

HART Services to Support Data Management

Research Concierge Services: Ensures appropriate support is leveraged to provide access, acquisition, storage and utilization of research and QA/QI data, including PHI

- Secure Shared Drive
- Epic Research Access
- Data Extracts (EPIC)
- Data Collection and Storage (REDCap)
- Data Analysis: Biostatistical Services
- Data Exploration: TriNetX



Definitions

- Medical records based research is frequently performed as Retrospective Research Studies
 - Data that was originally collected for purposes other than the current research study (i.e. clinical reasons), and which exists at the time of IRB submission.
 - For a study to be considered retrospective by the IRB, all data must be in existence at time of submission.
- Prospective Chart Reviews will share similar benefits and limitations
 - Patients have not yet occurred, may require consent



Medical Records Based Research: Privacy and IT Security Concerns.....

- Protected Health Information of our patients
- Most frequently requested PHI elements
 - Name
 - MRN, MPI, CSN or any account number
 - Date of Birth
 - SSN (rarely allowed)
 - **All** dates – e.g. procedure, admission, encounter
 - Email, Addresses, phone numbers
 - Zip code that is more specific than first three digits



Medical Records Based Research:

..... Privacy and IT Security Concerns.....

- Collect only what you need!
- Carilion Secured Shared Directory for limited (no patient names, maybe a date and a zip code) or de-identified datasets (no HIPAA element)
- Anyone with access to data must either be on the study team or a Carilion employee whose job duties stipulate access or have another approved agreement (R&D)



Medical Records Based Research:

.....Privacy and IT Security Concerns.....

- Acquire and use Epic Research access if you do chart review
- REDCap required for PHI that is manually collected
- Use Carilion or your institution's email, not private
- Obtain IRB approval BEFORE working with identifiable information (consent may be required, if feasible)



Data Management

- Manual chart review vs. data extraction
 - *Long*: data extraction
 - *Wide*: chart review
 - *Notes, free text*: chart review
 - *Field calculated on the fly*: chart review
- Data collection tool indicative of amount of work
- Use **MPI** (master patient index, unique) instead of MRN (not unique)
- PHI and sensitive information: if don't need, don't collect



Data Management

- What data are needed to answer the question? Are the data available?
- Are the data consistently collected? How confident are you in the accuracy?
- Are there inter-rater reliability issues?
- What are exact definitions? (ex/ ICD codes)



Data Management

- What are the other confounding factors?
- Have data instruments changed over the time period? Upgrades to the EMR?
- Are there other initiatives that overlap your timeframe that may also impact the outcomes?
- How to define clean cohorts without bias, attrition, impacts of changing medicine?
- If the project relies on follow-up, how will attrition, discharge status, death, and ongoing Carilion relationship impact your outcomes and measurement?



Data Sources

➤ **EPIC EMR** (electronic medical record)

- Most common data categories:
demographics, encounters, diagnosis, procedures, labs, flowsheets, medications, assessments
- Most general categories:
inpatient, outpatient, emergency, operating room...

➤ **Other data sources: (examples)**

- Infection Control Susceptibilities
- Trauma Registry
- TriNetX data sets
- All Payers Claim Database (APCD)



Data Source Limitations

- Inconsistent
 - workflow change
- Incomplete
 - medical history, employment/education history
- Inaccurate
 - manual entry

Solution:

- Understand the workflow: how, when, who
- Limit the time frame to consistent documentation
- Eliminate inaccurate data



EPIC Timeline at Carilion



Ambulatory

Started 2/2008

CMC

8/2008

NRV

4/2009

Optime

10/2010

Radiant

6/2012

Cardiant

5/2013

Stork

5/2013

Beacon

10/2014

Anesthesia

11/2016

**Home
Health**

2/2017

- Incomplete data prior to go live dates
- Paper chart reviews prior dates or may not even be available.



Data Management: Data Extraction

- Most appropriate for large numbers of patient records
- Most appropriate for less variety in data types
- Different types of data (medications versus demographics) may have different sources and become separate extracts
- Requires very complete data sets



Data Management: Data Abstraction

- Most appropriate for fewer numbers of patient records
- Most appropriate for more variety in data types from different parts of medical record
- Essential if data is not stored in extractable fields
- Better technique if data fields frequently incomplete and need review to complete



Data Management: Hybrid

- Data Extract for extractable data points
- File imported into REDCap including patient identifiers
- Additional fields for chart review built into REDCap
- Chart reviewed and data entered into REDCap
- De-identified data set exported from REDCap for analysis



- **R**esearch **E**lectronic **D**ata **C**apture
- **Secure, web-based** application for managing and collection of research data
- Origin: Vanderbilt University, 2004
- >2000 institutions, >100 countries

