Appendix A. Detailed Quantitative Methodology

This appendix provides a detailed explanation of the quantitative methods used in this study. This was a retrospective observational study using administrative Medicaid claims data.

Medicaid data obtainment process

To access Medicaid data at the beneficiary level, researchers must first submit a draft application packet to ResDAC for review that includes a detailed description of the proposed research project, specific data files and years of data being requested, IRB documentation, and a data use agreement. ResDAC reviews submitted packets and provides feedback and recommendations to researchers to ensure that the documentation meets CMS requirements. ResDAC also generates a cost invoice for the research team to review and approve. Researchers and university support staff complete all required edits and send the final, signed versions of all required materials in the application packet to ResDAC, who then submits the packet to the CMS Privacy Board for review. After CMS Privacy Board approval, researchers submit payment for their data on the U.S. government payment website. The data is then processed according to the specifications of the request, which takes around one month. We accessed this data through the Virtual Research Data Center (VRDC), which is a virtual, cloud-based research environment through which researchers can remotely and securely access data and conduct analyses. The VRDC is administered by CMS's Chronic Conditions Warehouse (CCW). Results of analyses conducted in the VRDC can be exported from the VRDC after a data privacy and security review by CCW staff.

Data sources

Our study uses the Medicaid Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAF) Research Identifiable Files (RIF) for 2017–2019. We draw from the Demographics and Eligibility (DE), Inpatient (IP), Other Services (OT), and the Medicaid Enrollee Supplemental File: National Death Index (NDI). The DE, IP, and OT files provide beneficiary demographic characteristics, Medicaid eligibility information, and all adjudicated healthcare claims, including diagnosis and procedure codes. The NDI files were used to identify all-cause and cause-specific mortality; these files include only beneficiaries who have a date of death reported in the annual Medicaid enrollment file. Our data was a "custom cohort" subset by the CCW to only include Medicaid beneficiaries whose enrollment information was submitted by Louisiana (versus other states).

In addition to Medicaid claims data, we used publicly available environmental datasets to construct exposure variables related to industrial pollution. Locations of petrochemical facilities and petroleum refineries in 2017 were obtained from the Louisiana Department of Environmental Quality (LDEQ) through a public records request and classified using North American Industry Classification System (NAICS) codes. Estimates

of toxicity-weighted air pollutant concentrations were obtained from the 2016 U.S. Environmental Protection Agency's Risk-Screening Environmental Indicators (RSEI) geographic microdata. RSEI combines data from the Toxics Release Inventory (TRI) with information on chemical toxicity and atmospheric dispersion modeling to estimate relative levels of potential human exposure to industrial air pollutants. These environmental datasets were linked to Medicaid beneficiary records using ZIP code tabulation areas (ZCTAs), as described in the sections below.

Creation of analytic sample

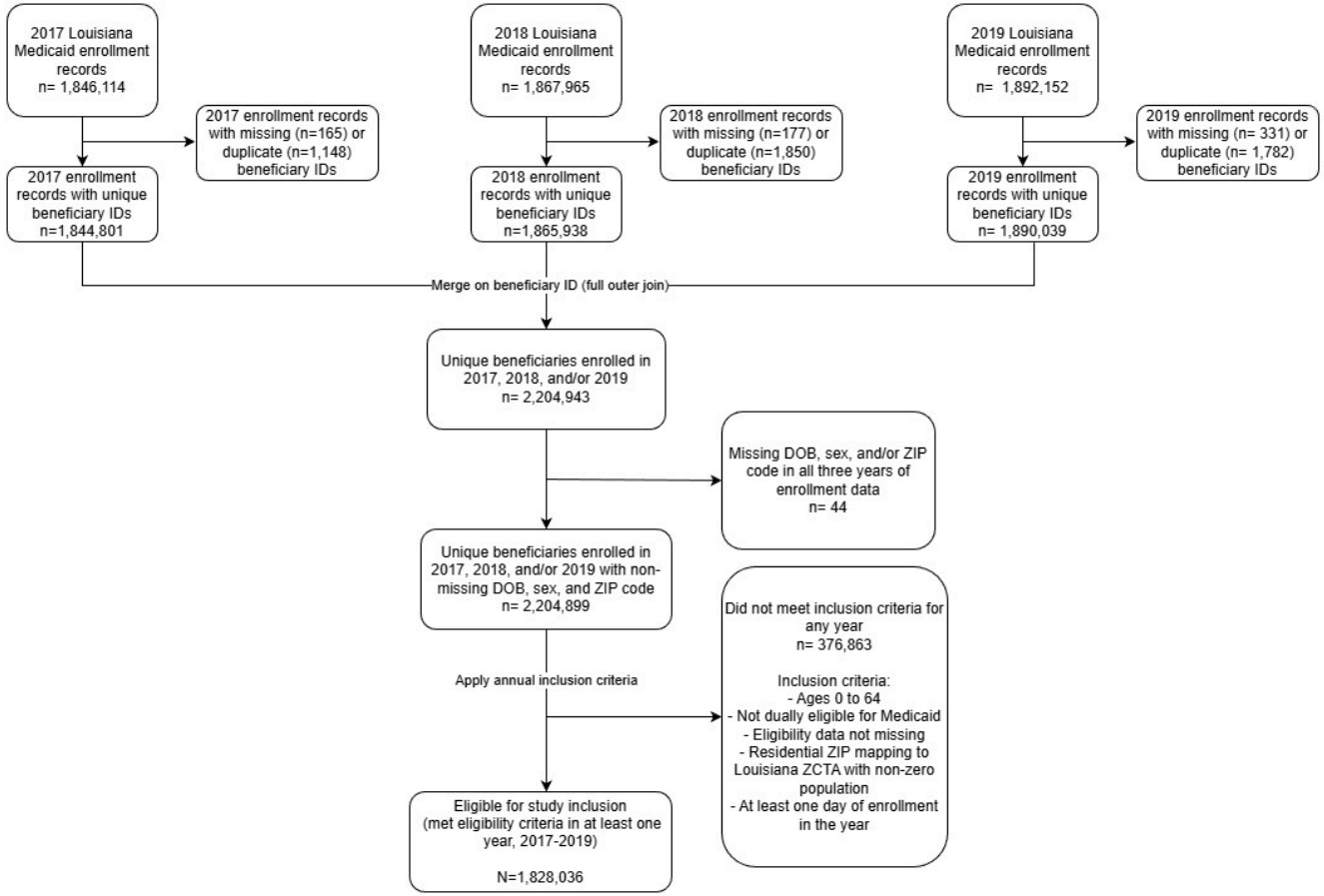


Figure 1. Flow diagram illustrating creation of analytic sample.

1. Identified unique, non-missing beneficiary identifiers (BENE_IDs) from within each annual DE files for the three years of data (2017, 2018, and 2019).
    a. BENE_ID is unique to each individual enrolled in Medicaid and is stable across years and different enrollment periods.
2. Within each annual file, de-duplicated beneficiary records by retaining one enrollment record (row) per unique BENE_ID.
    a. Each annual DE file may list a beneficiary multiple times due to data discrepancies in state Medicaid records and the procedures used to create the DE files from the original source data. The DE files are created at the beneficiary level using unique combinations of a state-specific identifier (MSIS ID) and submitting states (i.e. the state that submitted the enrollment record to Medicaid). To identify unique beneficiaries, we instead relied on the nationally-assigned beneficiary ID, which is more stable over time. Duplicate records shared the same beneficiary ID but had different MSIS IDs. Review of a subset of these duplicates suggests they represent the same individuals (identical date of birth, sex, residential ZIP code, and residential county) with discrepancies in enrollment information (e.g., managed care plan enrollment or eligibility category).
    b. Within-year variation in key eligibility variables among these duplicates (i.e., residential ZIP code, dual eligibility flag) was very rare (<0.05%), so keeping any single record per beneficiary was deemed acceptable.
3. Merged the three years of de-duplicated records into a single, person-level file, keeping each year's demographic variables (DOB, sex, race/ethnicity, household size, residential ZIP code, residential state, dual eligibility code, missing eligibility data flag, enrollment days in the year, and eligibility group code).
4. Created a canonical DOB, sex, and race/ethnicity, for each beneficiary by selecting the most frequently occurring value across years; in the case of ties, randomly assigned a non-missing value. Collapsed race/ethnicity into five categories due to missingness and small cell sizes (non-Hispanic White, non-Hispanic Black, Hispanic, Other, Not reported).
    a. Birth date and biological sex are fixed, so differences across years are due to recording errors. While race/ethnicity identity can change, we will assume it to be relatively stable over the three year period.
5. Dropped beneficiaries who had missing DOB, sex, and/or ZIP code in all three years.
6. Using the canonical DOB, calculated beneficiary age for each calendar year.
7. Applied year-specific eligibility criteria to each calendar year, creating annual eligibility flags indicating whether a beneficiary met inclusion criteria in at least one year during the study period:
    a. Aged 0 to 64 years
    b. Louisiana listed as the state in beneficiary's residential address (BENE_STATE_CD="22")
    c. Louisiana ZIP code listed as the ZIP code in beneficiary's residential address based on the list of Louisiana ZIP codes contained in the 2017 HUD ZIP-ZCTA crosswalk. (Will add 2017 HUD ZIP ZCTA crosswalk as an appendix)
    d. Not dually eligible for Medicare (DUAL_ELGBL_CD_LTST = "00")
    e. Not flagged as having missing eligibility data (MISG_ELGBLTY_DATA_IND = 1)

f. Beneficiaries were required to have ≥1 day of enrollment in the year; those with 0 days had no opportunity for diagnosis during that year, and so were excluded. This was done using a variable derived by CCW in the TAF using the sum of values (days) from the monthly Medicaid enrollment days variables (MDCD_ENRLMT_DAYS_YR >0).
g. We did not require continuous enrollment within a year for inclusion. Allowing partial-year enrollment improves representativeness by retaining beneficiaries with unstable coverage; however, this approach may lead to underestimation of prevalence if diagnoses are missed among individuals enrolled for only brief periods. If enrollment duration differs systematically between exposed and unexposed populations, prevalence estimates and risk ratios may be biased. Most conditions of interest in this study are chronic, and beneficiaries with chronic disease would usually be expected to have multiple encounters with the healthcare system during the three-year study period.
h. In Louisiana, Medicaid is delivered primarily through managed-care organizations (Louisiana Department of Health, 2019). Consistent with this, more than 90% of beneficiaries in the 2017–2019 TAF-RIF DE files used for our analysis were enrolled in managed care at some point during each year, with the majority (>60%) enrolled in managed care throughout the year. Therefore, we did not stratify or adjust analyses by delivery system type (managed care vs fee-for-service).

8. Assigned residential ZIP code as the ZIP code in the first year in which a beneficiary met eligibility criteria; if missing in that year, used the earliest non-missing value from subsequent years.
    a. Residential ZIP code may change if a beneficiary moves during the three-year study period. We used the first recorded ZIP code when the beneficiary was eligible because it most likely reflects where they experienced the majority of their exposure during/immediately prior to the study period and in an effort to minimize potential reverse causation (e.g., relocation due to illness).
    b. Because exposure measures are based on 2016 and 2017 facility and emissions data, we assume relative spatial patterns of industrial pollution remained reasonably stable over the 2017-2019 study period.
9. Assigned age as the age in the first year in which a beneficiary met eligibility criteria.
10. Created a categorical age group variable based on age when beneficiary met eligibility criteria.
11. Assigned household size using the value recorded in the first year in which a beneficiary met eligibility criteria; if missing in that year, used the earliest non-missing value from subsequent years.

Identifying health outcomes from claims data

1. Developed ICD-10-CM code lists for each health outcome of interest. When available, code lists were obtained from CMS Chronic Conditions Warehouse (CCW) Chronic Conditions and Other Chronic Health, Mental Health, and Potentially Disabling Conditions condition algorithms; for outcomes without existing lists, codes were identified through review of the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) codebook.

2. We used two approaches to operationalize health outcomes. The first used an adapted version of the CCW Chronic Conditions and Other Chronic or Disabling Conditions algorithms, which are standardized and widely used in Medicare and Medicaid claims research. These algorithms typically require at least one inpatient claim or a specified number and type of outpatient claims for a beneficiary to qualify as having a diagnosis. The CCW algorithms also specify a look-back period; in our analysis, we extended the look-back period to include the full three years of available data (2017-2019) to capture any diagnosis during this interval. Where possible, we applied the CCW algorithms directly. For conditions without a CCW algorithm, we either applied the rules of a similar CCW condition (e.g., applying the CCW anemia algorithm to nutritional anemia and other acquired anemias) or used a commonly applied rule requiring one inpatient claim or diagnoses on any two non-drug claims.
3. Our second operationalization used a less restrictive definition: beneficiaries were coded as having a condition if they had at least one relevant diagnostic code on any inpatient or outpatient Medicaid claim. This approach was intended to capture conditions that may be documented during only a single outpatient clinical encounter.
4. In general, we prioritize findings based on the CCW definitions, because they apply more conservative criteria intended to identify confirmed diagnoses. However, the CCW definition also likely underestimates the number of affected individuals. For some health conditions, the CCW approach excluded nearly all cases, suggesting that certain conditions may frequently be documented only during a single outpatient visit. As such, we present findings from both operationalizations in this report. Cross-tabulations comparing the number of cases identified by the two approaches for statistically significant findings are provided in Appendix G. Bar charts displaying the number of diagnosis codes recorded among those with ≥1 diagnosis code are shown in Appendix H.
5. Health outcomes were defined as person-level period prevalence across 2017-2019. Conditions were identified using all available claims from this period, regardless of whether the diagnosis occurred in a year when the beneficiary met inclusion criteria.
6. Created binary indicators for each condition using inpatient and outpatient claims and managed care encounter records included in the TAF-RIF IP and OT files.
7. Merged condition indicators with the analytic beneficiary file created above, retaining only diagnoses for beneficiaries included in the analytic sample.

*Note: This approach captures conditions that may have been diagnosed outside the year in which the beneficiary met eligibility criteria. The approach may underestimate prevalence if conditions were not captured in claims during the study period, particularly among beneficiaries with short enrollment durations.

Creating primary explanatory / exposure variables

1. The smallest geographic unit available in the Medicaid data is beneficiary ZIP code. Per CMS/HIPAA regulations, beneficiary data cannot be exported/published at the beneficiary ZIP code level. For this reason, we used various groupings of ZIP codes based on different ways of measuring industrial pollution exposure as our primary explanatory variables.
2. Because ZIP codes are not spatial polygons, we used the HUD ZIP-ZCTA crosswalk to approximate residential geography. This may introduce minor spatial misclassification where ZIP codes map imperfectly to ZCTAs. Where we refer to ZIP codes throughout the report and appendices, we are typically referring to ZCTAs, but use the term ZIP code because it is more familiar to lay readers.
3. We conducted two parallel sets of analyses using two different operationalizations of residential petrochemical industry pollution exposure- residential ZIP-code level distance to petrochemical facilities (as a binary variable indicating whether the ZIP code fell within a 5 mile buffer around a petrochemical facility) and toxicity-weighted chemical concentration in air from Risk Screening Environmental Indicators (RSEI) estimates.
4. Distance to petrochemical facilities captures potential exposure to multiple unmeasured hazards not included in RSEI air emissions data (e.g., particulate matter, odors, flaring, emergency releases, other routes of pollution exposure such as through water or food), whereas RSEI provides a toxicity-weighted estimate of chemical-specific air emissions, which may provide a closer proxy for potential biological exposure pathways between pollution and health.

Distance-based exposure

1. The distance-based exposure variable was derived from 2017 data from the Louisiana Department of Environmental Quality (LDEQ).
2. We identified 40 petrochemical facilities and petroleum refineries from the LDEQ dataset using North American Industry Classification System (NAICS) codes. The facilities and NAICS codes included are listed in Appendix C.
3. We used the Pairwise Buffer tool in ArcGIS Pro to generate proximity buffers around each of the 40 facilities at varying distances (2-mile, 5-mile, 10-mile). We then overlaid the distance bands created by this tool on ZIP code tabulation areas (ZCTAs) in Louisiana to determine which ZCTAs had any of their geographic areas within any of the set distances from each facility. We chose the 5-mile distance buffer as the unit of analysis and determined whether all, part, or none of each ZCTA fell within the radius of each facility. The 5-mile distance buffer was selected because the 2-mile buffer did not include a sufficient population size to conduct our analyses; the 5-mile buffer provided the best balance of exposure contrast and sample size.
4. We converted beneficiary's first non-missing recorded five digit ZIP code into ZCTAs using the US HUD's ZIP to ZCTA crosswalk files.
5. We then classified each beneficiary into one of the categories based on whether their ZIP code corresponded to a ZCTA that fell within/outside the 5 mile radius of one of the facilities of interest.
6. In the main report, distance exposure is described conceptually in terms of proximity between ZIP code centroids and facilities; operationally, exposure classification was based on whether any portion of the ZCTA polygon fell within the specified buffer distance.

Pollution-based exposure

The pollution-based exposure was derived from data from the Environmental Protection Agency (EPA)'s Risk Screening Environmental Indicators (RSEI) geographic microdata. RSEI geographic microdata are datasets from the US EPA that contain air pollution modeling results for a variety of different spatial aggregations in the US. These data contain values resulting from where chemical releases and potential environmental impacts may occur. RSEI models releases and transfers reported to TRI from stack and fugitive air emissions, transfers off site to publicly owned treatment works (POTW) facilities, and transfers off site to incineration. Toxicity-weighted chemical concentrations are adjusted for chemical toxicity to human health, weather patterns, and other relevant factors.

