Disclaimer

The following report(s) provides findings from an FDA-initiated query using Sentinel. While Sentinel queries may be undertaken to assess potential medical product safety risks, they may also be initiated for various other reasons. Some examples include determining a rate or count of an identified health outcome of interest, examining medical product use, exploring the feasibility of future, more detailed analyses within Sentinel, and seeking to better understand Sentinel capabilities.

Data obtained through Sentinel are intended to complement other types of evidence such as preclinical studies, clinical trials, postmarket studies, and adverse event reports, all of which are used by FDA to inform regulatory decisions regarding medical product safety. The information contained in this report is provided as part of FDA's commitment to place knowledge acquired from Sentinel in the public domain as soon as possible. Any public health actions taken by FDA regarding products involved in Sentinel queries will continue to be communicated through existing channels.

FDA wants to emphasize that the fact that FDA has initiated a query involving a medical product and is reporting findings related to that query does not mean that FDA is suggesting health care practitioners should change their prescribing practices for the medical product or that patients taking the medical product should stop using it. Patients who have questions about the use of an identified medical product should contact their health care practitioners.

The following report contains a description of the request, request specifications, and results from the modular program run(s).

If you are using a web page screen reader and are unable to access this document, please contact the Sentinel Operations Center for assistance at info@sentinelsystem.org.

Overview for Report: cder_iqp_wp021

Request ID: cder_iqp_wp021_tnx

Requester: Center for Drug Evaluation and Research (CDER)

Request Description: In this report, we characterized patients and encounters in the TriNetX Live™ USA Network. This is report three of three. Report one includes data from the past 20 years as of November 3, 2021. Report two includes data from the past five years as of November 18, 2021. Report three includes data from the past year as of December 1, 2021.

Data Source: We ran this query on December 1, 2021. This query contains data from 67 health care organizations (HCOs), provided through the TriNetX Live™ platform in the USA Network from the past year, as of December 1, 2021. We additionally included data from 49 HCOs in the USA Network with Minimal Shift, and data from 36 HCOs in the USA Network with No Shift.

TriNetX aggregates electronic health record (EHR) systems data from its partner HCOs to create queryable datasets. TriNetX datasets are primarily comprised of clinical patient data such as demographics, diagnoses, procedures, labs, and medications. The USA Network with Minimal Shift contains HCOs that date shift their data by 14 or fewer days (including 0). The USA Network with No Shift contains HCOs that do not date shift their data. For more information on the TriNetX Live™ platform and the TriNetX data visit their website here: https://trinetx.com/

Study Design: We identified individuals with any encounter using the Query Builder module in the TriNetX Live™ platform. We characterized the demographic and geographic distribution of the cohorts using the Explore Cohort module. We further utilized the Summary Statistics module to describe the span and number of clinical facts and the fact distribution by type. We repeated each analysis in the entire USA, USA with Minimal Shift, and USA with No Shift Networks.

Events of Interest: Our cohort of interest was defined by the presence of an encounter in the query period in the network of interest. We utilized the Query Builder module to obtain the number of patients with at least one visit term of any type in the prior year (the overall cohort). We additionally counted the subset of patients with at least one visit term of any type in each of the prior two years. We also specifically counted the number of patients with at least one record of the following visit types: ambulatory, emergency, home health, inpatient (defined as inpatient encounter or inpatient non-acute), observation, pre-admission, short stay, virtual, and unknown.

Cohort Eligibility Criteria: Patients with any encounter in the query period were eligible to be included in the cohorts. Patients of all ages were included.

Characteristics of Interest: Our characteristics of interest in this request were on both the patient- and the encounter-level. We utilized the Explore Cohort module to describe the age, sex, race, and ethnicity distribution of the overall cohort in each data network. We utilized the Summary Statistics module to describe the geographic distribution of patients of the overall cohort and the distribution of clinical facts (in a given time span and a given volume of facts) in each data network.

Please see Appendices A and B for the specifications of parameters used in this request.