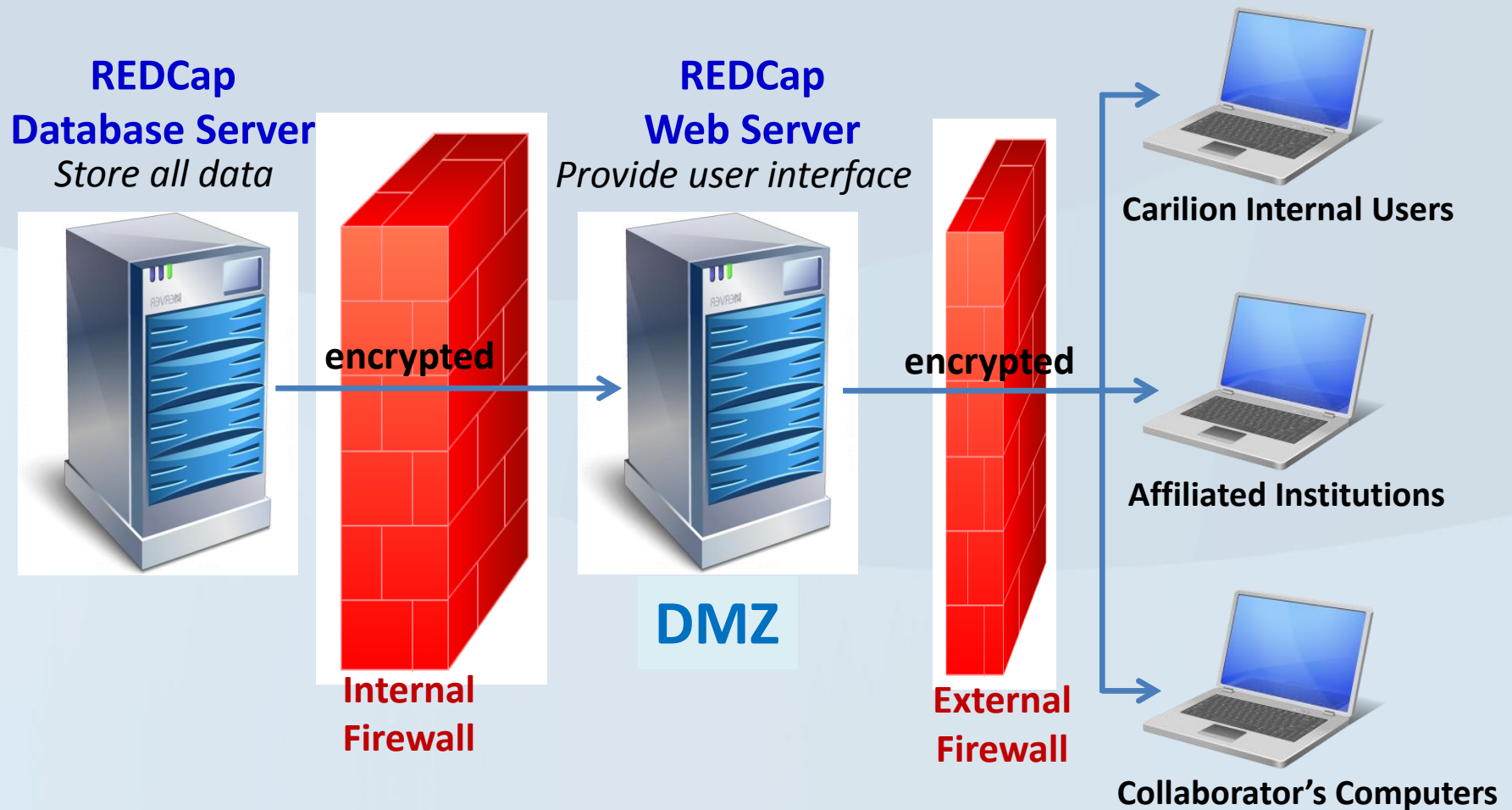
REDCap at Carilion



- **Go-live** at Carilion 2017
- Open to **Research, QA/QI and Grants**
- Installed and maintained by **HART**
- Free of charge



Carilion REDCap Infrastructure



REDCap Project Types

- **Survey**
 - Forms to collect data from respondents, typically *via* emailed links
- **Traditional project**
 - Data entry forms for consistent data collection
- **Longitudinal Study**
 - Data gathered for the same subjects over time
- **Combination of the above**



Field Types

Project Setup
Project status: **Development**

Data Collection [Edit instruments](#)

Record Status Dashboard
- View data collection status of all records

Add / Edit Records
- Create new records or edit/view existing ones

Study ID 1 [Select other record](#)

Data Collection Instruments:

Demographics

Baseline Data

Month 1 Data

Month 2 Data

Month 3 Data

Completion Data

Demographics 2

Lock all instruments

Applications

Calendar

Data Exports, Reports, and Stats

Data Import Tool

Data Comparison Tool

Logging

Field Comment Log

File Repository

User Rights and DAGs

Record Locking Customization

E-signature and Locking Mgmt

Data Quality

API and API Playground

REDCap Mobile App

Help & Information

Study ID 1 successfully edited

Save & Exit Form
Save & Stay
-- Cancel --

Editing existing Study ID 1

Study ID 1

Upload the patient's consent form [Upload document](#)

Contact Information

Name

Date of birth Today Y-M-D

Age (years) View equation

Phone number Include Area Code

E-mail

Ethnicity

☐ Hispanic or Latino

☐ NOT Hispanic or Latino

☐ Unknown / Not Reported

 reset

Race Unknown / Not Reported

Gender

☒ Female

☐ Male

☐ Transgender

☐ Unspecified

 reset

Has the patient given birth before?