We used the continuous ZIP code-level toxicity-weighted chemical concentrations for all RSEI modeled chemicals. 2016 exposure estimates were used so that modeled pollution levels preceded the observation window for health outcomes. We converted beneficiary's first non-missing recorded five digit ZIP code into ZCTAs using the US HUD's ZIP to ZCTA crosswalk files. Using the ZCTA as the linkage variable, we merged the RSEI toxicity-weighted air chemical concentrations into the main dataset. Individuals residing in three ZIP codes (71243, 71377, 71425) that were missing RSEI exposure values (n = 154, > 0.0001% of total sample) were excluded from analyses that used RSEI data as the primary exposure.

For visualization of the pollution based exposure measure in bar charts and maps, we created a five category version of the ZIP code level toxicity-weighted chemical concentrations. The classes were automatically generated through the Natural Breaks (Jenks) classification algorithm in ArcGIS Pro, which looks for natural groupings in the structure of the dataset. Breaking the data down into 5 classes is a common strategy for large datasets, but it also worked well with the structure of the dataset (five classes seemed to adequately represent variability without over-lumping or over-splitting) and from a cartographic point of view (five classes showed contrast well in our maps).

Analytic sub-samples

Individuals were assigned to analytic subgroups based on age at first eligibility. For child health outcomes, we restricted the analytic sample to beneficiaries whose age at study inclusion was 17 years or younger. For birth related outcomes (preterm birth, low birth weight, birth defects) we restricted the analytic sample to those whose age at study inclusion was 1 year or younger.

For adult health outcomes, we restricted the analytic sample to non-smoking adults aged 18 and older. Non-smoking individuals were defined as those with no ICD-10-CM diagnosis codes indicating tobacco use on any Medicaid claim during 2017–2019. We used the codes listed for "Tobacco Use Disorders" from CCW's Other Chronic Health, Mental Health, and Potentially Disabling Conditions code list for "Tobacco Use Disorders".

For analyses of pregnancy-related health outcomes, we restricted the analytic sample to non-smoking women aged 18 and older who have experienced pregnancy. Pregnancy history was defined based on diagnosis or procedure codes recorded in the Medicaid data during 2017-2019.

We used all ICD-10-CM and International Classification of Diseases, 10th Revision, Procedure Coding System (ICD-10-PCS) codes listed in the "Ever pregnant" tab of the CMS TAF-RIF "Maternal and Infant Health (MIH): Identifying Pregnant and Postpartum Beneficiaries in Medicaid and CHIP Administrative Data" reference codes document to identify women who had experienced pregnancy (any diagnosis on any claim). We did not account for multiple pregnancies, but rather, whether women had ever had a pregnancy-related diagnosis during the study period (had ever been pregnant).

Other variables

To assess the contribution of living in a food desert to the association between industrial pollution and anemia, we constructed a variable representing the proportion of residential area within each ZIP code designated as low-income and low-access to grocery stores ("food deserts"). These designations were based on 2015 data from the USDA Food Access Research Atlas. Because the Food Access Research Atlas data are available at the census tract level, we used the 2017 U.S. HUD ZIP Code Crosswalk to convert these data to ZIP code tabulation areas (ZCTAs). The Food Access Research Atlas provides a binary indicator for whether a census tract is classified as "low-income and low-access" (defined as no grocery store within 1 mile for urban tracts or within 10 miles for rural tracts). Using the crosswalk, we calculated the percentage of residential area within each ZCTA that falls within census tracts designated as low-income and low-access (i.e., a food desert). This variable was included in multivariate analyses of nutritional anemia and other acquired anemias as a covariate to control for the effect of living in a food desert.

Data analysis

We conducted two parallel sets of multivariate analyses with (1) the ZIP code distance-based binary variable and (2) the ZIP code toxicity-weighted air chemical concentration (log-transformed in the models, to adjust for skewed distribution of the data) as the primary explanatory variable.

We calculated adjusted risk ratios (RRs) using modified Poisson regression with robust (sandwich) standard errors, implemented through generalized estimating equations (GEE), with an exchangeable correlation structure to account for clustering of individuals within ZIP codes. ZIP code was selected as the clustering unit because exposure is defined at the ZIP level and individuals within ZIP codes may share unmeasured environmental and socioeconomic characteristics. Risk ratios were adjusted for age and sex, except where analyses were limited to one sex (e.g., analyses of women's health and pregnancy outcomes), in which cases, they were only adjusted for age. For anemia, risk ratios were also adjusted for food access (e.g., living in a food desert). The unit of analysis was the individual Medicaid beneficiary. Each health outcome was modeled as a binary indicator of ever diagnosed during the study period.

For the distance-based exposure models:

Let Y_{iz} denote a binary health outcome for individual *i* residing in ZIP code *z*. We model the expected prevalence of the outcome as:

$$\log(E[Y_{iz}]) = \beta_0 + \beta_1 \text{Exposure}_{iz} + \beta_2 \text{Age}_{iz} + \beta_3 \text{Sex}_{iz}$$

where:
- Exposure_{iz} is an indicator equal to 1 if the beneficiary's ZIP code lies within the 5-mile buffer of any petrochemical facility or refinery, and 0 otherwise;
- Age_{iz} is age in years in the first year that individual *i* met study inclusion criteria;
- Sex_{iz} is an indicator variable for sex;
- the log link ensures that $e^{\beta 1}$ is interpreted as the risk ratio comparing exposed vs. unexposed.

Model SAS code:

```
proc genmod data= analytic_dataset;
where /*(restrict to subsample using age, sex, and/or tobacco use variables) */;
class sex zipcode;
model outcome = exposed_5m age_inclusion sex / dist=poisson link=log;
repeated subject= zipcode;
estimate "RR" exposed_5m 1;
run;
```

For the RSEI toxicity-weighted chemical air concentration exposure models:

Let Y_{iz} denote a binary health outcome for individual *i* residing in ZIP code *z*. We model the expected prevalence of the outcome as:

$$\log(E[Y_{iz}]) = \beta_0 + \beta_1 \log(\text{Exposure}_{iz}) + \beta_2 \text{Age}_{iz} + \beta_3 \text{Sex}_{iz}$$

where:
- $\log(\text{Exposure}_{iz})$ is the natural logarithm of the continuous exposure variable;
- Age_{iz} is age in years in the first year that individual *i* met study inclusion criteria;

- Sex_{iz} is an indicator variable for sex;
- because the exposure is modeled on the log scale, a 10-fold increase in the original exposure corresponds to multiplying β_1 by ln(10), the natural logarithm of 10 (≈2.3026) so that

$$\exp(\beta_1 \ln(10))$$

represents the adjusted risk ratio associated with a 10-fold increase in exposure.

Model SAS code:

```
proc genmod data= analytic_dataset;
where /*(restrict to subsample using age, sex, and/or tobacco use variables) */;
class sex zipcode;
model outcome = log_toxconc age_inclusion sex / dist=poisson link=log;
repeated subject= zipcode;
estimate "RR" log_toxconc 2.3026;
run;
```

We used the case counts and adjusted risk ratios (RRs) from the distance-based exposure models to calculate the attributable fraction among the exposed (AFe) and estimate the number of excess cases among individuals living in ZIP codes within 5 miles of petrochemical facilities and petroleum refineries. The AFe quantifies the proportion of cases among exposed individuals that may be attributable to higher levels of industrial pollution exposure, under the assumption that the observed association reflects a causal relationship. Excess cases were calculated by multiplying the AFe by the number of observed cases among exposed individuals. These estimates are intended to provide a more interpretable measure of the potential public health burden associated with differences in industrial pollution exposure.

All analyses were conducted in SAS Enterprise Guide within the CMS VRDC, with the exception of calculations of the excess cases among the exposed, which were conducted in RStudio using exported/CMS reviewed output data.

Sensitivity analysis

To assess the sensitivity of our results to how the health outcome was defined, we conducted multivariate analyses where the health outcomes were defined using the CCW algorithms and where a less restrictive definition (one diagnosis on any claim) was used. To further probe the sensitivity of our statistically significant outcomes, we also computed crosstabulations of the prevalence of health conditions using each outcome definition (CCW and less restrictive) and constructed histograms of the number of diagnosis codes for a given health condition among those who had at least

one code (to determine how many individuals had been diagnosed on only one visit versus two or more). We also conducted sensitivity analyses using more restrictive requirements around periods of enrollment in Medicaid, where individuals were only included in the sample if they had at least 180 days of enrollment within 2017-2019 (entire 3 year period), or at least 180 days of enrollment in each year enrolled. Additionally, for outcomes that were not statistically significant, we conducted power calculations to determine whether we had sufficient sample size to detect a difference in risk of those health conditions between people living in ZIP codes within/outside of the 5-mile radius around the petrochemical facilities and refineries.

Expanded outcome definitions

We conducted analyses using models with a less restrictive outcome definition. Under this definition, a beneficiary was considered to have a given condition if any claim in any diagnosis position contained an ICD-10 code from the corresponding list at any point during 2017–2019 (instead of applying the claims requirements from the CCW algorithms).

Longer enrollment periods

Our first sensitivity analysis was to determine whether our results of analyses using the less restrictive outcome definition changed if we excluded beneficiaries with short enrollment periods, who would be less likely to have their condition captured in the data compared to those enrolled for longer periods. To conduct this analysis, we used the same inclusion criteria as in our main sample, but required at least 180 days of enrollment (was not required to be continuous) for each year they are present in the data; about 75% of the original sample met this criteria. We also conducted a set of analyses that only required a total of 180 days of enrollment across the three years of health data.

Power calculations

We conducted power calculations for health outcomes where findings were mixed or null. Power calculations were conducted using an online tool: Kane SP. Sample Size Calculator. ClinCalc: https://clincalc.com/stats/samplesize.aspx. Updated June 23, 2024. Further details and results of power calculations are provided in Appendix I.

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Anemias	Other acquired anemias	D59, D60, D61, D64 and subcodes, except: D59.0xx, D59.2xx, D59.3xx, D59.6xx, D61.0xx, D61.810, D61.811, D61.82x, D64.0xx, D64.1xx, D64.2xx	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Anemias	Nutritional anemias	D50, D51, D52, D53 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Autoimmune	Lupus	M32 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Autoimmune	Multiple sclerosis	G35, G36, G37 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Autoimmune	Systemic sclerosis	M34 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Autoimmune	Myositis	M33 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Autoimmune	Other connective tissue disorder	M35 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Birth outcomes	Low birth weight	P07.0 and subcodes, P07.1 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Birth outcomes	Preterm birth	P07.2 and subcodes, P07.3 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Birth outcomes	Birth defects	All codes beginning with Q	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Cancer	Breast cancer	C50.011-C50.012, C50.019, C50.021-C50.022, C50.029, C50.111-C50.112, C50.119, C50.121-C50.122, C50.129, C50.211-C50.212, C50.219, C50.221-C50.222, C50.229, C50.311-C50.312, C50.319, C50.321-C50.322, C50.329, C50.411-C50.412, C50.419, C50.421-C50.422, C50.429, C50.511-C50.512, C50.519, C50.521-C50.522, C50.529, C50.611-C50.612, C50.619, C50.621-C50.622, C50.629, C50.811-C50.812, C50.819, C50.821-C50.822, C50.829, C50.911-C50.912, C50.919, C50.921-C50.922, C50.929, D05.00-D05.02, D05.10-D05.12, D05.80-D05.82, D05.90-D05.92, Z17.0-Z17.1, Z19.1-Z19.2, Z85.3, Z86.000	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cancer	Colorectal cancer	C18, C18.0-C18.9, C19, C20, C49.A4, C49.A5, D01.0-D01.2, Z85.030, Z85.038, Z85.040, Z85.048	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cancer	Lung cancer	C34.00-C34.02, C34.10-C34.12, C34.30-C34.32, C34.80-C34.82, C34.90-C34.92, C34.2, D02.20-D02.22, Z85.110, Z85.118	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cancer	Prostate cancer	C61, D07.5, Z85.46	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cancer	Urologic (kidney, renal pelvis, ureter) cancer	C64.1-C64.2, C64.9, C65.1-C65.2, C65.9, C66.1-C66.2, C66.9, C68.8-C68.9, D09.10, D09.19, Z85.520, Z85.528, Z85.53-Z85.54, Z85.59	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cancer	Leukemias and lymphomas	C81.00-C81.49, C81.70-C81.79, C81.90-C81.99, C82.00-C82.69, C82.80-C82.99, C83.00-C83.19, C83.30-C83.39, C83.50-C83.59, C83.70-C83.99, C84.00-C84.19, C84.40-C84.49, C84.60-C84.79, C84.90-C84.99, C84.7A, C84.A0, C84.A1, C84.A2, C84.A3, C84.A4, C84.A5, C84.A6, C84.A7, C84.A8, C84.A9, C84.Z0, C84.Z1, C84.Z2, C84.Z3, C84.Z4, C84.Z5, C84.Z6, C84.Z7, C84.Z8, C84.Z9, C85.10-C85.29, C85.80-C85.99, C86.0-C86.6, C88.4, C90.10-C90.12, C91.00-C91.02, C91.10-C91.12, C91.30-C91.32, C91.40-C91.42, C91.50-C91.52, C91.60-C91.62, C91.90-C91.92, C91.A0, C91.A1, C91.A2, C91.Z0, C91.Z1, C91.Z2, C92.00-C92.02, C92.10-C92.12, C92.20-C92.22, C92.30-C92.32, C92.40-C92.42, C92.50-C92.52, C92.60-C92.62, C92.90-C92.92, C92.A0, C92.A1, C92.A2, C92.Z0, C92.Z1, C92.Z2, C93.00-C93.02, C93.10-C93.12, C93.30-C93.32, C93.90-C93.92, C93.Z0, C93.Z1, C93.Z2, C94.00-C94.02, C94.20-C94.22, C94.30-C94.32, C94.80-C94.82, C95.00-C95.02, C95.10-C95.12, C95.90-C95.92, C96.4, C96.9, C96.Z, D45, Z85.6, Z85.71-Z85.72, Z85.79	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Cancer	Female reproductive cancers	D06 and subcodes, C51, C51.0-C51.2, C51.8-C51.9, C52, C53, C53.0-C53.1, C53.8-C53.9, C54, C54.0-C54.3, C54.8-C54.9, C55, C56, C56.1-C56.3, C56.9, C57, C57.0-C57.4, C57.7-C57.9, C57.00-C57.02, C57.10-C57.12, C57.20-C57.22, C58, D07, D07.1-D07.3, D07.30, D07.39, Z85.4, Z85.40-Z85.41, Z85.43-Z85.44, Z86.001	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Cancer	Any cancer diagnosis/history	C00-C96, D00-D09, Z85 and all subcodes, Z08	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Cancer	Thyroid cancer	C73	none	At least 1 inpatient claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Cardiovascular disease	Acute myocardial infarction (AMI, "heart attack")	I21.01-I21.02, I21.09, I21.11, I21.19, I21.21, I21.29, I21.A1, I21.A9, I21.3-I21.4, I21.9, I21.B, I22.0-I22.2, I22.8-I22.9, I23.0-I23.8	none	At least 1 inpatient claim with DX codes	CCW 30 Chronic Conditions Algorithms