Overview for Report: cder_iqp_wp021
Limitations: Algorithms and used to define cohorts and outcomes, and mapping of source data to the data model are imperfect and susceptible to misclassification. Additionally, EHR data in the US lacks longitudinality. The information before or after patients' healthcare encounters could be missing, especially if patient care was administered across different HCOs that may or might not participate in the TriNetX USA network. We are unable to determine if absence of evidence of a condition implies a true absence of a condition or if the condition was not observed in the data. Therefore, data should be interpreted with these limitations in mind. All counts provided through the TriNetX Live™ platform are rounded up to the nearest 10 to protect patient privacy. This rounding affects error, especially as sample sizes decrease. Error due to rounding can range from <0.09% when sample sizes are >10,000 to nearly 20% as sample sizes drop. Thus, all estimates should be interpreted as ranges, and small sample sizes should be interpreted with caution. **Notes:** We ran this query on December 1, 2021. A re-run of this query for the same query period in the future may not yield the same results owing to the dynamic nature of the TriNetX Live™ network. Please contact the Sentinel Operations Center (info@sentinelsystem.org) for questions and to provide comments/suggestions for future enhancements to this document. For more information on Sentinel's querying in the TriNetX platform, please refer to the Sentinel Website (https://www.sentinelinitiative.org/methods-data-tools/methods/trinetx-rapid-querying).

Table of Contents	
Glossary	Terms for Analyses Using TriNetX Live™ Platform
Table 1	Patient Counts and Geographic Distribution in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)
Table 2	Number of Patients by Encounter Type in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)
Table 3	Distribution of Clinical Facts in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)
Figure 1a	Demographic Characterization of Patients in the TriNetX USA Network
Figure 1b	Demographic Characterization of Patients in the TriNetX USA Network: Minimal or No Date Shifting (Between 0-14 Days)
Figure 1c	Demographic Characterization of Patients in the TriNetX USA Network: Only HCOs With No Date Shifting
Figure 2a	Geographic Distribution of Patients in the TriNetX USA Network, Overall
Figure 2b	Geographic Distribution of Patients in the TriNetX USA Network: Minimal or No Date Shifting (Between 0-14 Days)
Figure 2c	Geographic Distribution of Patients in the TriNetX USA Network: Only HCOs With No Date Shifting
Appendix A	Specifications Defining Query Builder Parameters in this Request
Appendix B	Specifications Defining Analytics Parameters in this Request

Glossary of Terms for Analyses Using TriNetX Live™ Platform*

Characteristic - A medical fact (e.g., diagnosis, procedure, lab result) that occurred on or before the cohort-defining index event.

Explore Cohort - A description module on the TriNetX platform that presents a clinical profile of patients in a given cohort. Patient counts are rounded up to the nearest 10 before percentages are calculated, so the sum each of the values in one category may not total to 100%.

Date Shifting - A data obfuscation technique that some HCOs use to preserve patient privacy. Date shifting entails assigning each patient a random number of days (eg, -365 to +365 days) and consistently adjusting each of their dates by that number of days, thus maintaining temporal relationships between records within a single patient.

Fact - (Medical Fact) A unit of utilization that represents a medical observation on a patient (e.g., diagnosis, procedure, clinical observation).

Filter - A method of limiting terms included in queries to a specific subset of data. Filters include age at time of event, data source (electronic health record or natural language processing); brand name, route, and strength for medication terms; occurrence (first or most recent) for lab terms; and priority for diagnosis and procedure terms.

Group - A series of codes and terms defined with Boolean logic that are used to create a query cohort. For each group, users have the ability to specified time periods of interest, and the number of instances that the group must occur for cohort entry.

> **Subgroup -** Within a group, additional subgroups can be specified to define temporal relationships between the terms in the subgroup (e.g., terms in subgroup B must occur within 5 days after terms in subgroup A). Users can require that these temporal constraints be applied to the 1) first, 2) last, or 3) any instance of each subgroup.

Health Care Organization (HCO) - Organizations that contribute electronic healthcare record data to the TriNetX data networks. HCOs include academic institutions and community health provider systems and a single HCO may contain one or more individual sites or facilities.

Index - The first date when a patient meets all of the cohort-defining criteria. In Analytics modules, the index can be defined as the date when a patient meets all of the cohort criteria, or only one specific group's criteria.

Module - A subsection of the TriNetX platform that performs a distinct functionality. Cohorts are created using the Query Builder module. Descriptive modules include Healthcare Organizations, Explore Cohorts, Rate of Arrival, Summary Statistics, and Analyze Criteria. Advanced analytic modules include Analyze Outcomes, Compare Outcomes, Compare Cohorts, Treatment Pathways, and Incidence and Prevalence.