☐ Yes

☐ No

 reset

Attachment

Text

Date

Calculated field

Phone #

Email

Choice: radio button

Choice: drop down

Hidden question: only appear if the above answer is 'Female'



Matrix

The following qualities are important in my supervisor:

		Extremely Important	Very Important	Somewhat Important	Slightly Important	Not Important at All	
1)	BELIEVABLE	☐	☐	☐	☐	☐	reset
2)	APPROACHABLE	☐	☐	☐	☐	☐	reset
3)	HONEST	☐	☐	☐	☐	☐	reset
4)	SUCCESSFUL	☐	☐	☐	☐	☐	reset
5)	QUALIFIED	☐	☐	☐	☐	☐	reset

Please let us know your weekly schedule for the following:

		Monday	Tuesday	Wednesday	Thursday	Friday
6)	Gym (Weight Training)	☐	☐	☐	☐	☐
7)	Aerobics	☐	☐	☐	☐	☐
8)	Eat Out (Dinner/Lunch)	☐	☐	☐	☐	☐
9)	Drink (Alcoholic Beverages)	☐	☐	☐	☐	☐



Survey Distribution Methods

- Email invite
- Public survey link
- Embed survey on website
- QR code



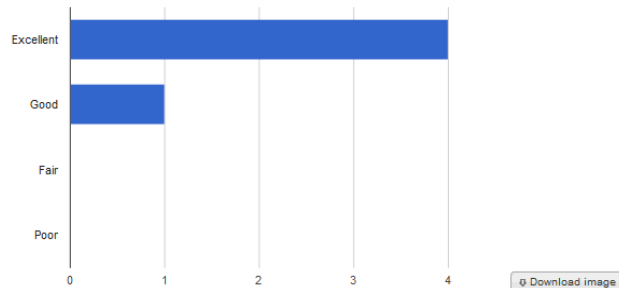
Analyzing Results

Completing a literature review.

[Refresh Plot](#) | [View as Bar Chart](#) ▼

Total Count (N)	Missing	Unique
5	0 (0,0%)	2

Counts/frequency: Excellent (4, 80,0%), Good (1, 20,0%), Fair (0, 0,0%), Poor (0, 0,0%)



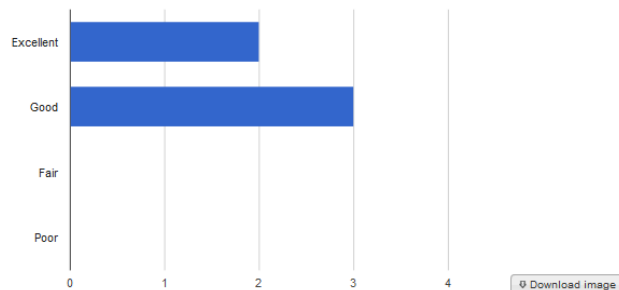
[Download image](#)

Utilizing online database search engines.

[Refresh Plot](#) | [View as Bar Chart](#) ▼

Total Count (N)	Missing	Unique
5	0 (0,0%)	2

Counts/frequency: Excellent (2, 40,0%), Good (3, 60,0%), Fair (0, 0,0%), Poor (0, 0,0%)



[Download image](#)

- Descriptive
- Export to Excel, R, SAS, SPSS, Stata for further analysis



REDCap Data Sharing

- **Multiple Sites**

Collective Data (de-identified)

Carilion REDCap

Carilion User

Carilion Data

Collaborator 1

Collaborator 1
Data

Collaborator 2

Collaborator 2
Data



REDCap Data Sharing

Define each individual user's specific access

Adding new user "eslockhart"

You may set the rights for the user below by checking the boxes next to the application tools to which you wish to grant them access. You may also grant them or deny them access to individual data collection instruments, if so desired. To save your selections, click the "Add user" button at the bottom of the page.

Basic Rights

Expiration Date (D/M/Y)
(if applicable)

Highest level privileges:

- ☐ Project Design and Setup
- ☐ User Rights
- ☐ Data Access Groups

Privileges for data exports (including PDFs and API exports), reports, and stats:

☒ Data Exports
* De-identified means that all free-form text fields will be removed, as well as any date/time fields and Identifier fields.

☐ No Access
☒ De-Identified*
☐ Remove all tagged Identifier fields
☐ Full Data Set

☒ Add / Edit Reports
Also allows user to view ALL reports (but not necessarily all data in the reports)

☒ Stats & Charts

Other privileges:

- ☒ Calendar
- ☐ Data Import Tool
- ☐ Data Comparison Tool

Data Entry Rights

NOTE: The data entry rights *only* pertain to a user's ability to view or edit data on a web page in REDCap (e.g., data entry forms, reports). It has no effect on data imports or data exports.

	No Access	Read Only	View & Edit
Demographics	☒	☐	☐
Baseline Data	☐	☐	☒
Month 1 Data	☐	☒	☐
Month 2 Data	☐	☒	☐
Month 3 Data	☐	☒	☐
Completion Data	☐	☐	☒
Demographics 2	☒	☐	☐

New User Notification

☒ Notify user of their project access via email? ☒

Add user

Cancel



Audit Trail

This module lists all changes made to this project, including data exports, data changes, and the creation or deletion of users.