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Cardiovascular disease	Atrial fibrillation and flutter	I48.0-I48.4, I48.11, I48.19-I48.21, I48.91	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cardiovascular disease	Heart failure and non-ischemic heart disease	I09.81, I11.0, I13.0, I13.2, I42.0, I42.5-I42.8, I43, I50.1, I50.9, I50.20-I50.23, I50.30-I50.33, I50.40-I50.43, I50.82-I50.84, I50.89, I50.810-I50.814, P29.0	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cardiovascular disease	Ischemic heart disease	I20.0-I20.2, I20.8, I20.81, I20.89, I24.0-I24.1, I24.8, I24.81, I24.89, I25.10, I25.41-I25.42, I25.82-I25.85, I25.89, I25.110-I25.112, I25.118-I25.119, I25.700-I25.702, I25.708, I25.710-I25.712, I25.718-I25.722, I25.728-I25.732, I25.738-I25.739, I25.750-I25.752, I25.758-I25.762, I25.768-I25.769, I25.790-I25.792, I25.798-I25.799, I25.810-I25.812, I25.3, I25.5-I25.6, I25.9	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cardiovascular disease	Peripheral Vascular Disease (PVD)	E08.51-E08.52, E09.51-E09.52, E10.51-E10.52, E11.51-E11.52, E13.51-E13.52, I70.0-I70.1, I70.201-I70.203, I70.208-I70.209, I70.211-I70.213, I70.218-I70.219, I70.221-I70.223, I70.228-I70.229, I70.231-I70.235, I70.238-I70.239, I70.241-I70.245, I70.248-I70.249, I70.291-I70.293, I70.298-I70.299, I70.25, I70.92, I73.81, I73.89, I73.9, I79.1, I79.8	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Cardiovascular disease	Hypertension	H35.031-H35.033, H35.039, I10, I11.0, I11.9, I12.0, I12.9, I13.0, I13.2, I13.10-I13.11, I15.0-I15.2, I15.8-I15.9, I1A.0, I67.4, N26.2	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Cardiovascular disease	Stroke / transient ischemic attack	G45.0-G45.3, G45.8-G45.9, G46.0-G46.8, G97.31-G97.32, I60.00-I60.02, I60.10-I60.12, I60.20-I60.22, I60.30-I60.32, I60.50-I60.52, I60.2, I60.4, I60.6-I60.9, I61.0-I61.6, I61.8-I61.9, I62.00-I62.02, I62.9, I63.00, I63.02, I63.09-I63.10, I63.12, I63.19-I63.20, I63.22, I63.29-I63.30, I63.39-I63.40, I63.49-I63.50, I63.59, I63.81, I63.89, I63.011-I63.013, I63.019, I63.031-I63.033, I63.039, I63.111-I63.113, I63.119, I63.131-I63.133, I63.139, I63.211-I63.213, I63.219, I63.231-I63.233, I63.239, I63.311-I63.313, I63.319, I63.321-I63.323, I63.329, I63.331-I63.333, I63.339, I63.341-I63.343, I63.349, I63.411-I63.413, I63.419, I63.421-I63.423, I63.429, I63.431-I63.433, I63.439, I63.441-I63.443, I63.449, I63.511-I63.513, I63.519, I63.521-I63.523, I63.529, I63.531-I63.533, I63.539, I63.541-I63.543, I63.549, I63.6, I63.8-I63.9, I67.841, I67.848, I67.89, I97.810-I97.811, I97.820-I97.821	S06.340A, S06.341A, S06.342A, S06.343A, S06.344A, S06.345A, S06.346A, S06.347A, S06.348A, S06.34AA, S06.349A, S06.350A, S06.351A, S06.352A, S06.353A, S06.354A, S06.355A, S06.356A, S06.357A, S06.358A, S06.35AA, S06.359A, S06.360A, S06.361A, S06.362A, S06.363A, S06.364A, S06.365A, S06.366A, S06.367A, S06.368A, S06.36AA, S06.369A, S06.370A, S06.371A, S06.372A, S06.373A, S06.374A, S06.375A, S06.376A, S06.377A, S06.378A, S06.37AA, S06.379A, S06.380A, S06.381A, S06.382A, S06.383A, S06.384A, S06.385A, S06.386A, S06.387A, S06.388A, S06.38AA, S06.389A, S06.5X0A, S06.5X1A, S06.5X2A, S06.5X3A, S06.5X4A, S06.5X5A, S06.5X6A, S06.5X7A, S06.5X8A, S06.5XAA, S06.5X9A, S06.6X0A, S06.6X1A, S06.6X2A, S06.6X3A, S06.6X4A, S06.6X5A, S06.6X6A, S06.6X7A, S06.6X8A, S06.6XAA, S06.6X9A, S06.810A, S06.811A, S06.812A, S06.813A, S06.814A, S06.815A, S06.816A, S06.817A, S06.818A, S06.81AA, S06.819A, S06.820A, S06.821A, S06.822A, S06.823A, S06.824A, S06.825A, S06.826A, S06.827A, S06.828A, S06.82AA, S06.829A, S06.890A, S06.891A, S06.892A, S06.893A, S06.894A, S06.895A, S06.896A, S06.897A, S06.898A, S06.89AA, S06.899A, S06.9X0A, S06.9X1A, S06.9X2A, S06.9X3A, S06.9X4A, S06.9X5A, S06.9X6A, S06.9X7A, S06.9X8A, S06.9XAA, S06.9X9A, S06.A0XA, S06.A1XA	At least 1 inpatient, hospital outpatient or carrier claim with DX codes	CCW 30 Chronic Conditions Algorithms
Eligibility variables	Tobacco use	F17.2 and all subcodes, O99.330-O99.335, T65.211A, T65.212A, T65.213A, T65.214A, T65.221A, T65.222A, T65.223A, T65.224A, T65.291A, T65.292A, T65.	none	Any diagnosis code on any claim	CCW Other Chronic Health, Mental Health,
Eligibility variables	Pregnancy	All ICD-10-CM codes in the "ever pregnant" tab of the Medicaid MIH Reference Codes: https://www.medicaid.gov/medicaid/data-and-systems/downloads/macbis/mih_reference_codes.pdf	none	Any diagnosis code on any claim	CMS TAF-RIF Maternal and Infant Health (MIH): Identifying Pregnant and Postpartum Beneficiaries in Medicaid and CHIP Administrative Data
Endocrine disorders	Hypothyroidism	E00.0-E00.2, E00.9, E01.8, E02, E03.0-E03.4, E03.8-E03.9, E89.0	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Endocrine disorders	Diabetes, type 1	E10, E10.1-E10.6, E10.8-E10.9, E10.10-E10.11, E10.21-E10.22, E10.29, E10.31-E10.37, E10.39-E10.44, E10.49, E10.51-E10.52, E10.59, E10.61-E10.65, E10.69, E10.311, E10.319, E10.321, E10.329, E10.331, E10.339, E10.341, E10.349, E10.351-E10.355, E10.359, E10.610, E10.618, E10.620-E10.622, E10.628, E10.630, E10.638, E10.641, E10.649	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Endocrine disorders	Diabetes, type 2	E11, E11.0-E11.6, E11.8-E11.9, E11.00-E11.01, E11.10-E11.11, E11.21-E11.22, E11.29, E11.31-E11.37, E11.39-E11.44, E11.49, E11.51-E11.52, E11.59, E11.61-E11.65, E11.69, E11.311, E11.319, E11.321, E11.329, E11.331, E11.339, E11.341, E11.349, E11.351-E11.355, E11.359, E11.610, E11.618, E11.620-E11.622, E11.628, E11.630, E11.638, E11.641, E11.649	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Endocrine disorders	Thyroid gland disorders (non-cancerous, includes hypothyroidism)	E00, E00.0-E00.2, E00.9, E01, E01.0-E01.2, E01.8, E02, E03, E03.0-E03.5, E03.8-E03.9, E04, E04.0-E04.2, E04.8-E04.9, E05, E05.0-E05.4, E05.8-E05.9, E05.00-E05.01, E05.10-E05.11, E05.20-E05.21, E05.30-E05.31, E05.40-E05.41, E05.80-E05.81, E05.90-E05.91, E06, E06.0-E06.5, E06.9, E07, E07.0-E07.1, E07.8-E07.9, E07.81, E07.89	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Endocrine disorders	Other glucose disorders (not diabetes)	E15, E16, E16.0-E16.4, E16.8-E16.9	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	ICD-10-CM codebook
Endocrine disorders	Early puberty among girls	E30.1	none	Any diagnosis code on any claim	ICD-10-CM codebook
Gastrointestinal	Irritable bowel syndrome (IBS)	K58 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Gastrointestinal	Gastrointestinal reflux disease (GERD)	K21 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Gastrointestinal	Irritable bowel disorders (IBDs)	K50, K51, K52 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Mental health	Anxiety disorders	F06.4, F40.00-F40.02, F40.10-F40.11, F40.210, F40.218, F40.220, F40.228, F40.230-F40.233, F40.240-F40.243, F40.248, F40.290-F40.291, F40.298, F40.8-F40.9, F41.0-F41.1, F41.3, F41.8-F41.9, F42, F42.2-F42.4, F42.8-F42.9, F43.0, F43.10-F43.12, F44.9, F45.8, F48.8-F48.9, F93.8, F99, R45.2, R45.5-R45.7	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Mental health	Bipolar disorder	F30, F30.1-F30.4, F30.8-F30.9, F30.10-F30.13, F31, F31.0, F31.2-F31.9, F31.10-F31.13, F31.30-F31.32, F31.60-F31.64, F31.70-F31.78, F31.81, F31.89, F33.8, F34.81, F34.89, F34.9, F39	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Mental health	Depressive disorders	F32, F32A, F32.0-F32.5, F32.9, F32.89, F33.0-F33.4, F33.8-F33.9, F33.40-F33.42, F34.1	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Musculoskeletal	Rheumatoid Arthritis/Osteoarthritis	L40.50-L40.51, L40.54, L40.59, M05.00, M05.09-M05.10, M05.19-M05.20, M05.29-M05.30, M05.39-M05.40, M05.49-M05.50, M05.59-M05.60, M05.69-M05.70, M05.79-M05.80, M05.89, M05.7A, M05.8A, M05.011-M05.012, M05.019, M05.021-M05.022, M05.029, M05.031-M05.032, M05.039, M05.041-M05.042, M05.049, M05.051-M05.052, M05.059, M05.061-M05.062, M05.069, M05.071-M05.072, M05.079, M05.111-M05.112, M05.119, M05.121-M05.122, M05.129, M05.131-M05.132, M05.139, M05.141-M05.142, M05.149, M05.151-M05.152, M05.159, M05.161-M05.162, M05.169, M05.171-M05.172, M05.179, M05.211-M05.212, M05.219, M05.221-M05.222, M05.229, M05.231-M05.232, M05.239, M05.241-M05.242, M05.249, M05.251-M05.252, M05.259, M05.261-M05.262, M05.269, M05.271-M05.272, M05.279, M05.311-M05.312, M05.319, M05.321-M05.322, M05.329, M05.331-M05.332, M05.339, M05.341-M05.342, M05.349, M05.351-M05.352, M05.359, M05.361-M05.362, M05.369, M05.371-M05.372, M05.379, M05.411-M05.412, M05.419, M05.421-M05.422, M05.429, M05.431-M05.432, M05.439, M05.441-M05.442, M05.449, M05.451-M05.452, M05.459, M05.461-M05.462, M05.469, M05.471-M05.472, M05.479, M05.511-M05.512, M05.519, M05.521-M05.522, M05.529, M05.531-M05.532, M05.539, M05.541-M05.542, M05.549, M05.551-M05.552, M05.559, M05.561-M05.562, M05.569, M05.571-M05.572, M05.579, M05.611-M05.612, M05.619, M05.621-M05.622, M05.629, M05.631-M05.632, M05.639, M05.641-M05.642, M05.649, M05.651-M05.652, M05.659, M05.661-M05.662, M05.669, M05.671-M05.672, M05.679, M05.711-M05.712, M05.719, M05.721-M05.722, M05.729, M05.731-M05.732, M05.739, M05.741-M05.742, M05.749, M05.751-M05.752, M05.759, M05.761-M05.762, M05.769, M05.771-M05.772, M05.779, M05.811-M05.812, M05.819, M05.821-M05.822, M05.829, M05.831-M05.832, M05.839, M05.841-M05.842, M05.849, M05.851-M05.852, M05.859, M05.861-M05.862, M05.869, M05.871-M05.872, M05.879, M05.9, M06.00, M06.08-M06.09, M06.20, M06.28-M06.30, M06.38-M06.39, M06.80, M06.88-M06.89, M06.0A, M06.8A, M06.011-M06.012, M06.019, M06.021-M06.022, M06.029, M06.031-M06.032, M06.039, M06.041-M06.042, M06.049, M06.051-M06.052, M06.059, M06.061-M06.062, M06.069, M06.071-M06.072, M06.079, M06.211-M06.212, M06.219, M06.221-M06.222, M06.229, M06.231-M06.232, M06.239, M06.241-M06.242, M06.249, M06.251-M06.252, M06.259, M06.261-M06.262, M06.269, M06.271-M06.272, M06.279, M06.311-M06.312, M06.319, M06.321-M06.322, M06.329, M06.331-M06.332, M06.339, M06.341-M06.342, M06.349, M06.351-M06.352, M06.359, M06.361-M06.362, M06.369, M06.371-M06.372, M06.379, M06.811-M06.812, M06.819, M06.821-M06.822, M06.829, M06.831-M06.832, M06.839, M06.841-M06.842, M06.849, M06.851-M06.852, M06.859, M06.861-M06.862, M06.869, M06.871-M06.872, M06.879, M06.1, M06.9, M08.00, M08.08-M08.09, M08.20, M08.28-M08.29, M08.40, M08.48, M08.80, M08.88-M08.90, M08.98-M08.99, M08.0A, M08.2A, M08.4A, M08.9A, M08.011-M08.012, M08.019, M08.021-M08.022, M08.029, M08.031-M08.032, M08.039, M08.041-M08.042, M08.049, M08.051-M08.052, M08.059, M08.061-M08.062, M08.069, M08.071-M08.072, M08.079, M08.211-M08.212, M08.219, M08.221-M08.222, M08.229, M08.231-M08.232, M08.239, M08.241-M08.242, M08.249, M08.251-M08.252, M08.259, M08.261-M08.262, M08.269, M08.271-M08.272, M08.279, M08.411-M08.412, M08.419, M08.421-M08.422, M08.429, M08.431-M08.432, M08.439, M08.441-M08.442, M08.449, M08.451-M08.452, M08.459, M08.461-M08.462, M08.469, M08.471-M08.472, M08.479, M08.811-M08.812, M08.819, M08.821-M08.822, M08.829, M08.831-M08.832, M08.839, M08.841-M08.842, M08.849, M08.851-M08.852, M08.859, M08.861-M08.862, M08.869, M08.871-M08.872, M08.879, M08.911-M08.912, M08.919, M08.921-M08.922, M08.929, M08.931-M08.932, M08.939, M08.941-M08.942, M08.949, M08.951-M08.952, M08.959, M08.961-M08.962, M08.969, M08.971-M08.972, M08.979, M08.1, M08.3, M15.0-M15.4, M15.8-M15.9, M16.0, M16.2, M16.4, M16.6-M16.7, M16.9, M16.10-M16.12, M16.30-M16.32, M16.50-M16.52, M17.0, M17.2, M17.4-M17.5, M17.9, M17.10-M17.12, M17.30-M17.32, M18.0, M18.2, M18.4, M18.9, M18.10-M18.12, M18.30-M18.32, M18.50-M18.52, M19.011-M19.012, M19.019, M19.021-M19.022, M19.029, M19.031-M19.032, M19.039, M19.041-M19.042, M19.049, M19.071-M19.072, M19.079, M19.111-M19.112, M19.119, M19.121-M19.122, M19.129, M19.131-M19.132, M19.139, M19.141-M19.142, M19.149, M19.171-M19.172, M19.179, M19.211-M19.212, M19.219, M19.221-M19.222, M19.229, M19.231-M19.232, M19.239, M19.241-M19.242, M19.249, M19.271-M19.272, M19.279, M19.09, M19.19, M19.29, M19.90-M19.93, M45.0-M45.9, M45.A0, M45.A1, M45.A2, M45.A3, M45.A4, M45.A5, M45.A6, M45.A7, M45.A8, M45.AB, M46.80-M46.99, M47.011-M47.016, M47.019, M47.021-M47.022, M47.029, M47.811-M47.819, M47.891-M47.899, M47.10-M47.16, M47.20-M47.28, M47.9, M48.8X1, M48.8X2, M48.8X3, M48.8X4, M48.8X5, M48.8X6, M48.8X7, M48.8X8, M48.8X9	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Musculoskeletal	Osteoporosis	M80.00XA, M80.011A, M80.012A, M80.019A, M80.021A, M80.022A, M80.029A, M80.031A, M80.032A, M80.039A, M80.041A, M80.042A, M80.049A, M80.051A, M80.052A, M80.059A, M80.061A, M80.062A, M80.069A, M80.071A, M80.072A, M80.079A, M80.08XA, M80.0AXA, M80.0B1A, M80.0B2A, M80.0B9A, M80.80XA, M80.811A, M80.812A, M80.819A, M80.821A, M80.822A, M80.829A, M80.831A, M80.832A, M80.839A, M80.841A, M80.842A, M80.849A, M80.851A, M80.852A, M80.859A, M80.861A, M80.862A, M80.869A, M80.871A, M80.872A, M80.879A, M80.88XA, M80.8AXA, M80.8B1A, M80.8B2A, M80.8B9A, M81.0, M81.6, M81.8	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/Carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Neurodegenerative diseases	Alzheimer's	G30.0-G30.1, G30.8-G30.9	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Neurodegenerative diseases	Non-Alzheimer's dementia	F01.50-F01.54, F01.A0, F01.A2, F01.A3, F01.A4, F01.B0, F01.B2, F01.B3, F01.B4, F01.C0, F01.C2, F01.C3, F01.C4, F01.511, F01.518, F01.A11, F01.A18, F01.B11, F01.B18, F01.C11, F01.C18, F02.80-F02.84, F02.A0, F02.A2, F02.A3, F02.A4, F02.B0, F02.B2, F02.B3, F02.B4, F02.C0, F02.C2, F02.C3, F02.C4, F02.811, F02.818, F02.A11, F02.A18, F02.B11, F02.B18, F02.C11, F02.C18, F03.90-F03.94, F03.A0, F03.A2, F03.A3, F03.A4, F03.B0, F03.B2, F03.B3, F03.B4, F03.C0, F03.C2, F03.C3, F03.C4, F03.911, F03.918, F03.A11, F03.A18, F03.B11, F03.B18, F03.C11, F03.C18, F05, G13.8, G31.01, G31.09, G31.83, G31.1-G31.2, G94, R41.81	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Neurodegenerative diseases	Parkinson's Disease and Secondary Parkinsonism	G20, G20.A1, G20.A2, G20.B1, G20.B2, G20.C, G21.11, G21.19, G21.3-G21.4, G21.8-G21.9, G31.83	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Neurodevelopmental disorders	ADHD, Conduct Disorders, and Hyperkinetic Syndrome	F63.0-F63.3, F63.9, F63.81, F63.89, F90.0-F90.2, F90.8-F90.9, F91.0-F91.3, F91.8-F91.9	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Neurodevelopmental disorders	Autism Spectrum Disorders	F84.0, F84.3, F84.5, F84.8-F84.9	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Neurodevelopmental disorders	Intellectual Disabilities and Related Conditions	E78.71-E78.72, F70, F71, F72, F73, F78, F78.A1, F78.A9, F79, P04.3, Q86.0, Q87.1-Q87.3, Q87.5, Q87.11, Q87.19, Q87.81, Q87.83-Q87.85, Q87.89, Q89.7-Q89.8, Q90.0-Q90.2, Q90.9, Q91.0-Q91.7, Q92.0-Q92.2, Q92.5, Q92.7-Q92.9, Q92.61-Q92.62, Q93.0-Q93.5, Q93.7, Q93.9, Q93.51-Q93.52, Q93.59, Q93.81, Q93.88-Q93.89, Q95.2-Q95.3, Q99.2	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Neurodevelopmental disorders	Learning Disabilities	F80.0-F80.2, F80.4, F80.9, F80.81-F80.82, F80.89, F81.0, F81.2, F81.9, F81.81, F81.89, F82, H93.25, R48.0	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Neurodevelopmental disorders	Other Developmental Delays	F81.9, F82, F88, F89	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Other	Chronic kidney disease	A18.11, A52.75, B52.0, E08.21-E08.22, E08.29, E09.21-E09.22, E09.29, E10.21-E10.22, E10.29, E11.21-E11.22, E11.29, E13.21-E13.22, E13.29, I12.0, I12.9, I13.0, I13.2, I13.10-I13.11, K76.7, M10.30, M10.38-M10.39, M10.311-M10.312, M10.319, M10.321-M10.322, M10.329, M10.331-M10.332, M10.339, M10.341-M10.342, M10.349, M10.351-M10.352, M10.359, M10.361-M10.362, M10.369, M10.371-M10.372, M10.379, M32.14-M32.15, M35.04, M350A, N01.0-N01.9, N01A, N02.0-N02.9, N02A, N02B1, N02B2, N02B3, N02B4, N02B5, N02B6, N02B9, N03.0-N03.9, N03A, N04.0-N04.9, N04.20-N04.22, N04.29, N04A, N05.0-N05.9, N05A, N06.0-N06.9, N06.20-N06.22, N06.29, N06A, N07.0-N07.9, N07A, N08, N14.0-N14.4, N14.11, N14.19, N15.0, N15.8-N15.9, N16, N18.1-N18.6, N18.9, N18.30-N18.32, N25.1, N25.9, N25.89, N26.1, N26.9, N99.0, Q61.02, Q61.11, Q61.19, Q61.2-Q61.5, Q61.8	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Other	Liver Disease, Cirrhosis and Other Liver Conditions (except Viral Hepatitis)	K70.0, K70.2, K70.9, K70.10-K70.11, K70.30-K70.31, K70.40-K70.41, K71.0, K71.7-K71.9, K71.11, K72.00-K72.01, K72.10-K72.11, K72.90-K72.91, K74.0-K74.5, K74.00-K74.02, K74.60, K74.69, K75.0-K75.1, K75.9, K75.81, K75.89, K76.0-K76.3, K76.5-K76.7, K76.9, K76.81-K76.82, K76.89, K77, K80.30-K80.37, K83.0, R16.0, R16.2, Z48.23, Z94.4. Procedure codes: 06L20ZZ, 06L23ZZ, 06L24ZZ, 06L30ZZ, 06L33ZZ, 06L34ZZ, 0DL57DZ, 0DL58DZ, 0D9S30Z, 0D9S3ZZ, 0D9S40Z, 0D9S4ZZ, 0D9T30Z, 0D9T3ZZ, 0D9T40Z, 0D9T4ZZ, 0D9V30Z, 0D9V3ZZ, 0D9V40Z, 0D9V4ZZ, 0D9W30Z, 0D9W3ZZ, 0D9W40Z, 0D9W4ZZ, 0W9F30Z, 0W9F3ZZ, 0W9F40Z, 0W9F4ZZ, 0W9G30Z, 0W9G3ZZ, 0W9G40Z, 0W9G4ZZ, 0W9J30Z, 0W9J3ZZ	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions

