# The TriNetX Linked EHR and Administrative Claims Network

Sarah Malek, MPPA[1], sarah.malek@trinetx.com, Michael Swartzbaugh, BS[1], michael.swartzbaugh@trinetx.com, Ellen Stein, MS[1], ellen.stein@trinetx.com,
Jeffrey Brown, PhD[1], jeff.brown@trinetx.com
[1]TriNetX, LLC., Cambridge, MA, USA.

## OBJECTIVES

Data resources that combine the clinical richness of Electronic Health Records (EHR) with administrative claims data have long been considered the gold standard for health economics and outcomes (HEOR) and other real-world data research. However, there are few well-characterized integrated EHR and claims systems widely available for research. TriNetX seeks to fill that void with creation of the TriNetX Linked EHR and Claims Network.

## METHODS

TriNetX's Dataworks-USA is a federated global research network that includes 64 US-based HealthCare Organizations (HCOs), contributing EHR data for over 112 million patients. The de-identified EHR data is sourced directly from the HCOs and mapped to standardized terminologies. A subset of the Dataworks-USA HCOs permit linking to external data sources through Datavant tokenization. The EHR data has been linked to closed claims and mortality data to create the TriNetX Linked Network.

## RESULTS

As of April 2024, the Network includes 14.0 million patients with 23.4 million unique medical enrollment periods. The Network is 51% female, 88% with a known race (61% white, 18% Black, 3% Asian, 4% other), 59% with a known ethnicity (49% not Hispanic or Latino, 10% Hispanic or Latino). The mean age is 43 years, with 14.3% under 18 and 18.9% 65+. Insurance types are Medicaid only (30%), Medicare only (6%), commercial insurance only (46%), and mixed coverage (17%). There are 6.3 million enrollment periods over 2 years, and 4.7 million are 3+ years (60% and 67%, respectively, contain at least one EHR encounter). In addition to the clinical depth of EHR data that includes laboratory results, prescriptions, medications administered, and vital signs, the claims data provides complementary data to the patient records including claim type, payment amounts, provider specialty, medication dispensed (NDC), quantity dispensed, and days supply.

### TABLE 1 . Baseline Characteristics

| Characteristics | All  TriNetX Linked n=13,981,103 | | Use Case (semaglutide or tirzepatide) n=220,239 | |
|---|---|---|---|---|
| | N or (Mean) | % or Std Dev | N or (Mean) | % or Std Dev |
| **Sex** | | | | |
| Male | 6,843,750 | 49.0% | 72,304 | 32.8% |
| Female | 7,137,353 | 51.1% | 147,935 | 67.2% |
| **Race** | | | | |
| White | 8,555,037 | 61.2% | 140,953 | 64.0% |
| Black or African American | 2,562,736 | 18.3% | 46,647 | 21.2% |
| Asian | 454,386 | 3.3% | 4,581 | 2.1% |
| Native Hawaiian or Other Pacific Islander | 62,915 | 0.5% | 947 | 0.4% |
| American Indian or Alaskan Native | 36,351 | 0.3% | 749 | 0.3% |
| Other | 609,576 | 4.4% | 8,017 | 3.6% |
| Unknown | 1,700,102 | 12.2% | 18,346 | 8.3% |
| **Ethnicity** | | | | |
| Not Hispanic or Latino | 6,915,054 | 49.5% | 117,145 | 53.2% |
| Hispanic or Latino | 1,367,352 | 9.8% | 17,773 | 8.1% |
| Unknown | 5,698,698 | 40.8% | 85,321 | 38.7% |
| **Age - Current** | | | | |
| ALL | 43 | 22 | 53 | 13 |
| 0 to 17 | 2,005,294 | 14.3% | 326 | 0.1% |
| 18 to 24 | 1,268,524 | 9.1% | 3,157 | 1.4% |
| 25 to 44 | 4,453,163 | 31.9% | 55,183 | 25.1% |
| 45 to 64 | 3,617,633 | 25.9% | 121,408 | 55.1% |
| 65+ | 2,636,489 | 18.9% | 40,165 | 18.2% |
| **Insurance** | | | | |
| Commercial | 6,500,489 | 46.5% | 116,472 | 52.9% |
| Managed Medicaid | 4,194,243 | 30.0% | 47,832 | 21.7% |
| Medicare Advantage | 860,286 | 6.2% | 10,953 | 5.0% |
| Multiple | 2,425,088 | 17.2% | 44,971 | 20.4% |
| Unknown | 997 | 0.1% | 11 | 0.0% |
| **Enrollment*** | | | | |
| < 1 year continuous enrollment | 7,815,829 | 33.4% | 48,436 | 20.4% |
| 1 to 2 years continuous enrollment | 5,098,524 | 21.8% | 43,501 | 18.3% |
| 2 to 3 years continuous enrollment | 3,382,432 | 14.5% | 36,156 | 15.2% |
| > 3 years continuous enrollment | 7,087,919 | 30.3% | 109,427 | 46.1% |
| **BMI**** | | | | |
| ALL | 29.3 | 7.2 | 35.8 | 6.5 |
| 0 to 18.49 - Under Weight | 378,419 | 6.7% | 301 | 0.2% |
| 18.5 to 24.9 - Normal | 1,645,679 | 29.3% | 6,528 | 5.0% |
| 25 to 29.9 - Overweight | 1,607,554 | 28.6% | 22,427 | 17.1% |
| 30 to 39.9 - Obese | 1,581,422 | 28.1% | 66,155 | 50.3% |
| 40 to 50 - Severe Obesity | 408,178 | 7.3% | 36,053 | 27.4% |

* Enrollments numbers and percentages are based on number of unique patient medical insurance enrollment periods.  There are 23.4M medical enrollment periods across 14.0M patients in February 2024.  Of those, 10.6M medical enrollment periods have at least 1 EHR encounter.