Filter by event: All event types (excluding page views) ▾

Filter by user name: All users ▾

Filter by record: All records ▾

Displaying events (by most recent): 1 - 12 ▾

Filter by time range from  to 

Time / Date	Username	Action	List of Data Changes OR Fields Exported
06/06/2017 6:11pm	mmwang	Manage/Design	Email survey link to self
06/06/2017 6:11pm	mmwang	Manage/Design	Set up survey
29/05/2017 11:14am	mmwang	Updated User mmtenzer	user = 'mmtenzer'
29/05/2017 11:14am	mmwang	Updated User mmwang	user = 'mmwang'
29/05/2017 11:13am	mmwang	Updated User mmwang	user = 'mmwang'
29/05/2017 11:13am	mmwang	Updated User sjbwondara	user = 'sjbwondara'
29/05/2017 11:13am	mmwang	Updated User sjbwondara	user = 'sjbwondara'
29/05/2017 11:12am	mmwang	Created User mmtenzer	user = 'mmtenzer'
29/05/2017 11:11am	mmwang	Created User sjbwondara	user = 'sjbwondara'
29/05/2017 11:09am	mmwang	Manage/Design	Modify project settings
29/05/2017 11:09am	mmwang	Manage/Design	Modify project settings
29/05/2017 11:08am	mmwang	Manage/Design	Create project using REDCap XML file



Exploring Data through The TriNetX Solution

Health Analytics Research Team



TriNetX Background

- 90+ Healthcare Organizations
- 10 out of 15 top Pharma companies
- 5 out of 10 top CROs (clinical research organization)
- 19 countries
- 14.5 billion facts
- 140M patients



TriNetX “Live” Business Model

- FREE to Healthcare Organizations (HCO)
- HCO contributes data for two main purposes
 - Increase number of incoming clinical trials
 - Utilize for self-service feasibility for investigator initiated trials
 - Data exploration without need for IRB approval
- HCO queries their own de-identified data, with the ability to re-identify a patient list with appropriate internal HCO approvals



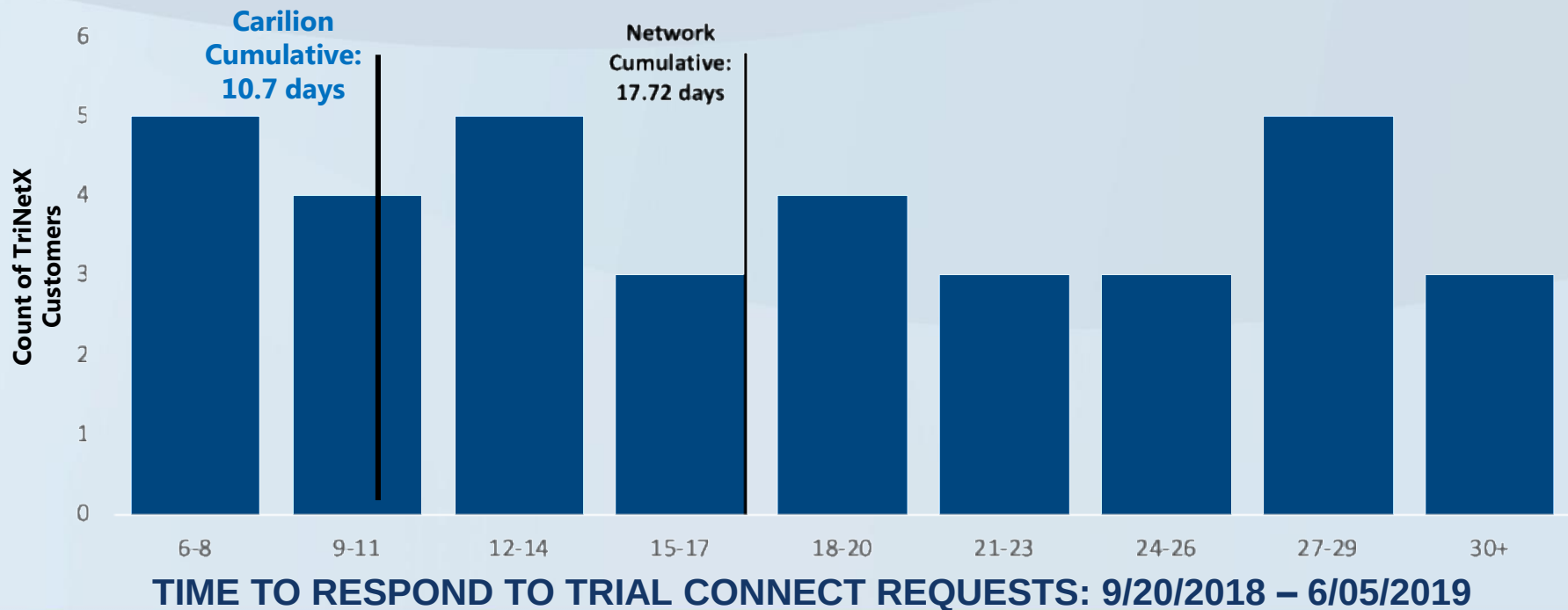
TriNetX “Live” Business Model

- CRO's and Pharma query aggregated data to determine available study populations for clinical research trials
- CRO's and Pharma PAY for this service
- HCOs respond to clinical trial queries either accepting or declining the trial request
- Efficient, consistent, timely



TriNetX “Live” Clinical Trial Statistics

- ~ 2000 trials offered; >600 trials accepted
- Turn around time improvement: Industry average 42 days



TriNetX “Live” Data Set

- Demographics
- Encounters
- Diagnoses
- Procedures
- Labs (numeric and positive/negative)
- Medications (outpatient prescription orders)
- Vitals



TriNetX “Live” Logic

- Search for terms with clear ontologies
 - Can search on “Diabetes” or codes E11.9, 250.*, etc
- Filter for specific values
 - A1C >8
- Apply time logic to events
 - Patients with diabetes with visits between 1/1/2018 and 8/31/2018
- AND / OR logic on every component
 - AND on metformin AFTER visit
 - AND not on metformin BEFORE visit
 - AND Creatinine >=60 BEFORE visit
 - AND patient does not have CHF BEFORE visit
- Can get very complicated. TriNetX provides FREE support



TriNetX “Live” Result

- X number of patients

1,300

PATIENTS

1

HCOS

May 17, 2019, 2:09 pm. Deidre Myers. Carilion Clinic.