Appendix B. ICD-10 Codes and Claims Algorithms Used to Define Outcomes and Eligibility Variables

Category	Health outcome / variable	Inclusion codes	Exclusion codes	Case definition for main analyses (adapted from CCW algorithms)	Source of code list
Other	Migraine and Other Chronic Headache	G43.001, G43.009, G43.011, G43.019, G43.101, G43.109, G43.111, G43.119, G43.401, G43.409, G43.411, G43.419, G43.501, G43.509, G43.511, G43.519, G43.601, G43.609, G43.611, G43.619, G43.701, G43.709, G43.711, G43.719, G43.801, G43.809, G43.811, G43.819, G43.821, G43.829, G43.831, G43.839, G43.901, G43.909, G43.911, G43.919, G43.E01, G43.E09, G43.E11, G43.E19, G43.A0, G43.A1, G43.B0, G43.B1, G43.C0, G43.C1, G43.D0, G43.D1, G44.001, G44.009, G44.011, G44.019, G44.021, G44.029, G44.031, G44.039, G44.041, G44.049, G44.051, G44.059, G44.091, G44.099, G44.201, G44.209, G44.211, G44.219, G44.221, G44.229, G44.301, G44.309, G44.311, G44.319, G44.321, G44.329, G44.1, G44.40-G44.41, G44.51-G44.53, G44.59, G44.81-G44.86, G44.89	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Other	Epilepsy	G40.001, G40.009, G40.011, G40.019, G40.101, G40.109, G40.111, G40.119, G40.201, G40.209, G40.211, G40.219, G40.301, G40.309, G40.311, G40.319, G40.401, G40.409, G40.411, G40.419, G40.501, G40.509, G40.801-G40.804, G40.811-G40.814, G40.821-G40.824, G40.833-G40.834, G40.901, G40.909, G40.911, G40.919, G40.A01, G40.A09, G40.A11, G40.A19, G40.B01, G40.B09, G40.B11, G40.B19, G40.C01, G40.C09, G40.C11, G40.C19, G40.42, G40.89	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Other	Obesity	E66.01, E66.09, E66.89, E66.1-E66.2, E66.8-E66.9, E66.811-E66.813, E88.82, Z68.30-Z68.39, Z68.41-Z68.45, Z68.55-Z68.56	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	CCW Other Chronic Health, Mental Health, and Potentially Disabling Conditions
Pregnancy outcomes	Preeclampsia/eclampsia	O14, O14.0-O14.2, O14.9, O14.00, O14.02-O14.05, O14.10, O14.12-O14.15, O14.20, O14.22-O14.25, O14.90, O14.92-O14.95, O15, O15.0-O15.2, O15.9, O15.00, O15.02-O15.03	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Pregnancy outcomes	Ectopic pregnancy	O00, O00.0-O00.2, O00.8-O00.9, O00.00-O00.01, O00.10-O00.11, O00.20-O00.21, O00.101-O00.102, O00.109, O00.111-O00.112, O00.119, O00.201-O00.202, O00.209, O00.211-O00.212, O00.219	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Pregnancy outcomes	Miscarriage	O03, O03.0-O03.9, O03.30-O03.39, O03.80-O03.89	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Pregnancy outcomes	Gestational diabetes	O24.4, O24.41-O24.43, O24.410, O24.414-O24.415, O24.419-O24.420, O24.424-O24.425, O24.429-O24.430, O24.434-O24.435, O24.439, Z86.32	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Pregnancy outcomes	Unspecified maternal hypertension (not preeclampsia/eclampsia)	O16, O16.1-O16.5, O16.9	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Reproductive health	Male infertility	N46, N46.0-N46.1, N46.8-N46.9, N46.01-N46.02, N46.11-N46.12, N46.021-N46.025, N46.029, N46.121-N46.125, N46.129, Z31.41	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Reproductive health	Female infertility	N97, N97.0-N97.2, N97.8-N97.9, Z31.41	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Reproductive health	Fibroids	D25, D25.0-D25.2, D25.9	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Reproductive health	Early menopause	E28.31	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Respiratory illness	Chronic rhinitis, nasopharyngitis, and pharyngitis	J31 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Respiratory illnesses	Asthma	J45.20-J45.22, J45.30-J45.32, J45.40-J45.42, J45.50-J45.52, J45.901-J45.902, J45.909, J45.990-J45.991, J45.998	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Respiratory illnesses	Chronic obstructive pulmonary disease (COPD)	J40, J41, J41.0-J41.1, J41.8, J42, J43, J43.0-J43.2, J43.9, J44, J44.0-J44.1, J44.9, J44.81, J44.89, J47, J47.0-J47.1, J47.9, J98.2-J98.3	none	At least 1 inpatient/home health claim OR 2 hospital outpatient/carrier claims with DX codes	CCW 30 Chronic Conditions Algorithms
Respiratory illnesses	Seasonal allergies / hay fever	J30 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Respiratory illnesses	Acute and/or chronic sinusitis	J32, J01 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Skin	Dermatitis and eczema	L20, L21, L22, L23, L24, L25, L26, L27, L28, L29, L30 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook
Skin	Psoriasis	L40 and subcodes	none	At least 1 inpatient claim OR 2 other non-drug claims of any service type	ICD-10-CM codebook

Appendix C. List of included NAICS codes, petrochemical facilities and petroleum refineries

Agency Interest No	Facility_Name	Company	NAIC_CODE	NAIC_DESC	NAIC_GROUP	SIC_CODE	SIC_DESC	Parish	Facility Address Line 1	Facility City	Facility ZIP	Facility_Phone_ Number
248	Deltech Corp - Baton Rouge Facility	Deltech Corp	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	East Baton Rouge	11911 Scenic Hwy	Baton Rouge	70807	2257750150
312	Calumet Cotton Valley Refining LLC - Cotton Valley Refinery	Calumet Cotton Valley Refining LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Webster	1756 Old Hwy 7	Cotton Valley	71018	3188324236
1096	Monsanto Co - Luling Plant	Monsanto Luling Plant	325320	Pesticide and Other Agricultural Chemical Manufacturing	Petrochem_manufacture	2899	Chemical preparations, nec	St. Charles	12501 River Rd	Luling	70070	9857858211
1136	Shell Chemical LP - Geismar Plant	Shell Chemical LP	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	Ascension	7594 Hwy 75	Geismar	70734	2252016278
1214	Calumet Shreveport Refining LLC - Calumet Shreveport Refinery	Calumet Shreveport Refining LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Caddo	3333 Midway Ave	Shreveport	71109	3186324264
1224	Calumet Princeton Refining LP - Calumet Princeton Refinery	Calumet Princeton Refining LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Bossier	10234 Hwy 157	Princeton	71067	3189492421
1238	Valero Refining - Meraux LLC - Meraux Refinery	Valero Refining - Meraux LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Bernard	2500 E St Bernard Hwy	Meraux	70075	5042714141
1250	CITGO Petroleum Corp - Lake Charles Manufacturing Complex	Citgo Petroleum Corp	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Calcasieu	4401 Hwy 108 S	Sulphur	70601	3377086079
1376	Chalmette Refining LLC	Chalmette Refining LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Bernard	500 W St Bernard Hwy	Chalmette	70044	5042811918
1406	Equilon Enterprises LLC - Shell Oil Products US Norco Refinery	Equilon Enterprises LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Charles	15536 River Rd	Norco	70079	5044657609
1607	TOTAL Petrochemicals & Refining USA Inc - Cos-Mar Styrene Monomer Plant	Cos-Mar Company	325110	Petrochemical Manufacturing	Petrochem_manufacture	2865	Cyclic organic crudes, intermediates, dyes and pigments	Iberville	6325 Hwy 75	Carville	70721	2256424702
2083	Union Carbide Corp - St Charles Operations Site	Union Carbide Corp - St Charles Operations	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	St. Charles	355 Hwy 3142	Hahnville	70057	98578345341 or 9857834294
2366	Placid Refining Co LLC - Placid Refining Co	Placid Refining Co LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	West Baton Rouge	1940 LA Hwy 1 N	Port Allen	70767	2253870278
2367	Syngenta Crop Protection LLC - St Gabriel Plant	Syngenta Crop Protection LLC	325320	Pesticide and Other Agricultural Chemical Manufacturing	Petrochem_manufacture	2879	Pesticides and agricultural chemicals, nec	Iberville	3905 Hwy 75	St. Gabriel	70776	2256421100
2416	CF Industries Nitrogen LLC - Donaldsonville Nitrogen Complex	CF Industries Nitrogen LLC	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2873	Nitrogenous fertilizers	Ascension	39018 Hwy 3089	Donaldsonville	703468627	2254738291
2418	Phillips 66 Co - Alliance Refinery	Phillips 66 Co - Alliance Refinery	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Plaquemines	15551 Hwy 23 S, 12 Mi S of	Belle Chasse	700370395	5046567711
2425	Mosaic Fertilizer LLC - Faustina Plant	Mosaic Fertilizer LLC	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2873	Nitrogenous fertilizers	St. James	9959 Hwy 18	St. James	70086	2254744098
2532	Mosaic Fertilizer LLC - Uncle Sam Plant	Mosaic Fertilizer LLC	325312	Phosphatic Fertilizer Manufacturing	Petrochem_manufacture	2819	Industrial inorganic chemicals, nec	St. James	7250 Hwy 44	Uncle Sam	70792	2254744098
2538	Phillips 66 Co - Lake Charles Refinery	Phillips 66 Co	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Calcasieu	2200 Old Spanish Trail	Westlake	70669	3374915234
2638	ExxonMobil Fuels & Lubricants Co - Baton Rouge Refinery	ExxonMobil Fuels & Lubricants Co	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	East Baton Rouge	4045 Scenic Hwy	Baton Rouge	70805	2259778183
2719	Equilon Enterprises LLC dba Shell Oil Products US - Convent Refinery	Equilon Enterprises LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. James	10700 Hwy 44	Convent	70723	2255627681

3116	Alon Refining Krotz Springs Inc	Alon Refining Krotz Springs Inc	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Landry	256 S Levee Rd Hwy 105	Krotz Springs	70750	3375560119
3133	TopChem Pollock LLC	TopChem Pollock LLC	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2075	Soybean oil mills	Grant	737 Abe Hall Rd	Pollock	71467	2252937270
3165	Marathon Petroleum Company LP - Louisiana Refining Division - Garyville Refinery	Marathon Petroleum Co LP	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. John the Baptist	4663 W Airline Hwy	Garyville	70051	9855357215
3271	Sasol Chemicals (USA) LLC - Lake Charles Chemical Complex	Sasol Chemicals (USA) LLC	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	Calcasieu	2201 Old Spanish Trail	Westlake	70669	3374945119
3462	St Rose Refinery LLC - St Rose Refinery	St Rose Refinery LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Charles	11842 River Rd Lot 2	St. Rose	70087	
3585	Calcasieu Refining Co - Lake Charles Crude Oil Refinery	Calcasieu Refining Co	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	Calcasieu	4359 W Tank Farm Rd	Lake Charles	70605	3374782130
3732	PCS Nitrogen Fertilizer LP - Geismar Facility	PCS Nitrogen Fertilizer LP	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2874	Phosphatic fertilizers	Ascension	5301 Hwy 3115	Geismar	70734	2256211500
4013	Axiall LLC - Westlake Lake Charles North	Axiall LLC	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	Calcasieu	1600 VCM Plant Rd	Westlake	70669	3377085000
4384	Shell Chemical LP - Norco Chemical Plant West Site	Shell Chemical LP	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	St. Charles	16122 River Rd	Norco	70079	5044657011
5565	NOVA Chemicals Olefins LLC - Geismar Ethylene Plant	NOVA Chemicals Olefins LLC	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	Ascension	5205 Hwy 3115	Geismar	70734	2256422100
6164	Westlake Chemical OpCo LP - Westlake Petrochemical Complex	Westlake Chemical OpCo LP	325110	Petrochemical Manufacturing	Petrochem_manufacture	2869	Industrial organic chemicals, nec	Calcasieu	900 Hwy 108	Sulphur	70665	3375832200
18070	Westlake Styrene LLC - Styrene Monomer Production Facility	Westlake Chemical Corp	325110	Petrochemical Manufacturing	Petrochem_manufacture	2865	Cyclic organic crudes, intermediates, dyes and pigments	Calcasieu	900 Hwy 108	Sulphur	70665	3375832200
26003	Valero Refining - New Orleans LLC - St Charles Refinery	Valero Refining - New Orleans LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Charles	14902 River Rd	Norco	70079	9857648611
26336	Shell Pipeline Company LP - Norco Chemical Plant - East Site	Shell Chemical LP	324110	Petroleum Refineries	Petrol_refinery	2869	Industrial organic chemicals, nec	St. Charles	15536 River Rd	Norco	70079	5044657609
184235	Dyno Nobel LA Ammonia LLC - Ammonia Production Facilty	Dyno Nobel LA Ammonia LLC	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2873	Nitrogenous fertilizers	Jefferson	10800 River Rd	Waggaman	70094	
195909	Valero Partners LA LLC - St Charles Terminal	Valero Partners LA LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Charles	14902 River Rd	Norco	70079	9857644607
200116	Tampa Port Services LLC - Faustina Plant	Tampa Port Services LLC	325311	Nitrogenous Fertilizer Manufacturing	Petrochem_manufacture	2873	Nitrogenous fertilizers	St. James	9959 Hwy 18	St. James	70086	2254741336
202737	Valero Partners Meraux LLC Tank Farm	Valero Partners Meraux LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. Bernard	2500 E St Bernard Hwy	Meraux	70075	5042714141
209489	Garyville Refining Logistics LLC	Garyville Refining Logistics LLC	324110	Petroleum Refineries	Petrol_refinery	2911	Petroleum refining	St. John the Baptist	4663 W Airline Hwy	Garyville	70051	

Appendix D. Prevalence of health outcomes in Medicaid sample

Table D1. Prevalence of infant health outcomes among infants aged less than 1 year on Medicaid in Louisiana, 2017 to 2019

	Total infant sample (n=42,526)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 15,966, 37.5%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 26,560, 62.5%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Birth defects	7,231	17.0	2,728	17.1	4,503	17.0
Low birthweight	3,280	7.7	1,309	8.2	1,971	7.4
Preterm birth	5,072	11.9	1,953	12.2	3,119	11.7
Death before age 1, any cause	101	0.2	37	0.2	64	0.2

Table D2. Prevalence of children's health outcomes among children aged 0-17 years on Medicaid in Louisiana, 2017 to 2019