Network - An aggregation of HCOs contributing data to the platform. Multiple networks are available for querying on the platform; the different networks represent subsets of HCOs organized by date-shifting practices or availability of downloadable datasets.

Outcome - A medical fact (e.g., diagnosis, procedure, lab result) that occurred on or after the cohort-defining index event.

Query - In the TriNetX platform, a query is a distinct cohort with a unique set of terms and logic. Query cohorts are created using the Query Builder platform module.

Risk - In Advanced Analytics modules, risk refers to the percentage of patients in each cohort with the specified outcome of interest.

Priority - An indication whether the code was the condition that the provider spent the most time evaluating or treating during a visit. Possible values include primary, secondary, or unknown.

Term - The codes used to specify patient cohort criteria in a query. Code options include diagnoses, procedures, medications, labs, demographics, genomics, and visits. Terms can be linked together using and/or Boolean logic. TriNetX also creates terms that group together multiple medical codes into single clinical concepts.

Cannot Have Term - A category of terms within a query group that patients must not have evidence of to be included in the cohort.

Must Have Term - A category of terms within a query group that patients must have evidence of to be included in the cohort.

Time Constraint - used to define time periods of interest for each group within a query. Time constraints can be defined relative to the date the query was run (e.g., any time before today), or defined based on specific dates (e.g., January 1, 2015 to September 30, 2020).

Treatment Pathway - In Advanced Analytics modules, the Treatment Pathways module returns the order in which patients received treatment and the prevalence of treatments, including combination of medications, following an index event.

TriNetX Codes - For commonly used laboratory terms, TriNetX aggregates Logical Observation Identifiers Names and Codes (LOINC) laboratory codes at a clinically significant level to new queryable TNX:LAB terms.

Visit - A type of term used to specify the type of medical encounter or facility where the encounter was recorded. Visit terms are derived by TriNetX from the source data. Visits are recorded separately from the codes or labs that occurred during the encounter; care settings are not attached to individual codes. Values for visit terms include: ambulatory, emergency, field, home health, inpatient encounter, inpatient acute, inpatient non-acute, laboratory, observation, pharmacy, pre-admission, short stay, virtual, and unknown.

*all terms may not be used in this report

Table 1. Patient Counts and Geographic Distribution in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)

Characteristic	USA Network					
	Overall		Only HCOs with minimal or no date shifting (between 0-14 days)		Only HCOs with no date shifting	
	Number	Percent	Number	Percent	Number	Percent
Patient Counts						
Any Encounter	34,315,520	100	23,726,690	100	18,028,720	100
Encounter in Both of the Prior Two Years	21,375,070	62.29	14,790,660	62.34	11,243,500	62.36
Death Record	253,140	0.74	204,640	0.86	151,180	0.84
Geographic Distribution						
New England	2,828,410	8.24	1,314,320	5.54	1,314,320	7.29
Middle Atlantic	7,398,460	21.56	6,293,480	26.52	4,916,780	27.27
East North Central	4,175,940	12.17	3,691,280	15.56	2,274,880	12.62
West North Central	1,742,780	5.08	399,500	1.68	0	0.00
South Atlantic	7,796,080	22.72	4,699,650	19.81	3,454,580	19.16
East South Central	1,405,380	4.10	1,405,380	5.92	997,640	5.53
West South Central	5,204,200	15.17	3,201,530	13.49	3,201,530	17.76
Mountain	2,929,420	8.54	1,904,380	8.03	1,692,030	9.39
Pacific	770,740	2.25	177,000	0.75	177,000	0.98
Unknown	64,110	0.19	640,170	2.70	0	0.00

NOTE: All counts provided through the TriNetX Live™ platform are rounded up to the nearest 10 to protect patient privacy. Thus, all estimates should be interpreted as ranges, with the lower value of the range ≤ 0.09% less than the presented value unless otherwise noted.