Count Patients

All changes saved

Browse Network

Data updated Jun 19, 2019, 10:05 am

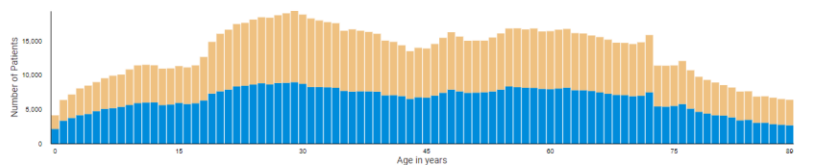
Network Carilion Clinic 1 of 1 HCOs online

Demographics

Range Months (0-36) Years (0-85)

Type of Graph Grouped Stacked

Sex Female Male



Patients Over 90: 64,509

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
1,278,520	0	90	46	24

Sex	Percentage	Race	Percentage
Female	52%	White	75%
Male	48%	Unknown Race	16%
Unknown Gender	0%	Black or African American	8%
		Asian	1%
		American Indian or Alaska Native	0%

Analyze Criteria

Criteria Order Most to least impact All HCOs Exclude HCOs with 0 patients in the cohort

	Patients	HCOS
Network	1,278,520	1
Base Population	7,020 (-99%)	1
Population 18 - 80 years, Any sex	6,240 (-11%)	1
Criteria Not Selected for Analysis	6,240 (0%)	1
Event 2A This event occurred between 1 year and 5 years ago. Must Have: 1007622 Endoscopy Procedures on the Rectum. Event 2B Any instance of Event 2B	1,900 (-70%)	1
Cannot Have: 327361 Adalimumab [key word]	1,580 (-17%)	1
Event 1A This event occurred since 1 year ago. Cannot Have: 1007622 Endoscopy Procedures on the Rectum	1,370 (-13%)	1
Cannot Have: 9063 C reactive protein [Mass/volume] in Serum, Plasma or Blood > 1 mg/L, most recent value	1,320 (-4%)	1
	1,320 Patients	1 HCOS

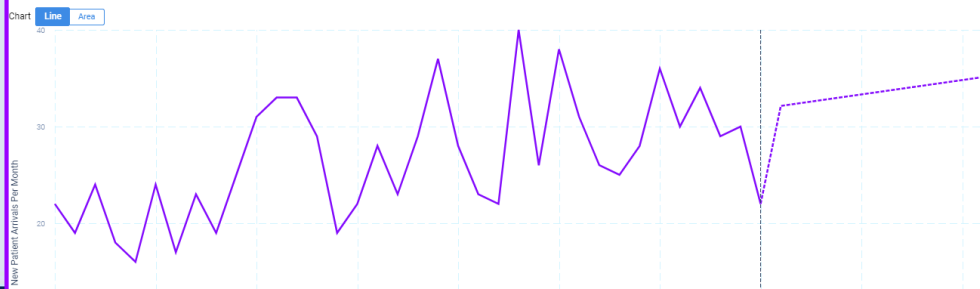
Diagnoses

Diagnoses	Patients	% of Network
R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	743,720	58%
Z00-Z99 Factors influencing health status and contact with health services	707,810	55%
J00-J99 Diseases of the respiratory system	522,190	41%
J00-J06 Acute upper respiratory infections	350,810	27%
J30-J39 Other diseases of upper respiratory tract	217,770	17%
J20-J22 Other acute lower respiratory infections	146,320	11%
J40-J47 Chronic lower respiratory diseases	144,670	11%
J09-J18 Influenza and pneumonia	106,450	8%
J96-J99 Other diseases of the respiratory system	79,620	6%
J90-J94 Other diseases of the pleura	39,040	3%
J80-J84 Other respiratory diseases principally affecting the interstitium	22,570	2%
J60-J70 Lung diseases due to external agents	8,140	1%
J95-J95 Intraoperative and postprocedural complications and disorders of respiratory system, not elsewhere classified (J95)	4,510	0%
J85-J86 Suppurative and necrotic conditions of the lower respiratory tract	4,100	0%
M00-M99 Diseases of the musculoskeletal system and connective tissue	521,910	41%

Patient Arrival Rate

How is this projection calculated?