	Total children (n=804,357)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 294,430, 36.6%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 509,927, 63.4%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Nutritional anemia	3,768	0.5	1,648	0.6	2,120	0.4
Other acquired anemias	6,738	0.8	2,612	0.9	4,126	0.8
All cancers	889	0.1	298	0.1	591	0.1
Leukemia and lymphoma	41	0.0	18	0.0	23	0.0
Asthma	54,857	6.8	19,980	6.8	34,877	6.8
Seasonal allergies/hay fever	104,369	13.0	33,331	11.3	71,038	13.9
Acute/chronic sinusitis	60,606	7.5	13,607	4.6	46,999	9.2
Chronic rhinitis, nasopharyngitis, and pharyngitis	12,813	1.6	4,413	1.5	8,400	1.7
ADHD, Conduct Disorders, and Hyperkinetic Syndrome	121,222	15.1	41,482	14.1	79,740	15.6
Autism spectrum disorders	11,356	1.4	4,120	1.4	7,236	1.4
Learning disabilities	44,116	5.5	18,054	6.1	26,062	5.1
Intellectual disabilities and related conditions	4,348	0.5	1,633	0.6	2,715	0.5
Other developmental delays	19,748	2.5	6,715	2.3	13,033	2.6

Anxiety Disorders	27,761	3.5	9,153	3.1	18,608	3.7
Bipolar Disorder	10,851	1.4	4,105	1.4	6,746	1.3
Depressive Disorders	24,417	3.0	8,805	3.0	15,612	3.1
Hypothyroidism	1,515	0.2	503	0.2	1,012	0.2
Diabetes type 1	1,149	0.1	385	0.1	764	0.2
Diabetes type 2	1,323	0.2	513	0.2	810	0.2
Thyroid gland disorders	2,190	0.3	711	0.2	1,479	0.3
Other disorders of glucose regulation	960	0.1	309	0.1	651	0.1
Dermatitis and eczema	86,319	10.7	33,640	11.4	52,679	10.3
Chronic kidney disease	828	0.1	308	0.1	520	0.1
Liver disease, cirrhosis, and other liver conditions (non-viral)	1,057	0.1	349	0.1	708	0.1
Migraine and other chronic headache	5,876	0.7	1,776	0.6	4,100	0.8
Epilepsy	6,303	0.8	2,097	0.7	4,206	0.8
Obesity	9,916	1.2	3,634	1.2	6,282	1.2
Death before age 18, any cause	826	0.1	293	0.1	533	0.1
Early puberty among girls (n=394,592)	1,825	0.46	757	0.5	1,068	0.4

Table D3. Prevalence of women's health outcomes among non-smoking women aged 18 to 65 on Medicaid in Louisiana, 2017 to 2019

	Total non-smoking women (n= 494,920)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 182,716, 36.9%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 312,204, 63.1%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Fibroid	8,273	1.7	3,317	1.8	4,956	1.6
ıale reproductive system cancers	2,385	0.5	861	0.5	1,517	0.5
Female infertility	518	0.1	208	0.1	310	0.1
Breast cancer	3,213	0.7	1,225	0.7	1,988	0.6
Early menopause	66	0.0	15	0.0	51	0.0
Death before age 65, any cause	4,388	0.9	1,593	0.9	2,795	0.9

Table D4. Prevalence of pregnancy-related health outcomes among non-smoking women who have experienced pregnancy aged 18 to 65 on Medicaid in Louisiana, 2017 to 2019

	Non-smoking women who have experienced pregnancy (n=96,482)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 36,284, 37.6%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 60,198, 62.4%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Preeclampsia or eclampsia	6,416	6.7	2,744	7.6	3,672	6.1
Other maternal hypertension	3,538	3.7	2,157	3.8	1,381	3.6
Ectopic pregnancy	1,094	1.1	462	1.3	632	1.1
Miscarriage	4,458	4.6	1,804	5.0	2,654	4.4
Gestational diabetes	5,158	5.4	1,880	5.2	3,278	5.5
Death before age 64, any cause	155	0.2	65	0.2	90	0.2

Table D5. Prevalence of men's health outcomes among non-smoking men aged 18 to 65 on Medicaid in Louisiana, 2017 to 2019

	Non-smoking men (n=315,812)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 114,164, 36.2%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 201,648, 63.9%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Male infertility	31	0.0	*	*	*	*
Prostate cancer	841	0.3	308	0.3	533	0.3
*= small case counts (less than 11) hidden for confidentially						

Table D6. Prevalence of adult health outcomes among non-smoking adults aged 18 to 65 on Medicaid in Louisiana, 2017 to 2019						
	Total non-smoking adult sample (n=810,732)		Lived in ZIP codes within 5 mi of petrochem/refinery (n= 296,880, 36.6%)		Lived in ZIP codes > 5 mi away from petrochem/refinery (n= 513,852, 63.4%)	
Health outcome	**Case count**	**Percent**	**Case count**	**Percent**	**Case count**	**Percent**
Nutritional anemia	19,229	2.4	7,747	2.6	11,482	2.2
Other acquired anemia	30,301	3.7	12,363	4.2	17,938	3.5
Lupus	1,699	0.2	680	0.2	1,019	0.2
Multiple sclerosis	995	0.1	392	0.1	603	0.1
Systemic sclerosis	130	0.0	55	0.0	75	0.0
Myositis	112	0.0	46	0.0	66	0.0
Other connective tissue disorders	855	0.1	312	0.1	543	0.1
Breast cancer	3,235	0.4	1,234	0.4	2,001	0.4
Colorectal cancer	1,329	0.2	464	0.2	865	0.2
Lung cancer	864	0.1	294	0.1	570	0.1
Urologic cancer	499	0.1	157	0.1	342	0.1
Leukemia and lymphoma	483	0.1	168	0.1	315	0.1
Thyroid cancer	548	0.1	182	0.1	366	0.1
All cancers	14,292	1.8	5,109	1.7	9,183	1.8
Heart attack/ AMI	2,444	0.3	940	0.3	1,504	0.3
Atrial fibrillation	4,855	0.6	1,635	0.6	3,220	0.6
Heart failure and nonischemic heart disease	11,623	1.4	4,435	1.5	7,188	1.4
Ischemic heart disease	12,892	1.6	3,734	1.3	9,158	1.8
Peripheral vascular disease	5,348	0.7	1,612	0.5	3,736	0.7
Hypertension	144,557	17.8	51,857	17.5	92,700	18.0
Stroke	8,374	1.0	3,162	1.1	5,212	1.0
Hypothyroidism	21,033	2.6	6,557	2.2	14,476	2.8
Diabetes, type 1	4,516	0.6	1,632	0.6	2,884	0.6
Diabetes, type 2	56,578	7.0	21,046	7.1	35,532	6.9
Thyroid gland disorder	29,126	3.6	9,183	3.1	19,943	3.9
Other glucose metabolism disorder	1,304	0.2	425	0.1	879	0.2
Irritable bowel syndrome (IBS)	2,327	0.3	813	0.3	1,514	0.3

Gastroesophageal reflux disease	36,741	4.5	11,589	3.9	25,152	4.9
Irritable bowel disorders	11,911	1.5	3,705	1.3	8,206	1.6
Anxiety disorders	83,220	10.3	26,094	8.8	57,126	11.1
Bipolar disorder	24,210	3.0	8,790	3.0	15,420	3.0
Depressive disorders	74,335	9.2	25,098	8.5	49,237	9.6
Arthritis	33,499	4.1	10,699	3.6	22,800	4.4
Osteoporosis	1,527	0.2	536	0.2	991	0.2
Early onset Alzheimer's disease	161	0.0	55	0.0	106	0.0
Early onset Non-Alzheimer's dementia	1,296	0.2	464	0.2	832	0.2
Early onset Parkinison's disease and secondary parkinsonsism	291	0.0	91	0.0	200	0.0
Chronic kidney disease	13,599	1.7	5,232	1.8	8,367	1.6
Liver disease, cirrhosis, and other liver conditions (non-viral)	11,982	1.5	4,301	1.5	7,681	1.5
Migraine and chronic headache	19,686	2.4	6,778	2.3	12,908	2.5
Epilepsy	10,305	1.3	3,498	1.2	6,807	1.3
Obesity	62,452	7.7	22,964	7.7	39,488	7.7
Asthma	28,917	3.6	11,217	3.8	17,700	3.4
COPD	15,132	1.9	4,409	1.5	10,723	2.1
Seasonal allergies/hay fever	33,209	4.1	11,408	3.8	21,801	4.2
Acute/chronic sinusitis	54,239	6.7	15,473	5.2	38,766	7.5
Chronic rhinitis, nasopharyngitis, and pharyngitis	2,078	0.3	700	0.2	1,378	0.3
Dermatitis and eczema	26,944	3.3	10,688	3.6	16,256	3.2
Psoriasis	2,593	0.3	858	0.3	1,735	0.3
Death before age 65, any cause	10,409	1.3	3,889	1.3	6,520	1.3

Appendix E. Results of multivariate analyses

Table E1. Results of multivariate analysis of infant health outcomes (risk ratios and 95% confidence intervals), (n=42,526).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - 180 days of enrollment in each year in sample		Sensitivity analysis - 180 days of enrollment total across the 3 years	
Category	Health outcome	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)
Reproductive and pregnancy	Birth defect	1.01 (0.95-1.07)	1.03 (1.01-1.06)	0.98 (0.94-1.02)	1.01 (0.99-1.02)	0.95 (0.91-1.00)	1.00 (0.98-1.03)	0.95 (0.91-1.00)	1.00 (0.98-1.03)
Reproductive and pregnancy	Low birth weight	1.10 (0.99-1.23)	1.00 (0.96-1.04)	1.08 (0.98-1.20)	1.00 (0.96-1.03)	1.09 (0.97-1.24)	0.99 (0.94-1.03)	1.09 (0.97-1.24)	0.99 (0.94-1.03)
Reproductive and pregnancy	Preterm birth	1.04 (0.97-1.12)	1.01 (0.98-1.04)	1.03 (0.95-1.10)	1.00 (0.97-1.03)	1.03 (0.94-1.12)	1.01 (0.97-1.04)	1.03 (0.95-1.12)	1.01 (0.97-1.04)
Mortality	Death before age 1, any cause	0.96 (0.65-1.44)	1.03 (0.86-1.22)	0.96 (0.65-1.44)	1.03 (0.86-1.22)	1.37 (0.69-2.72)	0.92 (0.72-1.18)	1.37 (0.69-2.72)	0.92 (0.72-1.18)

Table E2. Key findings of multivariate analysis of children's health outcomes (risk ratios and 95% confidence intervals), among all children aged 0-17 years (n=804,357).

		Primary analyses		CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years		Sensitivity analysis -CCW algorithm case definition, controlled for food access	
Category	Health outcome	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)
Anemia	Nutritional anemia	1.34 (1.13-1.59)	1.16 (1.08-1.23)	1.34 (1.13-1.59)	1.16 (1.09-1.23)	1.22 (1.03-1.44)	1.12 (1.05-1.19)	1.22 (1.04-1.43)	1.12 (1.06-1.19)	1.22 (1.03-1.44)	1.12 (1.05-1.19)	1.34 (1.13-1.59)	1.16 (1.08-1.23)
Neurodevelopmental	Learning disabilities	1.20 (1.05-1.37)	1.12 (1.06-1.18)	1.20 (1.05-1.37)	1.12 (1.06-1.18)	1.24 (1.04-1.48)	1.16 (1.10-1.23)	1.47 (1.16-1.85)	1.17 (1.11-1.24)	1.24 (1.04-1.48)	1.17 (1.10-1.23)	-	-
Others	Dermatitis and eczema	1.10 (1.01-1.20)	1.06 (1.03-1.08)	1.10 (1.01-1.20)	1.06 (1.03-1.08)	1.03 (0.98-1.09)	1.02 (1.01-1.04)	1.04 (0.98-1.09)	1.02 (1.01-1.04)	1.03 (0.98-1.09)	1.02 (1.01-1.04)	-	-
Endocrine	Early puberty (among girls only)	1.24 (1.10-1.41)	1.09 (1.04-1.16)	1.04 (0.35-3.07)	1.12 (0.79-1.60)	1.24 (1.10-1.41)	1.09 (1.04-1.16)	1.25 (1.10-1.41)	1.10 (1.04-1.16)	1.23 (1.09-1.39)	1.10 (1.04-1.15)	-	-

Note: primary risk ratio for nutritional anemia includes CCW algorithm case definition and adjusting for food access; primary risk ratios for learning disabilities and dermatitis/eczema includes CCW algorithm case definition; primary risk ratio for early puberty includes less restrictive case definition since early puberty is an event that occurs once.

Table E3. Mixed findings of multivariate analysis of children's health outcomes (risk ratios and 95% confidence intervals), among all children aged 0-17 years (n=804,357).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Cancers	Cancer (any type)	0.88 (0.75-1.03)	0.92 (0.87-0.99)	0.88 (0.78-1.01)	0.91 (0.86-0.96)	0.90 (0.79-1.02)	0.91 (0.86-0.96)	0.90 (0.78-1.02)	0.91 (0.86-0.96)
Endocrine	Hypothyroidism	0.87 (0.74-1.02)	0.88 (0.83-0.92)	0.94 (0.80-1.11)	0.89 (0.85-0.93)	0.93 (0.78-1.10)	0.88 (0.84-0.92)	0.96 (0.81-1.13)	0.89 (0.85-0.94)
Endocrine	Other disorders of glucose regulation	0.82 (0.71-0.93)	0.94 (0.89-0.99)	0.85 (0.77-0.93)	0.94 (0.91-0.97)	0.85 (0.77-0.94)	0.94 (0.90-0.98)	0.85 (0.77-0.94)	0.94 (0.90-0.97)
Endocrine	Thyroid gland disorders	0.84 (0.73-0.96)	0.89 (0.85-0.92)	0.88 (0.77-1.00)	0.88 (0.85-0.92)	0.86 (0.75-0.99)	0.88 (0.84-0.91)	0.89 (0.78-1.01)	0.88 (0.85-0.92)
Mental health	Anxiety Disorders	0.86 (0.79-0.94)	1.02 (0.99-1.05)	0.87 (0.80-0.94)	0.98 (0.95-1.01)	0.86 (0.79-0.93)	0.98 (0.95-1.01)	0.86 (0.80-0.94)	0.98 (0.95-1.01)
Neurodevelopmental	ADHD, Conduct Disorders, and Hyperkinetic Syndrome	0.91 (0.86-0.97)	0.95 (0.93-0.97)	0.91 (0.87-0.96)	0.95 (0.93-0.97)	0.92 (0.87-0.96)	0.95 (0.93-0.97)	0.91 (0.87-0.96)	0.95 (0.93-0.97)
Neurodevelopmental	Intellectual disabilities and related condition	1.04 (0.92-1.18)	1.07 (1.02-1.12)	1.03 (0.91-1.15)	1.05 (1.01-1.10)	1.03 (0.92-1.16)	1.05 (1.01-1.10)	1.03 (0.92-1.16)	1.05 (1.01-1.10)
Neurodevelopmental	Other developmental delays	0.89 (0.82-0.95)	1.00 (0.96-1.03)	0.93 (0.87-0.99)	1.02 (0.99-1.05)	0.94 (0.88-1.00)	1.02 (0.99-1.05)	0.93 (0.87-0.99)	1.02 (0.99-1.05)
Others	Epilepsy	0.87 (0.81-0.92)	0.94 (0.92-0.97)	0.86 (0.81-0.91)	0.94 (0.91-0.96)	0.85 (0.80-0.91)	0.94 (0.92-0.97)	0.86 (0.81-0.91)	0.94 (0.91-0.96)
Others	Liver disease, cirrhosis, and other liver conditions (non-viral)	0.86 (0.74-0.99)	0.90 (0.85-0.96)	0.77 (0.69-0.87)	0.87 (0.83-0.91)	0.76 (0.67-0.87)	0.85 (0.81-0.90)	0.77 (0.68-0.87)	0.87 (0.83-0.91)
Others	Migraine and other chronic headaches	0.76 (0.68-0.85)	0.93 (0.90-0.97)	0.80 (0.74-0.86)	0.95 (0.92-0.98)	0.80 (0.74-0.86)	0.95 (0.92-0.98)	0.80 (0.74-0.86)	0.95 (0.92-0.98)
Respiratory	Acute/chronic sinusitis	0.50 (0.44-0.57)	0.78 (0.74-0.83)	0.63 (0.57-0.68)	0.84 (0.81-0.87)	0.63 (0.58-0.69)	0.84 (0.81-0.88)	0.63 (0.57-0.69)	0.84 (0.81-0.88)
Respiratory	Seasonal allergies / hay fever	0.81 (0.76-0.88)	0.90 (0.87-0.93)	0.87 (0.82-0.91)	0.93 (0.91-0.95)	0.87 (0.83-0.92)	0.93 (0.92-0.95)	0.87 (0.30-0.91)	0.93 (0.91-0.95)

