients within a cohort?

Review Characteristics | Compare Outcomes

● Single Cohort

Burden of Illness

Index Event
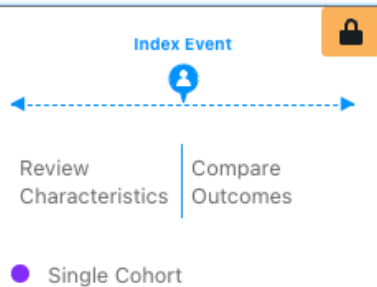
What is the burden of a disease on patients following a diagnosis?

Review baseline health | Review burden

● Single Cohort

To request TriNetX access visit ***https://is.gd/INSIGHTS***

After account is created, can find training materials
https://support.trinetx.com/hc/en-us/categories/115000239168

Other Networks

- Other Networks
  - Pedsnet
  - N3C
  - PHIS
  - PCORNET
  - All of US
  - UK Biobank

For more information

- Visit the CTSICN Research Launcher
  https://ctsicn.org/RL

- Cerner specific information https://ulearn.cerner.com

- GW Library offers several excellent workshops for Python and R

- GW Library also subscribes to O'Reilly Online Learning
  https://learning.oreilly.com/home-new/

- Children's has a Special Interest Group called Docs in Data Science and are hoping to expand to GW

- GW Coders meets regularly and may be an opportunity to find students interested in capstone projects

- Email me hiroki@gwu.edu