Graph Show Hide



TriNetX “Live” Result

- Use results to verify sufficient population
- Number is approximate
- De-identified data extracts (free of own data, fee for network-wide)
- Seek approval (R&D, IRB) to re-identify patient list, if needed for study (not necessarily needed for prospective cohorts yet to occur)



TriNetX Advanced Analytics

- Further explore the de-identified cohort
 - Ability to design comparative effectiveness studies by defining one or multiple cohorts
 - Comparisons between groups with outcome analysis and identification of impact weight of other variables
 - Provides statistics on outcomes, including p values, confidence intervals, variable impact, Kaplan-Meier survival curves

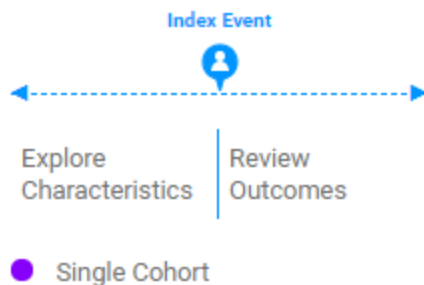


TriNetX Advanced Analytics

Advanced Analytics

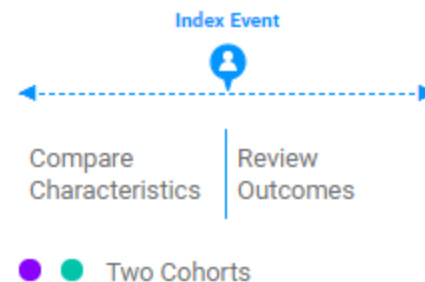
Analyze Outcomes

How do patients in a cohort experience outcomes?



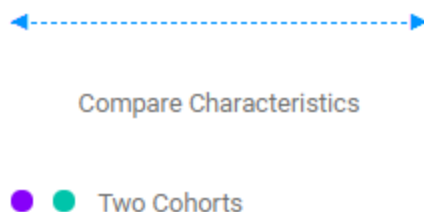
Compare Outcomes

How do outcomes compare between cohorts?



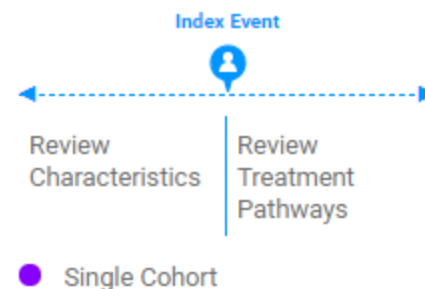
Compare Cohorts

How do patient characteristics compare between cohorts?



Treatment Pathways

In what order do patients receive treatments following a diagnosis?



TriNetX Advanced Analytics

COMPARE COHORTS...

38



TriNetX Advanced Analytics

COMPARE COHORTS...

39



Comparing Cohorts

1

Diabetes A1c > 9%
2,950
Patients

2

Diabetes A1c never > 7%
27,840
Patients

Show Diagnoses Within (Months) [What's this?](#)

☐ 3M ☐ 6M ☐ 12M ☐ 24M ☐ Custom ☒ Anytime

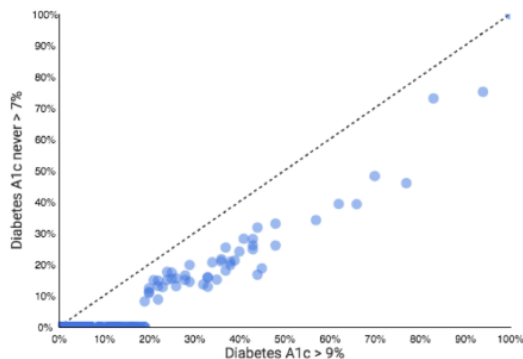
Show [What's this?](#)