Table E4. Null findings of multivariate analysis of children's health outcomes (risk ratios and 95% confidence intervals), among all children aged 0-17 years (n=804,357).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Anemia	Other acquired anemia	1.09 (0.94-1.27)	1.04 (0.98-1.11)	1.09 (0.96-1.24)	1.02 (0.97-1.08)	1.10 (0.97-1.25)	1.03 (0.97-1.08)	1.10 (0.97-1.25)	1.02 (0.97-1.08)
Cancers	Leukemia and lymphoma	1.38 (0.73-2.59)	1.13 (0.85-1.50)	1.12 (0.65-1.93)	1.05 (0.83-1.33)	0.88 (0.48-1.61)	1.01 (0.79-1.28)	1.12 (0.65-1.93)	1.05 (0.83-1.33)
Endocrine	Diabetes type 1	0.88 (0.78-1.01)	0.97 (0.92-1.02)	0.90 (0.80-1.02)	0.96 (0.92-1.01)	0.91 (0.80-1.03)	0.97 (0.92-1.02)	0.91 (0.80-1.03)	0.97 (0.92-1.02)
Endocrine	Diabetes type 2	1.11 (0.97-1.27)	0.99 (0.94-1.04)	1.47 (1.16-1.86)	1.14 (1.05-1.25)	1.47 (1.16-1.85)	1.14 (1.05-1.25)	1.47 (1.17-1.86)	1.14 (1.05-1.25)
Mental health	Bipolar Disorder	1.07 (0.96-1.19)	1.00 (0.97-1.04)	1.02 (0.93-1.11)	0.99 (0.96-1.02)	1.01 (0.92-1.11)	0.99 (0.96-1.02)	1.01 (0.93-1.11)	0.99 (0.96-1.02)
Mental health	Depressive Disorders	0.99 (0.94-1.05)	1.00 (0.98-1.02)	0.96 (0.92-1.00)	0.99 (0.98-1.01)	0.95 (0.91-1.00)	0.99 (0.98-1.01)	0.96 (0.92-1.00)	0.99 (0.98-1.01)
Mortality	Death before age 18, any cause	0.95 (0.82-1.11)	1.02 (0.96-1.09)	0.95 (0.82-1.11)	1.02 (0.96-1.09)	1.00 (0.79-1.26)	1.01 (0.93-1.11)	1.01 (0.86-1.20)	1.04 (0.98-1.12)
Neurodevelopmental	Autism spectrum disorders	0.99 (0.91-1.08)	1.03 (1.00-1.07)	0.98 (0.91-1.06)	1.02 (0.98-1.05)	0.98 (0.90-1.07)	1.02 (0.98-1.05)	0.98 (0.91-1.06)	1.02 (0.98-1.05)
Others	Chronic kidney disease	1.03 (0.89-1.19)	0.95 (0.90-1.01)	1.00 (0.90-1.11)	0.95 (0.91-0.99)	1.01 (0.90-1.13)	0.96 (0.92-1.01)	1.01 (0.91-1.12)	0.95 (0.91-1.00)
Others	Obesity	1.01 (0.89-1.15)	1.03 (0.98-1.07)	0.96 (0.86-1.08)	1.02 (0.99-1.06)	0.95 (0.85-1.07)	1.02 (0.98-1.06)	0.96 (0.86-1.08)	1.02 (0.99-1.06)
Respiratory	Asthma	0.99 (0.91-1.09)	1.00 (0.96-1.03)	0.95 (0.88-1.03)	0.98 (0.95-1.01)	0.96 (0.89-1.03)	0.98 (0.95-1.02)	0.95 (0.88-1.03)	0.98 (0.95-1.01)
Respiratory	Chronic rhinitis, nasopharyngitis, and pharyngitis	0.90 (0.72-1.13)	1.04 (0.96-1.13)	0.98 (0.83-1.17)	1.07 (1.00-1.14)	0.99 (0.83-1.18)	1.07 (1.00-1.14)	0.99 (0.83-1.17)	1.07 (1.00-1.14)

Table E5. Key findings of multivariate analysis of women's health outcomes (risk ratios and 95% confidence intervals), among all non-smoking women aged 18-64 years (n= 494,920).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Cancers	Breast cancer	**1.10 (1.02-1.19)**	**1.06 (1.03-1.09)**	1.09 (1.02-1.17)	1.06 (1.03-1.09)	1.09 (1.01-1.18)	1.07 (1.04-1.10)	1.10 (1.03-1.19)	1.06 (1.04-1.09)
Reproductive	Uterine fibroids	**1.16 (1.05-1.28)**	**1.06 (1.03-1.10)**	1.19 (1.09-1.31)	1.08 (1.05-1.12)	1.21 (1.10-1.32)	1.09 (1.06-1.12)	1.20 (1.09-1.31)	1.09 (1.05-1.12)

Table E6. Mixed findings of multivariate analysis of women's health outcomes (risk ratios and 95% confidence intervals), among all non-smoking women aged 18-64 years (n= 494,920).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Reproductive	Early menopause	0.52 (0.28-0.97)	0.92 (0.80-1.05)	0.68 (0.50-0.92)	0.87 (0.80-0.95)	0.64 (0.46-0.89)	0.88 (0.80-0.97)	0.68 (0.50-0.93)	0.87 (0.80-0.95)

Table E7. Null findings of multivariate analysis of women's health outcomes (risk ratios and 95% confidence intervals), among all non-smoking women aged 18-64 years (n= 494,920).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition (one diagnostic code on any claim during 2017-2019)		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Mortality	Death before age 65 among women who have been pregnant, any cause	-	-	1.16 (0.85-1.58)	1.01 (0.88-1.15)	1.02 (0.91-1.14)	1.03 (0.87-1.21)	1.01 (0.91-1.11)	0.98 (0.95-1.01)
Pregnancy	Other maternal hypertension	1.04 (0.93-1.16)	1.03 (0.99-1.08)	1.06 (0.98-1.15)	1.04 (1.00-1.08)	1.05 (0.96-1.15)	1.04 (1.00-1.08)	1.06 (0.98-1.15)	1.04 (1.01-1.08)
Cancers	Female reproductive system cancers	0.99 (0.90-1.10)	0.97 (0.94-1.00)	0.94 (0.87-1.02)	0.96 (0.94-0.99)	0.96 (0.88-1.05)	0.96 (0.93-0.99)	0.95 (0.88-1.03)	0.97 (0.94-1.00)
Reproductive	Female infertility	1.14 (0.90-1.45)	1.00 (0.92-1.09)	1.12 (0.95-1.31)	1.07 (1.01-1.13)	1.11 (0.95-1.31)	1.08 (1.01-1.15)	1.12 (0.96-1.31)	1.07 (1.01-1.14)
Pregnancy	Gestational diabetes	0.92 (0.85-1.00)	0.96 (0.93-1.00)	0.94 (0.87-1.01)	0.97 (0.94-1.00)	0.94 (0.86-1.01)	0.97 (0.94-1.01)	0.93 (0.86-1.00)	0.97 (0.94-0.99)

Table E8. Key findings of multivariate analysis of pregnancy-related health outcomes (risk ratios and 95% confidence intervals), among all non-smoking women who have been pregnant aged 18-64 years (n= 96,482).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Pregnancy	Preeclampsia / eclampsia	1.24 (1.14-1.35)	1.08 (1.05-1.12)	1.22 (1.13-1.32)	1.06 (1.03-1.10)	1.23 (1.14-1.33)	1.07 (1.04-1.11)	1.22 (1.12-1.31)	1.06 (1.03-1.10)
Pregnancy	Ectopic pregnancy	1.18 (1.02-1.37)	1.09 (1.02-1.16)	1.09 (0.95-1.25)	1.00 (0.95-1.05)	1.07 (0.89-1.29)	1.00 (0.93-1.06)	1.11 (0.94-1.32)	1.00 (0.95-1.07)
Pregnancy	Miscarriage	1.11 (1.02-1.21)	1.05 (1.01-1.08)	1.06 (1.00-1.12)	1.02 (1.00-1.05)	1.04 (0.99-1.11)	1.02 (0.99-1.04)	1.05 (0.99-1.12)	1.02 (1.00-1.04)

Table E9. Null findings of multivariate analysis of pregnancy-related health outcomes (risk ratios and 95% confidence intervals), among all non-smoking women who have been pregnant aged 18-64 years (n= 96,482).

		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition (one diagnostic code on any claim during 2017-2019)		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	**Health outcome**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**	**Distance model RR (95% CI)**	**Air pollution model RR (95% CI)**
Mortality	Death before age 65 among women who have been pregnant, any cause	-	-	1.16 (0.85-1.58)	1.01 (0.88-1.15)	1.02 (0.91-1.14)	1.03 (0.87-1.21)	1.01 (0.91-1.11)	0.98 (0.95-1.01)
Pregnancy	Gestational diabetes	0.92 (0.85-1.00)	0.96 (0.93-1.00)	0.94 (0.87-1.01)	0.97 (0.94-1.00)	0.94 (0.86-1.01)	0.97 (0.94-1.01)	0.93 (0.86-1.00)	0.97 (0.94-0.99)
Pregnancy	Other maternal hypertension	1.04 (0.93-1.16)	1.03 (0.99-1.08)	1.06 (0.98-1.15)	1.04 (1.00-1.08)	1.05 (0.96-1.15)	1.04 (1.00-1.08)	1.06 (0.98-1.15)	1.04 (1.01-1.08)

Table E10. Key findings of multivariate analysis of adult health outcomes (risk ratios and 95% confidence intervals), among non-smoking adults (n=810,732).											
		Primary analysis: adapted CCW algorithm case definition, controlled for food access		Adapted CCW algorithm case definition, unadjusted for food access		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	Health outcome	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)
Anemia	Nutritional anemia	1.15 (1.08-1.23)	1.06 (1.04-1.09)	1.17 (1.09-1.25)	1.06 (1.03-1.08)	1.20 (1.12-1.28)	1.06 (1.03-1.08)	1.20 (1.13-1.28)	1.06 (1.04-1.09)	1.20 (1.13-1.28)	1.06 (1.04-1.09)
Anemia	Other acquired anemia	1.18 (1.10-1.26)	1.07 (1.05-1.10)	1.19 (1.11-1.27)	1.07 (1.04-1.09)	1.16 (1.08-1.23)	1.04 (1.01-1.07)	1.16 (1.09-1.23)	1.05 (1.02-1.07)	1.16 (1.09-1.23)	1.05 (1.02-1.07)

Table E11. Mixed findings of multivariate analysis of adult health outcomes (risk ratios and 95% confidence intervals), among non-smoking adults (n=810,732).									
		Primary analysis: adapted CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	Health outcome	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)
Autoimmune disorders	Lupus	1.15 (1.03-1.30)	1.00 (0.96-1.04)	1.04 (0.93-1.15)	0.96 (0.93-1.00)	1.05 (0.94-1.18)	0.96 (0.92-1.00)	1.04 (0.93-1.16)	0.96 (0.93-1.00)
Cancers	Colorectal cancer	0.96 (0.86-1.08)	0.95 (0.91-0.99)	0.90 (0.82-1.00)	0.94 (0.91-0.98)	0.87 (0.77-0.99)	0.93 (0.89-0.98)	0.87 (0.78-0.98)	0.95 (0.91-0.98)
Cancers	Lung cancer	0.93 (0.81-1.07)	0.95 (0.90-1.00)	0.89 (0.78-1.01)	0.93 (0.89-0.97)	0.88 (0.73-1.05)	0.94 (0.89-1.00)	0.90 (0.78-1.04)	0.93 (0.88-0.97)
Cardiovascular	Atrial fibrillation	0.91 (0.85-0.98)	0.96 (0.93-0.98)	0.94 (0.88-1.00)	0.96 (0.94-0.99)	0.92 (0.85-1.00)	0.96 (0.93-0.99)	0.93 (0.87-1.00)	0.96 (0.94-0.99)
Cardiovascular	Non-ischemic heart failure	1.10 (1.03-1.19)	1.01 (0.98-1.04)	1.10 (1.02-1.18)	0.99 (0.97-1.02)	1.09 (1.01-1.19)	0.99 (0.97-1.02)	1.09 (1.01-1.18)	0.99 (0.97-1.02)
Cardiovascular	Ischemic heart disease	0.73 (0.67-0.80)	0.86 (0.84-0.89)	0.73 (0.67-0.79)	0.85 (0.83-0.88)	0.73 (0.67-0.79)	0.86 (0.83-0.88)	0.72 (0.66-0.79)	0.85 (0.83-0.88)
Cardiovascular	Peripheral vascular disease (PVD)	0.78 (0.70-0.86)	0.90 (0.87-0.93)	0.81 (0.75-0.88)	0.90 (0.87-0.92)	0.80 (0.73-0.87)	0.89 (0.87-0.92)	0.80 (0.74-0.87)	0.90 (0.87-0.92)
Cardiovascular	Hypertension	0.99 (0.94-1.04)	0.96 (0.94-0.98)	1.00 (0.96-1.05)	0.97 (0.95-0.99)	1.00 (0.96-1.04)	0.97 (0.96-0.99)	1.00 (0.96-1.05)	0.97 (0.96-0.99)
Cardiovascular	Heart attack	1.12 (1.01-1.24)	1.02 (0.98-1.06)	1.06 (0.97-1.16)	0.99 (0.97-1.03)	1.03 (0.92-1.14)	0.99 (0.95-1.02)	1.04 (0.95-1.15)	0.99 (0.96-1.02)
Endocrine	Hypothyroidism	0.79 (0.71-0.88)	0.86 (0.83-0.89)	0.85 (0.76-0.94)	0.88 (0.86-0.90)	0.84 (0.76-0.94)	0.89 (0.86-0.91)	0.85 (0.76-0.94)	0.88 (0.86-0.91)
Endocrine	Diabetes type 1	0.98 (0.90-1.08)	0.94 (0.91-0.97)	1.01 (0.92-1.11)	0.96 (0.92-0.99)	1.02 (0.92-1.13)	0.97 (0.94-1.00)	1.02 (0.93-1.12)	0.96 (0.93-1.00)
Endocrine	Diabetes type 2	1.05 (1.01-1.10)	0.99 (0.97-1.01)	1.11 (1.07-1.16)	1.00 (0.99-1.02)	1.12 (1.08-1.16)	1.01 (1.00-1.03)	1.12 (1.07-1.16)	1.01 (0.99-1.03)
Endocrine	Thyroid gland disorders	0.80 (0.75-0.87)	0.88 (0.86-0.90)	0.85 (0.80-0.91)	0.90 (0.88-0.92)	0.85 (0.79-0.92)	0.91 (0.89-0.93)	0.86 (0.80-0.92)	0.90 (0.88-0.92)
Endocrine	Other disorders of glucose regulation	0.84 (0.75-0.95)	0.95 (0.91-1.00)	0.81 (0.73-0.89)	0.91 (0.87-0.94)	0.81 (0.72-0.90)	0.90 (0.87-0.94)	0.80 (0.72-0.89)	0.90 (0.87-0.94)
Gastrointestinal	Irritable bowel syndrome (IBS)	0.93 (0.82-1.05)	0.95 (0.91-0.99)	0.88 (0.79-0.98)	0.93 (0.90-0.96)	0.88 (0.79-0.98)	0.94 (0.91-0.97)	0.88 (0.79-0.98)	0.93 (0.90-0.97)
Gastrointestinal	Gastrointestinal reflux disease (GERD)	0.81 (0.76-0.86)	0.88 (0.87-0.90)	0.81 (0.77-0.85)	0.89 (0.87-0.91)	0.81 (0.77-0.85)	0.89 (0.88-0.91)	0.81 (0.77-0.85)	0.89 (0.88-0.91)
Gastrointestinal	Inflammatory bowel disease (IBD)	0.78 (0.73-0.83)	0.90 (0.87-0.92)	0.80 (0.76-0.85)	0.90 (0.88-0.92)	0.80 (0.76-0.85)	0.91 (0.89-0.93)	0.80 (0.76-0.85)	0.90 (0.88-0.93)
Mental health	Anxiety Disorders	0.79 (0.74-0.84)	0.91 (0.89-0.93)	0.81 (0.77-0.86)	0.92 (0.90-0.94)	0.82 (0.78-0.87)	0.93 (0.91-0.94)	0.81 (0.77-0.86)	0.92 (0.90-0.94)
Mental health	Depressive Disorders	0.88 (0.84-0.92)	0.95 (0.93-0.96)	0.86 (0.83-0.90)	0.94 (0.92-0.95)	0.87 (0.84-0.91)	0.95 (0.93-0.96)	0.86 (0.83-0.90)	0.94 (0.93-0.96)
Musculoskeletal	Rheumatoid Arthritis/Osteoarthritis	0.84 (0.79-0.88)	0.91 (0.90-0.93)	0.88 (0.84-0.92)	0.92 (0.90-0.93)	0.88 (0.84-0.92)	0.93 (0.91-0.94)	0.88 (0.84-0.92)	0.92 (0.91-0.94)
Musculoskeletal	Osteoporosis	0.98 (0.86-1.11)	0.96 (0.91-1.00)	0.93 (0.84-1.03)	0.94 (0.91-0.98)	0.92 (0.83-1.03)	0.94 (0.91-0.98)	0.94 (0.85-1.04)	0.95 (0.91-0.98)
Neurodegenerative	Early-onset Alzheimers	0.94 (0.66-1.35)	0.88 (0.79-0.98)	0.91 (0.66-1.24)	0.85 (0.77-0.94)	0.90 (0.65-1.24)	0.84 (0.76-0.94)	0.89 (0.65-1.23)	0.87 (0.78-0.97)
Neurodegenerative	Early-onset dementia (non-Alzheimer's)	1.00 (0.85-1.19)	0.97 (0.90-1.03)	0.89 (0.76-1.03)	0.90 (0.85-0.95)	0.87 (0.74-1.02)	0.90 (0.84-0.95)	0.88 (0.75-1.03)	0.90 (0.85-0.96)
Others	Chronic kidney disease	1.12 (1.05-1.19)	1.02 (1.00-1.05)	1.12 (1.05-1.19)	1.02 (1.00-1.04)	1.11 (1.05-1.18)	1.02 (1.00-1.05)	1.12 (1.06-1.19)	1.02 (1.00-1.04)
Others	Migraine and other chronic headaches	0.90 (0.85-0.96)	0.95 (0.93-0.97)	0.92 (0.88-0.97)	0.97 (0.95-0.98)	0.92 (0.88-0.97)	0.98 (0.96-1.00)	0.93 (0.88-0.97)	0.97 (0.95-0.99)
Others	Epilepsy	0.90 (0.83-0.96)	0.93 (0.91-0.96)	0.88 (0.83-0.95)	0.92 (0.90-0.95)	0.89 (0.83-0.95)	0.93 (0.91-0.96)	0.88 (0.82-0.94)	0.92 (0.90-0.95)
Others	Psoriasis	0.86 (0.76-0.99)	1.01 (0.96-1.06)	0.84 (0.74-0.94)	0.98 (0.94-1.03)	0.88 (0.77-1.00)	1.00 (0.96-1.05)	0.84 (0.74-0.94)	0.98 (0.94-1.03)
Respiratory	Chronic obstructive pulmonary disorder (COPD)	0.73 (0.67-0.80)	0.85 (0.83-0.87)	0.75 (0.70-0.82)	0.86 (0.84-0.89)	0.76 (0.70-0.82)	0.87 (0.85-0.89)	0.76 (0.70-0.82)	0.87 (0.84-0.89)
Respiratory	Seasonal allergies / hay fever	0.90 (0.82-0.99)	0.95 (0.92-0.99)	0.93 (0.88-0.99)	0.96 (0.93-0.98)	0.94 (0.89-0.99)	0.97 (0.94-0.99)	0.94 (0.88-1.00)	0.96 (0.94-0.99)
Respiratory	Acute/chronic sinusitis	0.69 (0.62-0.76)	0.83 (0.80-0.86)	0.76 (0.70-0.81)	0.87 (0.85-0.89)	0.76 (0.70-0.81)	0.88 (0.85-0.90)	0.76 (0.70-0.82)	0.87 (0.85-0.90)
Respiratory	Asthma	1.09 (1.02-1.17)	1.01 (0.99-1.04)	1.03 (0.97-1.09)	0.99 (0.97-1.01)	1.03 (0.97-1.09)	1.00 (0.98-1.02)	1.04 (0.98-1.09)	1.00 (0.98-1.02)