Table 2. Number of Patients by Encounter Type in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)

Characteristic	USA Network					
	Overall		Only HCOs with minimal or no date shifting (between 0-14 days)		Only HCOs with no date shifting	
	Number	Percent	Number	Percent	Number	Percent
Number of Patients with Any Encounter	34,315,520	100	23,726,690	100	18,028,720	100
Number of Patients with at Least One:						
Ambulatory Encounter	26,008,780	75.79	18,184,270	76.64	14,734,090	81.73
Emergency Encounter	5,101,070	14.87	3,818,370	16.09	3,063,950	16.99
Home Health Encounter	368,770	1.07	304800	1.28	0	0.00
Inpatient Encounter	2,803,120	8.17	1,939,730	8.18	1,587,120	8.80
Observation Encounter	384,200	1.12	370,550	1.56	346,120	1.92
Pre-Admission Encounter	8,980	0.03	8,930	0.04	5,300	0.03
Short Stay Encounter	158,450	0.46	99,900	0.42	76,970	0.43
Virtual Encounter	578,990	1.69	165,920	0.70	60,890	0.34
Unknown Encounter Type	16,381,890	47.74	11,184,650	47.14	7,639,670	42.37

NOTE: All counts provided through the TriNetX Live™ platform are rounded up to the nearest 10 to protect patient privacy. Thus, all estimates should be interpreted as ranges, with the lower value of the range ≤ 0.09% less than the presented value unless otherwise noted.

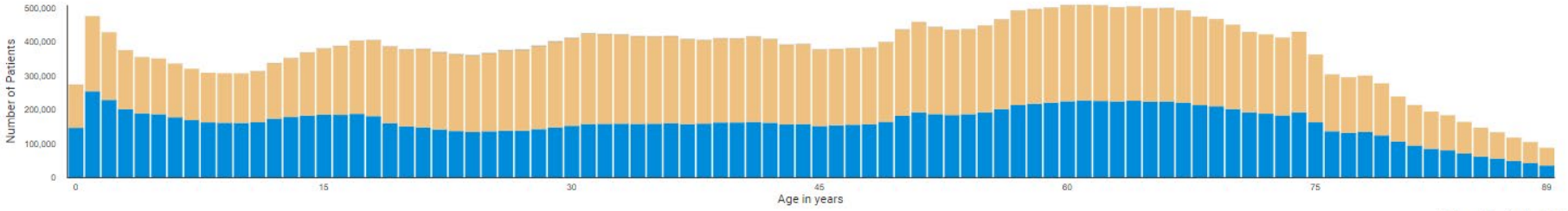
Table 3. Distribution of Clinical Facts in the TriNetX USA Networks, in the Past Year (as of December 1, 2021)

Characteristic	USA Network					
	Overall		Only HCOs with minimal or no date shifting (between 0-14 days)		Only HCOs with no date shifting	
	Number	Percent	Number	Percent	Number	Percent
Number of Patients with Any Encounter	34,315,520	100	23,726,690	100	18,028,720	100
Patients Having a Given Span of Clinical Facts						
At least 0 months[1]	34,108,970	99.40	23,591,400	99.43	17,928,540	99.44
At least 3 months	27,812,120	81.05	19,247,080	81.12	14,571,680	80.82
At least 6 months	26,767,170	78.00	18,510,870	78.02	13,998,770	77.65
At least 9 months	25,869,610	75.39	17,883,420	75.37	13,510,420	74.94
At least 12 months	25,048,370	72.99	17,334,340	73.06	13,075,800	72.53
At least 15 months	24,269,220	70.72	16,797,350	70.80	12,649,070	70.16
At least 18 months	23,622,400	68.84	16,343,290	68.88	12,290,610	68.17
At least 21 months	22,997,710	67.02	15,890,930	66.97	11,939,070	66.22
At least 24 months	22,372,900	65.20	15,437,880	65.07	11,588,310	64.28
At least 36 months	20,034,820	58.38	13,733,770	57.88	10,336,930	57.34
At least 48 months	17,596,210	51.28	11,941,060	50.33	9,059,970	50.25
At least 60 months	15,046,050	43.85	10,010,360	42.19	7,445,280	41.30
Patients Having A Given Volume of Clinical Facts						
One or more clinical fact	34,108,970	99.40	23,591,400	99.43	17,928,540	99.44
20 or more clinical facts	28,611,920	83.38	19,965,900	84.15	15,093,430	83.72
40 or more clinical facts	26,233,820	76.45	18,357,580	77.37	13,894,530	77.07
60 or more clinical facts	24,498,700	71.39	17,172,250	72.38	12,995,890	72.08
80 or more clinical facts	23,004,830	67.04	16,147,460	68.06	12,230,620	67.84
100 or more clinical facts	21,713,130	63.27	15,254,830	64.29	11,569,480	64.17
120 or more clinical facts	20,596,090	60.02	14,484,180	61.05	10,995,750	60.99
140 or more clinical facts	19,612,740	57.15	13,803,910	58.18	10,485,680	58.16
160 or more clinical facts	18,740,000	54.61	13,197,690	55.62	10,030,790	55.64
180 or more clinical facts	17,951,480	52.31	12,649,170	53.31	9,619,560	53.36
200 or more clinical facts	17,232,450	50.22	12,146,700	51.19	9,241,650	51.26
Average Number of Clinical Facts per Patient, by Type						
Diagnoses	110	N/A	105	N/A	100	N/A
Procedures	61	N/A	61	N/A	55	N/A
Medications	343	N/A	321	N/A	352	N/A
Labs	225	N/A	226	N/A	215	N/A
Vital Signs	83	N/A	86	N/A	74	N/A
Genomic Data	<1	N/A	<1	N/A	<1	N/A
Oncology	<1	N/A	<1	N/A	<1	N/A
Chemotherapy Lines	<1	N/A	<1	N/A	<1	N/A