☒ All ☐ Acute ☐ Chronic



SCATTER PLOT

9,439 Unique Diagnoses



RESET CHART

BACK



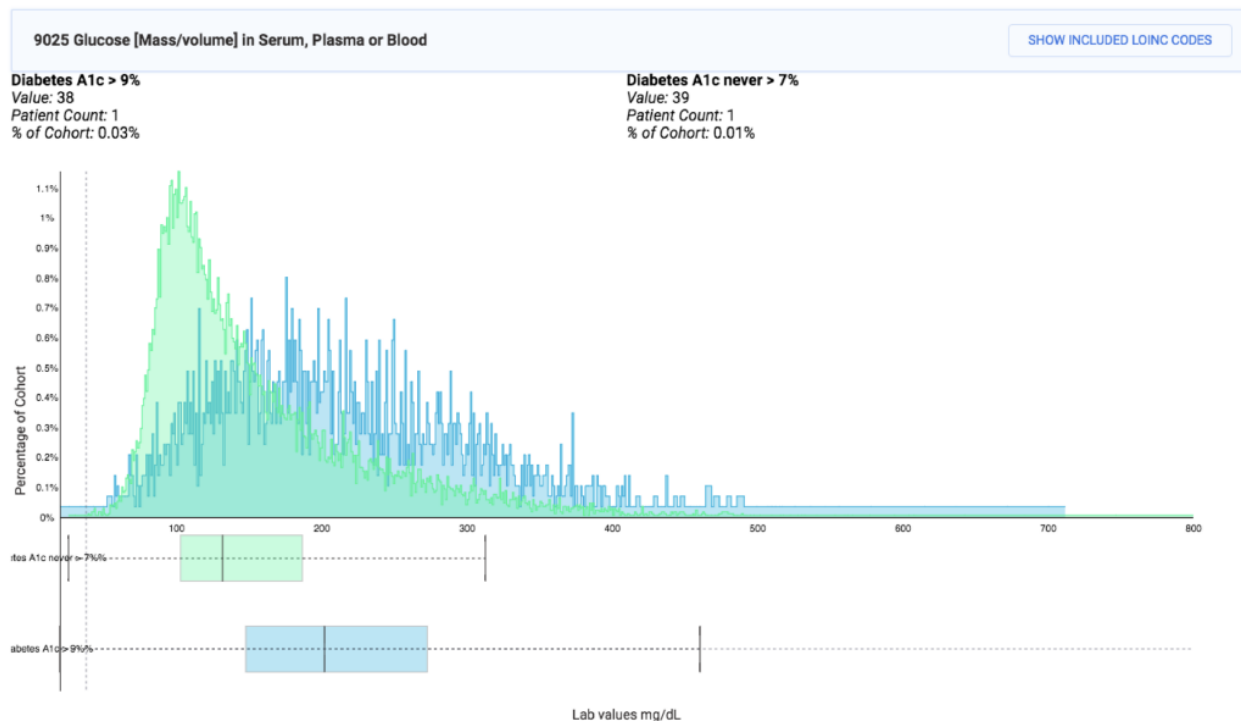
COPYRIGHT © 2017 TRINETX, INC. ALL RIGHTS RESERVED. CONFIDENTIAL



TriNetX Advanced Analytics

COMPARE COHORTS...

40



TriNetX Advanced Analytics

- Ability to adjust for confounding factors

Cohort 1				Outcome no CHF any encounter type				74,990		Cohort 2				Outcome CHF any encounter type				870	
Age		Mean +/- SD		Min		Max		Patients		% of Cohort				Save					
Current Age		61.9 ± 8.26		49		80		74,990		100%									
		67.4 ± 8.47		49		80		870		100%									
Age at Index		55.5 ± 8.28		40		79		74,990		100%									
		63.3 ± 8.66		41		79		870		100%									
Sex								Patients		% of Cohort				Save					
Female								43,710		58%									
								470		54%									
Male								31,280		42%									
								410		47%									

Select Propensity Score Covariates

Select 1+ covariates from your saved characteristics list.

☐ Check All (select at least one to balance cohorts)

☐ I49 Other cardiac arrhythmias	Type Present/Absent	Time Window Anytime before to same day as index event
☐ I48 Atrial fibrillation and flutter	Type Present/Absent	Time Window Anytime before to same day as index event
☐ I60-I69 Cerebrovascular diseases	Type Present/Absent	Time Window Anytime before to same day as index event
☐ I70-I79 Diseases of arteries, arterioles and capillaries	Type Present/Absent	Time Window Anytime before to same day as index event

- Ability to generate balanced cohorts through propensity score matching



TriNetX Advanced Analytics

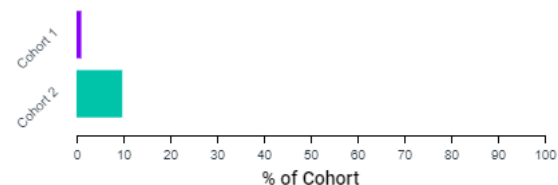
2a : Measures of Association Excluding Patients with Outcome Prior to the Time Window

Cohort	Cohort Statistics			Risk
	Patients in Cohort	Patients with Outcome		
1 No CHF	2,970	30		1.01%
2 CHF	2,970	290		9.764%

Risk Difference				Risk Ratio		Odds Ratio	
Risk Difference	95 % CI	z	p	Risk Ratio	95 % CI	Odds Ratio	95 % CI
-8.754%	(-9.881%,-7.628%)	-14.943	< 0.0001	0.103	(0.071,0.15)	0.094	(0.065,0.138)

20 patients in Cohort 1 and 10 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

Risk of Outcome



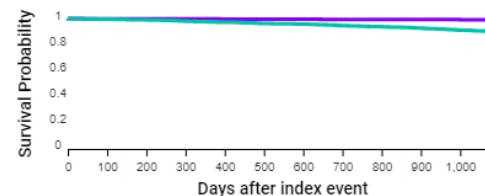
2b : Kaplan-Meier Analysis Excluding Patients with Outcome Prior to the Time Window

Cohort	Cohort Statistics			Survival Probability at End of Time Window
	Patients in Cohort	Patients with Outcome	Median Survival (Days)	
1 No CHF	2,970	30	-	99.17%
2 CHF	2,970	290	-	89.758%

Log-Rank Test		
χ^2	df	p
228.939	1	< 0.0001

20 patients in Cohort 1 and 10 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

Kaplan-Meier Survival Curve

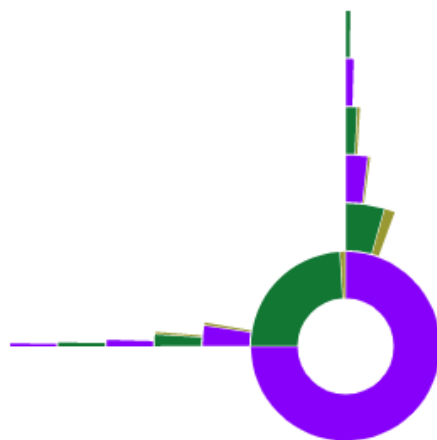


TriNetX Advanced Analytics

Cohort 1

alcohol 5 yr TNX

348,401



- Psychiatry services and procedures
- Alcohol deterrents
- Alcohol deterrents | Psychiatry services and procedures



TriNetX Research Network



Network

Research
24 of 24 HCOs online

...