Table E12. Null findings of multivariate analysis of adult health outcomes (risk ratios and 95% confidence intervals), among non-smoking adults (n=810,732).									
		CCW algorithm case definition		Less restrictive case definition		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment in each year enrolled		Sensitivity analysis - Less restrictive case definition, 180 days of enrollment total across all 3 years	
Category	Health outcome	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)	Distance model RR (95% CI)	Air pollution model RR (95% CI)
Autoimmune disorders	Multiple sclerosis	1.13 (0.98-1.30)	0.98 (0.92-1.04)	1.13 (1.00-1.28)	0.97 (0.92-1.02)	1.13 (0.98-1.30)	0.97 (0.91-1.02)	1.14 (1.01-1.29)	0.98 (0.93-1.03)
Autoimmune disorders	Systemic sclerosis	1.29 (0.89-1.87)	1.02 (0.88-1.18)	1.31 (1.00-1.73)	1.04 (0.93-1.17)	1.26 (0.91-1.74)	1.07 (0.94-1.21)	1.32 (1.00-1.74)	1.05 (0.94-1.18)
Autoimmune disorders	Myositis	1.22 (0.84-1.76)	1.11 (0.95-1.29)	1.13 (0.82-1.55)	1.00 (0.89-1.13)	1.19 (0.81-1.74)	1.05 (0.90-1.22)	1.15 (0.83-1.61)	1.02 (0.90-1.15)
Autoimmune disorders	Other connective tissue disorder	1.00 (0.86-1.17)	0.99 (0.93-1.07)	0.95 (0.85-1.07)	0.97 (0.91-1.03)	0.91 (0.80-1.03)	0.97 (0.91-1.04)	0.95 (0.85-1.06)	0.98 (0.92-1.04)
Cancers	Urologic cancer	0.82 (0.67-1.00)	0.97 (0.91-1.05)	0.87 (0.73-1.03)	0.96 (0.90-1.03)	0.82 (0.66-1.01)	0.94 (0.87-1.02)	0.87 (0.73-1.04)	0.96 (0.90-1.03)
Cancers	Leukemia and lymphoma	0.94 (0.79-1.11)	1.00 (0.94-1.07)	0.94 (0.80-1.09)	1.02 (0.96-1.08)	1.04 (0.87-1.25)	1.04 (0.97-1.12)	0.91 (0.77-1.07)	1.01 (0.95-1.07)
Cancers	Cancer (any type)	0.99 (0.95-1.03)	1.00 (0.98-1.02)	0.97 (0.94-1.01)	0.99 (0.98-1.01)	0.98 (0.94-1.02)	1.00 (0.98-1.02)	0.97 (0.93-1.01)	1.00 (0.98-1.01)
Cancers	Thyroid cancer	0.87 (0.72-1.04)	0.96 (0.89-1.04)	0.86 (0.73-1.01)	0.95 (0.89-1.02)	0.89 (0.73-1.08)	0.97 (0.90-1.04)	0.85 (0.72-1.00)	0.95 (0.89-1.02)
Cardiovascular	Stroke	1.08 (1.00-1.16)	1.01 (0.98-1.04)	1.08 (1.00-1.16)	1.01 (0.98-1.04)	1.07 (1.00-1.16)	1.01 (0.98-1.05)	1.07 (0.99-1.15)	1.01 (0.98-1.04)
Mental health	Bipolar Disorder	0.98 (0.92-1.06)	0.98 (0.95-1.01)	0.97 (0.90-1.04)	0.98 (0.96-1.01)	0.97 (0.90-1.04)	0.99 (0.96-1.01)	0.97 (0.90-1.03)	0.98 (0.96-1.01)
Neurodegenerative	Early-onset Parkinson's and secondary parkinsonsism	0.82 (0.62-1.09)	0.91 (0.82-1.01)	0.86 (0.68-1.08)	0.92 (0.84-1.00)	0.94 (0.74-1.21)	0.94 (0.85-1.04)	0.87 (0.69-1.10)	0.93 (0.85-1.02)
Others	Obesity	1.01 (0.93-1.09)	0.98 (0.95-1.02)	0.98 (0.92-1.04)	0.97 (0.95-1.00)	0.98 (0.93-1.04)	0.98 (0.96-1.01)	0.98 (0.92-1.04)	0.98 (0.95-1.00)
Mortality	Death before age 65, any cause	-	-	1.06 (0.99-1.14)	1.01 (0.98-1.03)	1.06 (0.97-1.16)	1.03 (0.99-1.06)	1.05 (0.97-1.13)	1.01 (0.98-1.03)
Cancers	Male infertility	0.84 (0.39-1.82)	0.82 (0.59-1.15)	1.05 (0.69-1.62)	1.02 (0.86-1.20)	0.93 (0.56-1.55)	1.01 (0.82-1.25)	1.06 (0.69-1.63)	1.02 (0.86-1.21)
Reproductive	Prostate cancer	1.04 (0.89-1.21)	1.01 (0.96-1.07)	1.04 (0.91-1.18)	1.02 (0.98-1.07)	1.02 (0.88-1.19)	1.04 (0.98-1.10)	1.06 (0.93-1.21)	1.04 (0.99-1.09)
Respiratory	Chronic rhinitis, nasopharyngitis, and pharyngitis	0.88 (0.76-1.02)	0.94 (0.88-1.00)	0.85 (0.77-0.95)	0.97 (0.93-1.02)	0.85 (0.77-0.95)	0.98 (0.93-1.02)	0.86 (0.77-0.95)	0.98 (0.93-1.02)
Others	Liver disease, cirrhosis, and other liver conditions (non-viral)	0.99 (0.93-1.06)	0.98 (0.95-1.00)	0.92 (0.86-0.99)	0.95 (0.93-0.97)	0.92 (0.85-0.99)	0.95 (0.93-0.97)	0.91 (0.85-0.98)	0.95 (0.92-0.97)

Appendix F. Comparison of census tracts that are food deserts to pollution exposure maps

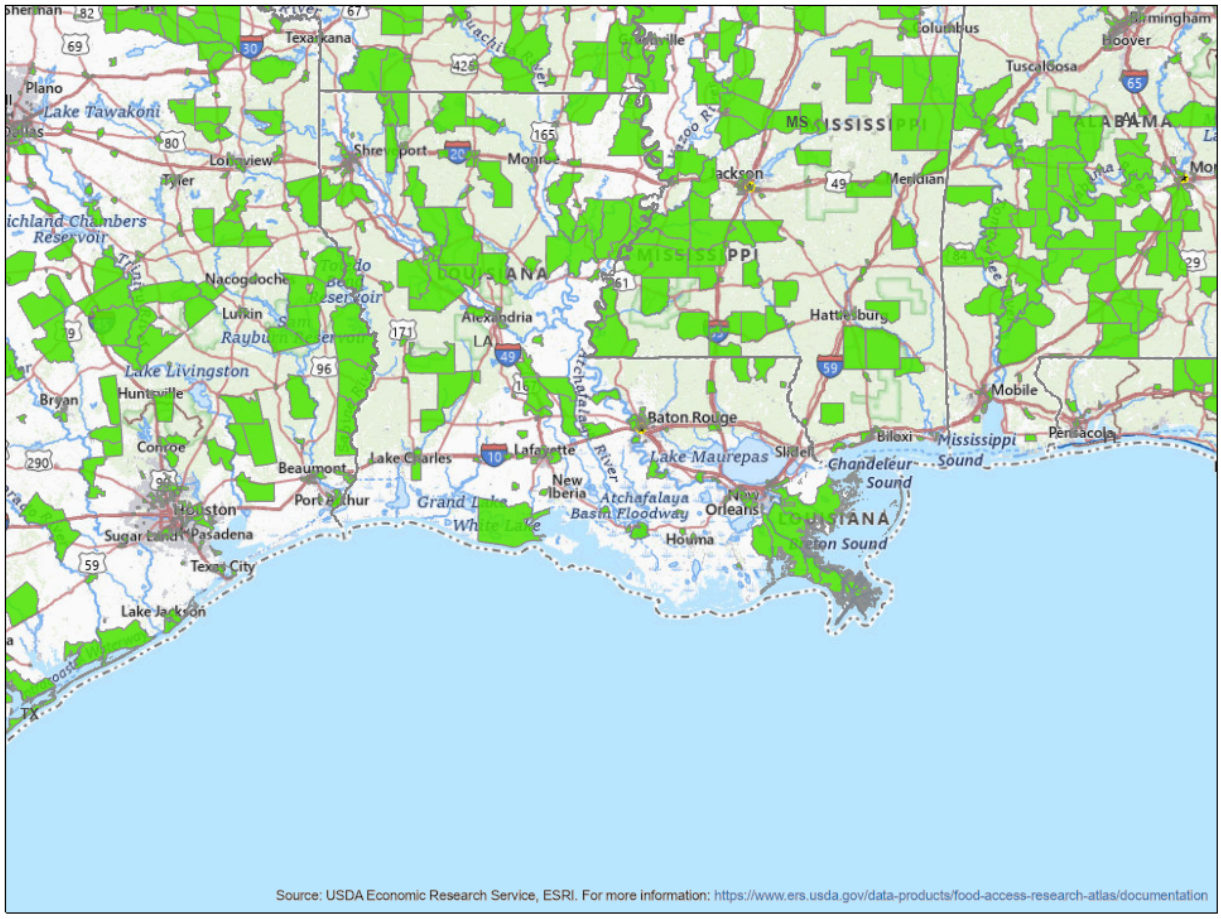


Figure F1a. Map of census tracts designated as low-income and low access to grocery stores using the USDA Food Access Research Atlas definition of “low-income, low-access” (often referred to as food deserts). These tracts meet income-based criteria and have at least 500 residents or ≥33% of the population living farther than 1 mile (urban) or 10 miles (rural) from the nearest supermarket or large grocery store.

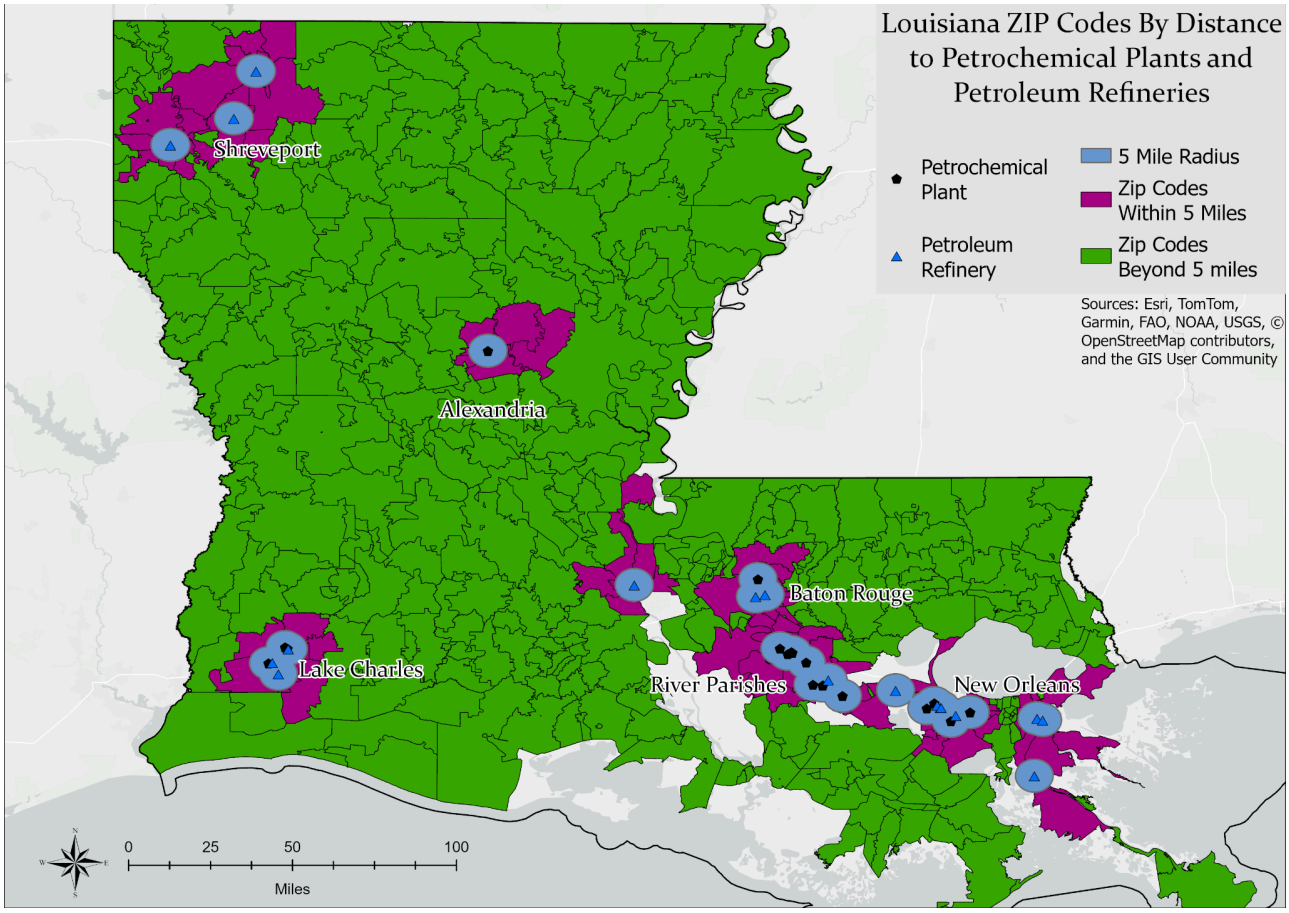


Figure 3a. Map of locations of 40 petrochemical facilities and petroleum refineries in Louisiana in 2017 LDEQ data, with ZIP codes by distance to those 40 petrochemical plants and petroleum refineries.