NOTE: All counts provided through the TriNetX Live™ platform are rounded up to the nearest 10 to protect patient privacy. Thus, all estimates should be interpreted as ranges, with the lower value of the range ≤ 0.09% less than the presented value unless otherwise noted.

[1] Patients without a span of clinical facts are patients with at least one encounter but without any clinical facts associated with that encounter.

Figure 1a. Demographic Characterization of Patients in the TriNetX USA Network



Patients 90 and Older: 383,063

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
34,315,520	0	90	44	24

Sex

Female	57%
Male	43%
Unknown	0%

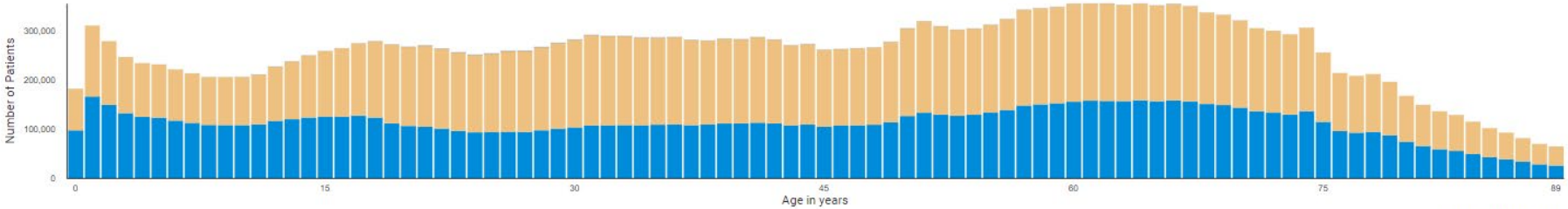
Ethnicity

Not Hispanic or Latino	64%
Unknown Ethnicity	26%
Hispanic or Latino	10%

Race

White	61%
Unknown Race	21%
Black or African American	15%
Asian	3%
American Indian or Alaska Native	0%
Native Hawaiian or Other Pacific Islander	0%

Figure 1b. Demographic Characterization of Patients in the TriNetX USA Network: Minimal or No Date Shifting (Between 0-14 Days)



Patients 90 and Older: 286,804

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
23,726,690	0	90	44	24

Sex

Female	57%
Male	43%
Unknown	0%

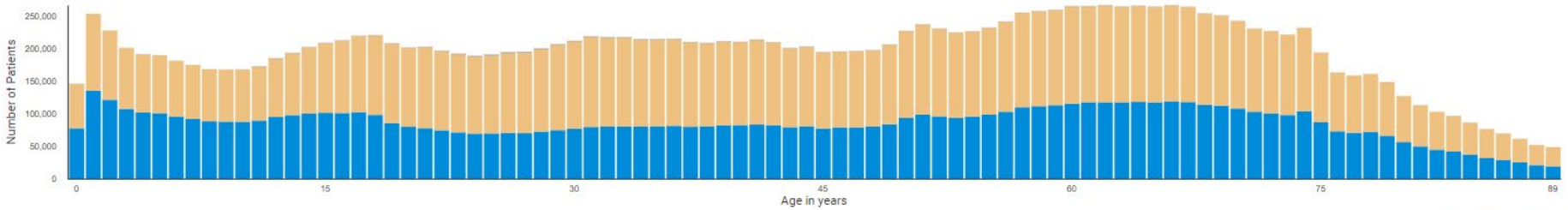
Ethnicity

Not Hispanic or Latino	66%
Unknown Ethnicity	27%
Hispanic or Latino	7%

Race

White	61%
Unknown Race	20%
Black or African American	15%
Asian	3%
American Indian or Alaska Native	1%
Native Hawaiian or Other Pacific Islander	0%