Population

Any age / Any sex
33,639,584 patients on network



Population

Research

Measurement of Time

Months (0-36)

Years (0-89)

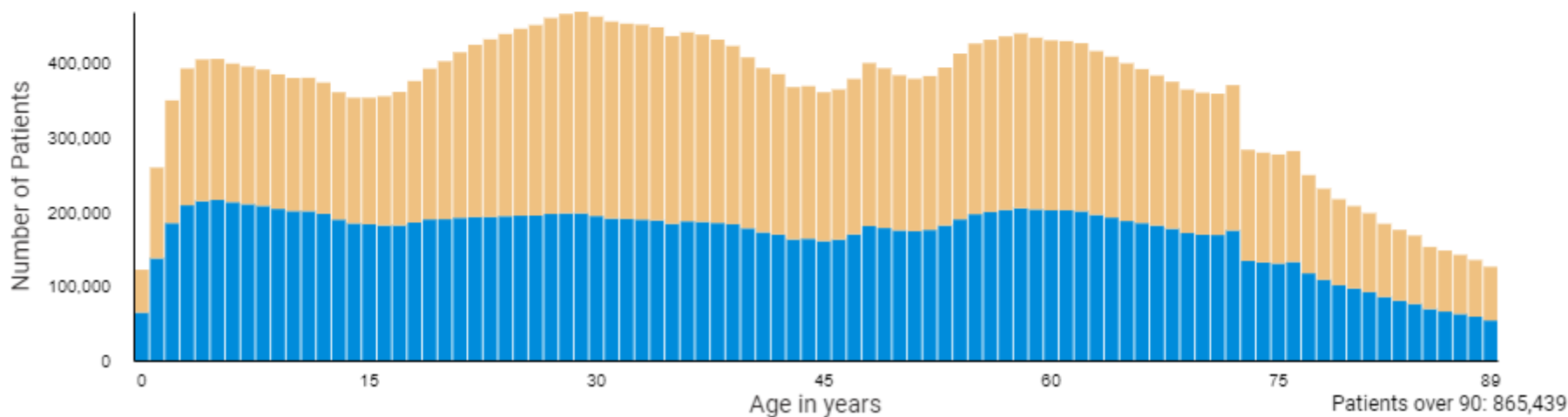
Type of Graph

Grouped

Stacked



Excluded Male Female



Data Management Checklist

- Explore your cohort through TriNetX
- Determine if you need identified or prospective data and its data source
- Determine if data is an extract, chart review, or both (or survey!)
- Determine who will have access, what level of access, and if they need agreements (don't forget the biostatistician!)
- Write up protocol to include Epic research access, REDCap and secure shared drive



Data Management Process

Post IRB and R&D approval, our concierge services will contact you to submit all necessary requests for:

- Secure Shared Drive
- Epic Research Access
- Data Extracts (EPIC)
- Data Collection and Storage (REDCap)
- Data Analysis: Biostatistical Services
- TriNetX Files or Reidentification



HART Drop-In Office Hours

- *CRMH, 1 South Training Room:*
 - Mondays and Wednesdays, 9 a.m. - 5 p.m.
 - Other days by appointment
-
- *3 Riverside, Conference Room 2B:*
 - Mondays and Wednesdays, drop-in for Biostatistics support, 9 a.m. - noon
 - Tuesdays and Thursdays, drop-in for Solution Developer support, 9 a.m. - noon
 - Other days and times by appointment



Contact Information

Mattie Tenzer, Director
mmtenzer@carilionclinic.org

Min Wang, Manager
mmwang@carilionclinic.org

HART@carilionclinic.org

<https://carilionclinic.org/health-analytics-research-team>



FAQ

Q: I want to just keep my data in an Excel spreadsheet. Can I do that?

A: Only if approved by the IRB, which is typically if the dataset is a limited dataset or fully de-identified. Unless it is fully deidentified and not company confidential, excel spreadsheets must be stored in the secure shared drive set up for your project.

Best practice is to use REDCap, and it is required for fully identified data sets. The patient name is allowed to be stored in REDCap, which facilitates data collection. In addition, REDCap offers protections against corruption and loss that Excel does not. It also enforces consistent data collection.



FAQ

Q: How do I get access to TriNetX and what approvals do I need?

A: Email HART@carilionclinic.org for Carilion Clinic TriNetX End User Agreement. Complete the form and have your chair/manager sign. If you are not a Carilion employee, send to Francis Farrell fxfarrell@carilionclinic.org for approval.



FAQ

Q: How do I get my project built into REDCap?
What if I also need a data extract?

A: Once you have IRB approval or exemption, HART is notified. Through our Concierge Services, we will confirm the services you need for your project, enter the requests and assign them out to our team. Our team member will develop the extract and build your REDCap project for you, working closely with you to ensure they meet your project's approved data collection tool. If needed, we will import the extract into REDCap. We will train you on how to use your REDCap project.



FAQ

Q: I think my project is just QA/QI. Do I have to go through all of this?

A: Not necessarily, but it is best practice.

The best QA/QI projects are often similar to research in terms of ensuring sufficient appropriate patient population (TriNetX), data extracts (Epic), data management (REDCap), chart review, and statistics. Often QA/QI projects take as much or more time and resources to accomplish. Therefore, safe storage of data is essential to protect against loss and corruption. When in doubt, consult with the IRB to ensure your project meets QA/QI definitions. In addition, Research takes priority when our workload is too heavy, so your QA/QI project may not get the timely it would if structured as Research.