** BMI numbers and percentages are based on the total number of patients that have a BMI in the Linked Network on each respective medication, based on their most recent BMI reading (5.6M).

### TABLE 2. Use Case (semaglutide or tirzepatide)

| | Semaglutide* 199,512 | | | | Tirzepatide: 39,770 | | | |
|---|---|---|---|---|---|---|---|---|
| | n | % | Mean | Std Dev | n | % | Mean | Std Dev |
| **Comorbidities**** | | | | | | | | |
| Diabetes | 144,369 | 72.4% | - | - | 30,386 | 76.4% | - | - |
| Hypertension | 155,732 | 78.1% | - | - | 30,981 | 77.9% | - | - |
| Obesity | 92,337 | 46.3% | - | - | 20,237 | 50.9% | - | - |
| **Lipid Profile**** | | | | | | | | |
| Cholesterol (mg/dL) | 92,964 | 46.6% | 173.6 | 48.9 | 20,133 | 50.6% | 176.3 | 48.7 |
| HDL (mg/dL) | 98,587 | 49.4% | 44.1 | 103.1 | 22,171 | 55.7% | 45.3 | 15.0 |
| LDL (mg/dL) | 97,886 | 49.1% | 97.1 | 40.0 | 22,113 | 55.6% | 100.2 | 39.4 |
| VLDL (mg/dL) | 55,500 | 27.8% | 33.7 | 21.4 | 12,064 | 30.3% | 33.9 | 22.4 |
| Triglycerides (mg/dL) | 98,115 | 49.2% | 182.3 | 189.3 | 21,958 | 55.2% | 185.7 | 207.0 |
| Lipoprotein a (mg/dL) | 553 | 0.3% | 80.9 | 80.4 | 114 | 0.3% | 48.5 | 48.1 |
| **Other Labs**** | | | | | | | | |
| A1c (%) | 98,638 | 49.4% | 7.4 | 1.9 | 22,271 | 56.0% | 7.3 | 1.9 |
| Blood-Glucose (mg/dL) | 129,862 | 65.1% | 165.6 | 78.6 | 27,806 | 69.9% | 167.3 | 81.5 |
| Fasting Glucose (mg/dL) | 6,229 | 3.1% | 122.9 | 51.7 | 1,678 | 4.2% | 121.3 | 52.7 |
| GFR < 60 (mL/min/(1.73_m2)) | 38,139 | 19.1% | 38.1 | 17.2 | 7,699 | 19.4% | 37.9 | 16.6 |
| GFR > = 60 AND <= 200 (mL/min/(1.73_m2)) | 119,181 | 59.7% | 99.2 | 14.7 | 25,557 | 64.3% | 97.9 | 25.7 |

*The number of patients that prescribed or dispensed semaglutide or tirzepatide will not add up to the total 220,239 patients, because there are instances of patients taking both drugs.

**Comorbidities:
Diabetes: Patients must have at least one diabetes related fact, such as ICD-10- CM code E11 (Type 2 diabetes mellitus) or a laboratory value for hemoglobin A1c greater than 6.5%.
Hypertension: Patients must have at least one hypertension related fact, such as ICD-10-CM code E78 (Disorders of lipoprotein metabolism and other lipidemias) or a laboratory value for cholesterol greater than 240.1 mg/dL or cholesterol in LDL greater than 160.1 mg/dL.
Obesity: Patients must have at least one obesity lab result, such as BMI (LOINC: 39156-5) or height (LOINC: 8302-2) and weight (3141-9) or related codes to either represent or calculate BMI.
***Lab test LOINC codes: Cholesterol (2093-3, 48620-9); HDL (18263-4, 2085-9, 49130-8) ; LDL (13457-7, 2089-1,18261-8, 49132-4, 18262-6); VLDL ( 13458-5, 2091-7, 49133-2); Triglycerides ( 12951-0, 3043-7, 2571-8); Lipoprotein a: (35388-8, 10835-7); A1c (4548-4,4549-2,17655-8,17856-6); Blood-Glucose (2340-8, 2339-0, 74774-1, 2341-6, 41653-7, 2345-7, 32016-8); Fasting Glucose (1558-6) GFR (50044-7, 77147-7, 69405-9, 70969-1, 48643-1, 48642-3, 76633-7, 33914-3).

## USE CASE

To describe possible uses of the TriNetX Linked Network, we examined the demographic and clinical characteristics of all patients in the Network who were prescribed or who filled a prescription for semaglutide and tirzepatide from January 1, 2022 through February 29, 2024. During the study period, there were over 220,000 unique patients prescribed or dispensed semaglutide and tirzepatide. Approximately two-thirds were female, the majority were white (64%) and non-Hispanic or Latino (53%). Most patients were 45-64 years old (55%) followed by 25-44 years (25%). Commercial insurance was the most common type (53%), followed by Medicaid (22%). Almost half (46%) of the observed enrollment periods were 3+ years of continuous enrollment with at least 1 EHR encounter documented. Among both semaglutide and tirzepatide patients, most had documented diabetes (72% and 76%, respectively), and hypertension (78% for both). The percentage of patients with lipid panel and other lab results available, and the findings, are similar between the two groups.

## CONCLUSION

The demographic characteristics in our use case are similar to other studies examining patients receiving glucagon-like peptide-1 (GLP-1) agonists, though our population is slightly more white, less Hispanic, and more female. Our use case, however, did not examine patients by indication, meaning it may not be directly comparable to studies examining patients receiving those medications for diabetes or weight loss treatments, specifically. Overall, this use case demonstrates that the TriNetX Linked Network combines EHR data with administrative claims providing a clinically rich, longitudinal, data network ideal for conducting HEOR research.

**Download Poster PDF**