Figure 1c. Demographic Characterization of Patients in the TriNetX USA Network: Only HCOs With No Date Shifting



Patients 90 and Older: 212,194

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
18,028,720	0	90	43	24

Sex

Female	57%
Male	43%
Unknown	0%

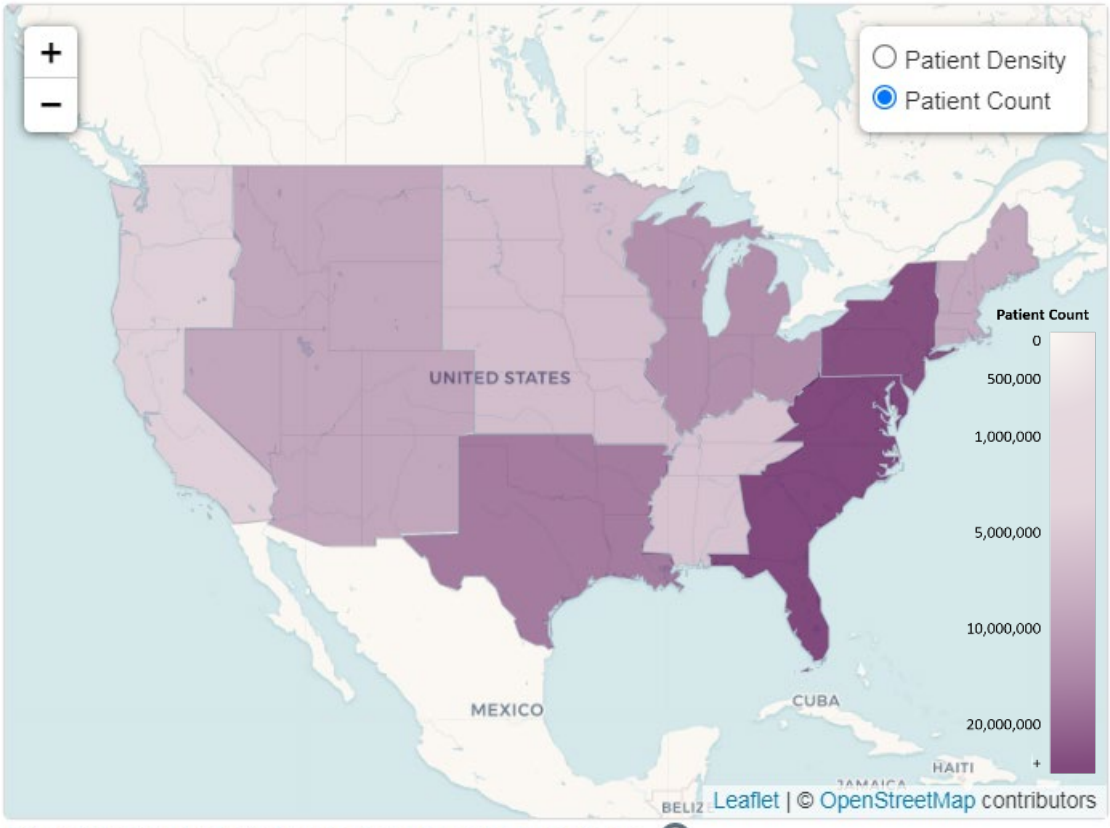
Ethnicity

Not Hispanic or Latino	66%
Unknown Ethnicity	27%
Hispanic or Latino	7%

Race

White	62%
Unknown Race	20%
Black or African American	15%
Asian	3%
American Indian or Alaska Native	0%
Native Hawaiian or Other Pacific Islander	0%

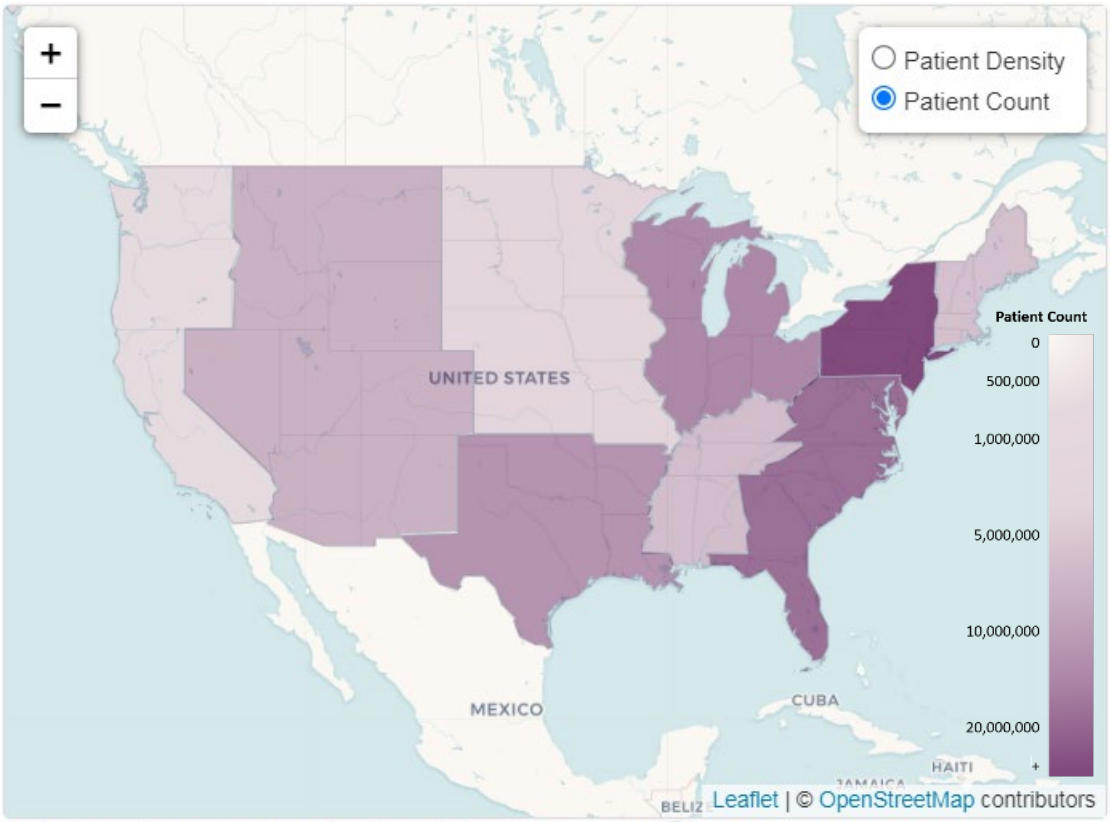
Figure 2a. Geographic Distribution of Patients in the TriNetX USA Network, Overall



Patient counts can be obfuscated to protect patient privacy

Patient location is determined by location of HCO headquarters

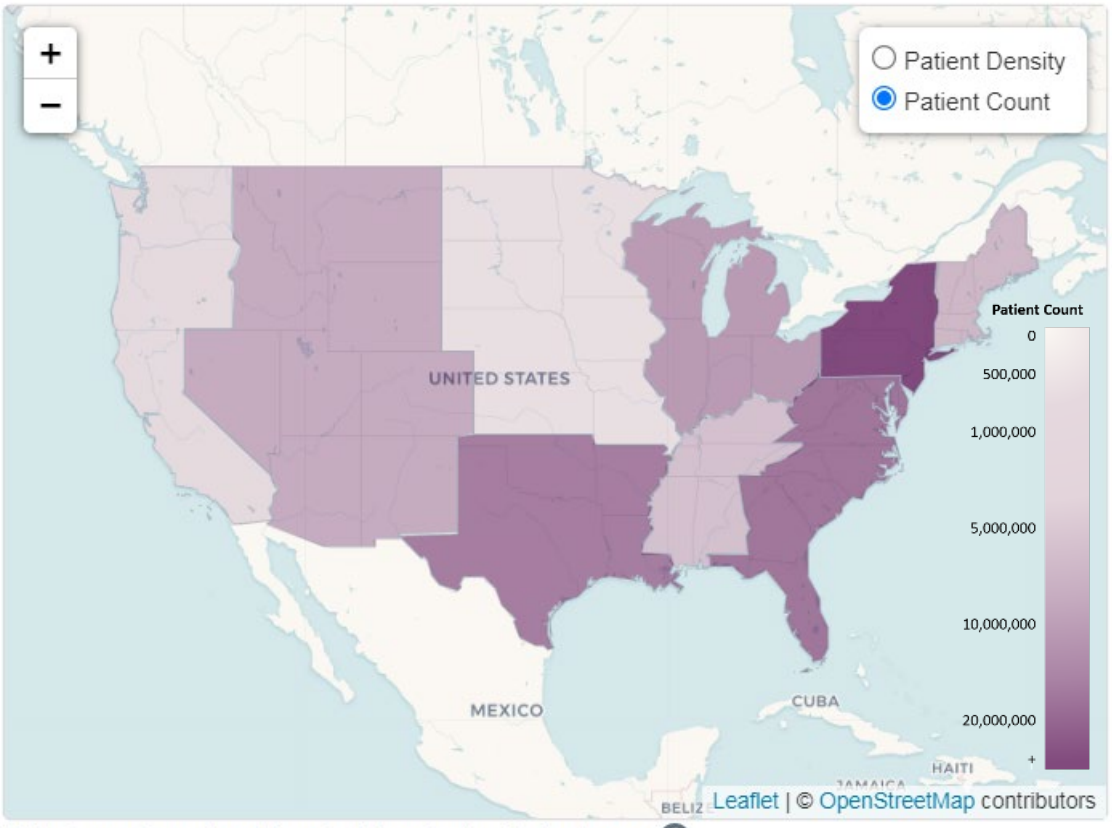
Figure 2b. Geographic Distribution of Patients in the TriNetX USA Network: Minimal or No Date Shifting (Between 0-14 Days)



Patient counts can be obfuscated to protect patient privacy

Patient location is determined by location of HCO headquarters

Figure 2c. Geographic Distribution of Patients in the TriNetX USA Network: Only HCOs With No Date Shifting



Patient counts can be obfuscated to protect patient privacy

Patient location is determined by location of HCO headquarters

Appendix A. Specifications Defining Query Builder Parameters in this Request

All cohorts in each stratification will be executed three times - once on each of the three available USA Networks:

A: All HCOs, **B:** Only HCOs with minimal or no date shifting (between 0-14 days), **C:** Only HCOs with no date shifting

Stratification C: Past Year

Cohort A1. Total Number with any Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit	
Cohort A3. Number with an Encounter in Each of the Prior Two Years	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit	
Group 2: Time Constraint = Between One and Two Years Ago	
Must Have:	*Cannot Have:*
Visit	
Cohort A4. Number with Death Record	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit Deceased	
Cohort A5. Ambulatory Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Ambulatory	
Cohort A6. Emergency Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Emergency	
Cohort A7. Home Health Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Home Health	
Cohort A8. Inpatient Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Inpatient Encounter OR Visit: Inpatient Non-acute	

Appendix A. Specifications Defining Query Builder Parameters in this Request

Cohort A9. Observation Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Observation Encounter	
Cohort A10. Pre-admission Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Pre-admission	
Cohort A11. Short Stay Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Short Stay	
Cohort A12. Unknown Encounter Type	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Unknown	
Cohort A13. Virtual Encounter	
Group 1: Time Constraint = Past Year	
Must Have:	*Cannot Have:*
Visit: Virtual	

Appendix B. Specifications Defining Analytics Parameters in this Request

	Description	Analysis Type	Cohorts[1]	Module	Window
1a.	Age distribution[2]	Explore Cohorts	A1, B1, C1	Demographics	[0,0] days
1b.	Racial distribution	Explore Cohorts	A1, B1, C1	Demographics	[0,0] days
1c.	Ethnic distribution	Explore Cohorts	A1, B1, C1	Demographics	[0,0] days
1d.	Sex distribution	Explore Cohorts	A1, B1, C1	Demographics	[0,0] days
1e.	Geographic distribution	Summary Statistics	A1, B1, C1	Geographic Distribution	[0,0] days
1f.	Span of clinical facts	Summary Statistics	A1, B1, C1	Length of Pt. Record	All available history
1g.	Number of clinical facts	Summary Statistics	A1, B1, C1	Fact Distribution	All available history
1h.	Fact distribution by type	Summary Statistics	A1, B1, C1	Facts by Type	All available history

[1]Each analysis will be performed three times (once for each stratification) for each of the three databases queried - for a total of 9 per analysis.

[2]Ages are rounded to the nearest year