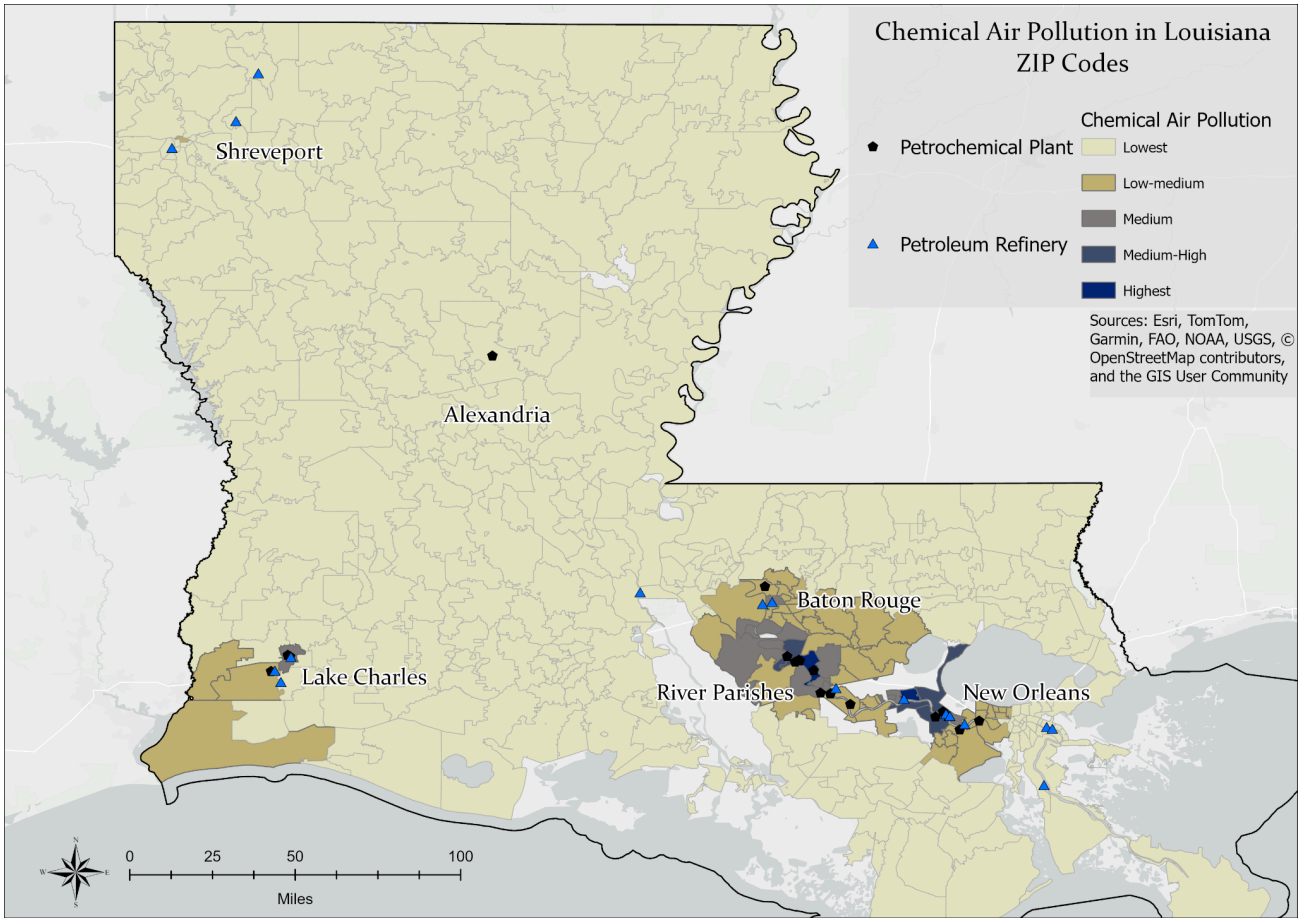


Figure 4a. Louisiana ZIP code tabulation areas (ZCTAs) by estimated toxicity-weighted concentration of chemicals in the air ("Chemical Air Pollution"). Lighter colors indicate lower pollution levels. From 2016 RSEI geographic microdata. Air pollution concentration categories, measured in $\mu g/m^3$, are: lowest: 0-23,963; low-medium: 23,964-80,751; medium: 80,752-203,654; medium-high: 203,655-363,305; highest: 363,306-783,144.

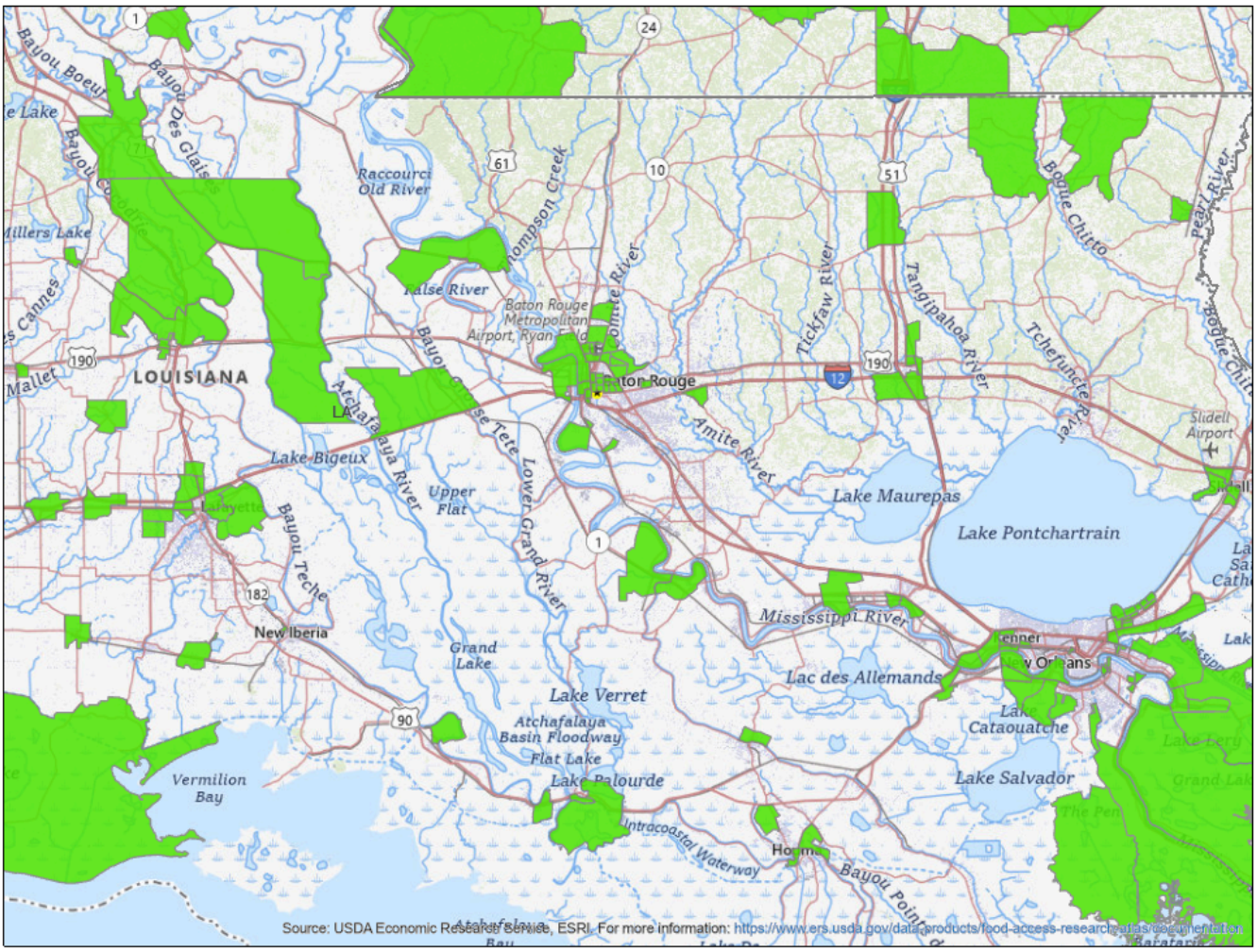


Figure F2a. Inset of Cancer Alley region. Map of census tracts designated as low-income and low access to grocery stores using the USDA Food Access Research Atlas definition of “low-income, low-access” (often referred to as food deserts). These tracts meet income-based criteria and have at least 500 residents or ≥33% of the population living farther than 1 mile (urban) or 10 miles (rural) from the nearest supermarket or large grocery store.

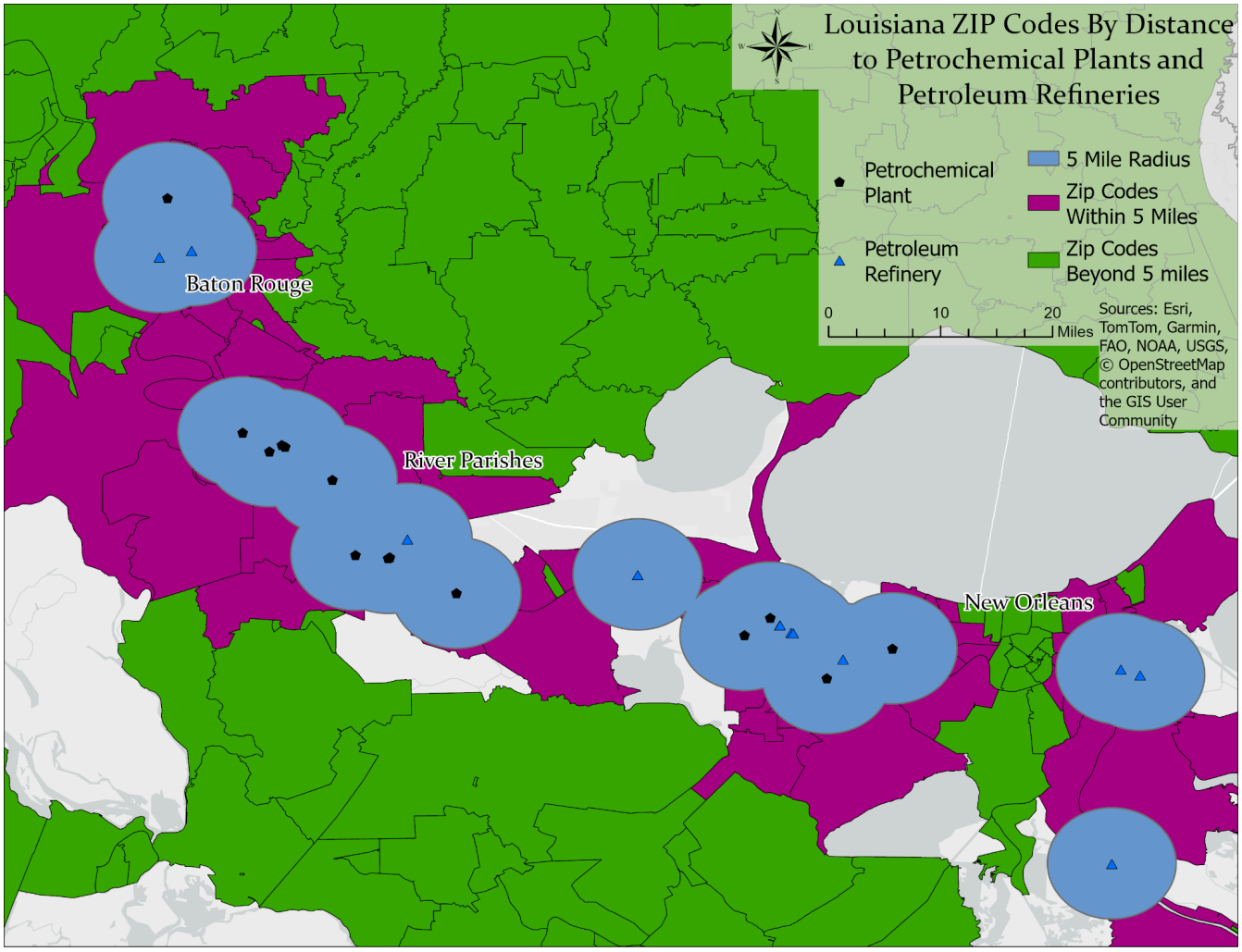


Figure 3b. Inset of Cancer Alley region. Map of locations of 40 petrochemical facilities and petroleum refineries in Louisiana in 2017 LDEQ data, with ZIP codes by distance to those 40 petrochemical plants and petroleum refineries.

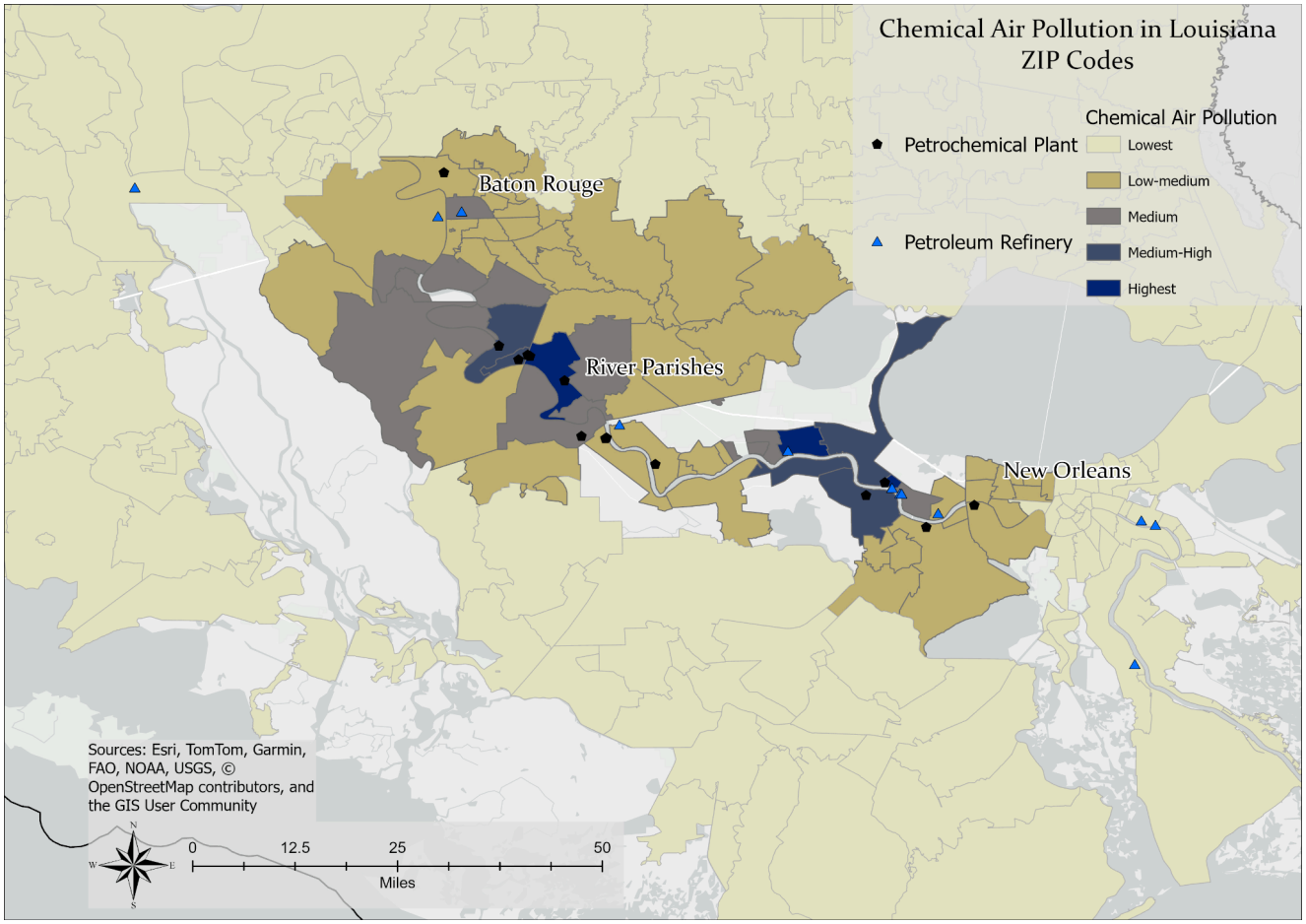


Figure 4b. Inset of Cancer Alley region. Louisiana ZIP code tabulation areas (ZCTAs) by estimated toxicity-weighted concentration of chemicals in the air ("Chemical Air Pollution"). Lighter colors indicate lower pollution levels. From 2016 RSEI geographic microdata. Air pollution concentration categories, measured in $\mu g/m^3$, are: lowest: 0-23,963; low-medium: 23,964-80,751; medium: 80,752-203,654; medium-high: 203,655-363,305; highest: 363,306-783,144.

Appendix G. Crosstabs of case counts using adapted CCW algorithm case definitions compared to less-restrictive case definition for key health outcomes

Table G1. Cross-tabulation of nutritional anemia case identification methods among children.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	792118	0	792118
	Percent	98.48	0	98.48
	Row Pct	100	0	
	Col Pct	98.94	0	
Met case defintion	Frequency	8471	3768	12239
	Percent	1.05	0.47	1.52
	Row Pct	69.21	30.79	
	Col Pct	1.06	100	
Total	Frequency	800589	3768	804357
	Percent	99.53	0.47	100

Table G2. Cross-tabulation of learning disability case identification methods among children.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	735861	0	735861
	Percent	91.48	0	91.48
	Row Pct	100	0	
	Col Pct	96.79	0	
Met case defintion	Frequency	24380	44116	68496
	Percent	3.03	5.48	8.52
	Row Pct	35.59	64.41	
	Col Pct	3.21	100	
Total	Frequency	760241	44116	804357
	Percent	94.52	5.48	100

Table G3. Cross-tabulation of dermatitis and eczema case identification methods among children.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	606440	0	606440
	Percent	75.39	0	75.39
	Row Pct	100	0	
	Col Pct	84.46	0	
Met case defintion	Frequency	111598	86319	197917
	Percent	13.87	10.73	24.61
	Row Pct	56.39	43.61	
	Col Pct	15.54	100	
Total	Frequency	718038	86319	804357
	Percent	89.27	10.73	100

Table G4. Cross-tabulation of early puberty case identification methods among girls.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	392767	0	392767
	Percent	99.54	0	99.54
	Row Pct	100	0	
	Col Pct	99.54	0	
Met case defintion	Frequency	1806	19	1825
	Percent	0.46	0	0.46
	Row Pct	98.96	1.04	
	Col Pct	0.46	100	
Total	Frequency	394573	19	394592
	Percent	100	0	100

Table G5. Cross-tabulation of breast cancer case identification methods among non-smoking women.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	491055	0	491055
	Percent	99.22	0	99.22
	Row Pct	100	0	
	Col Pct	99.87	0	
Met case defintion	Frequency	652	3213	3865
	Percent	0.13	0.65	0.78
	Row Pct	16.87	83.13	
	Col Pct	0.13	100	
Total	Frequency	491707	3213	494920
	Percent	99.35	0.65	100

Table G6. Cross-tabulation of uterine fibroid case identification methods among non-smoking women.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	480350	0	480350
	Percent	97.06	0	97.06
	Row Pct	100	0	
	Col Pct	98.71	0	
Met case defintion	Frequency	6296	8274	14570
	Percent	1.27	1.67	2.94
	Row Pct	43.21	56.79	
	Col Pct	1.29	100	
Total	Frequency	486646	8274	494920
	Percent	98.33	1.67	100

Table G7. Cross-tabulation of preeclampsia and eclampsia case identification methods among non-smoking women who have been pregnant.

		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	88897	0	88897
	Percent	92.14	0	92.14
	Row Pct	100	0	
	Col Pct	98.7	0	
Met case defintion	Frequency	1169	6416	7585
	Percent	1.21	6.65	7.86
	Row Pct	15.41	84.59	
	Col Pct	1.3	100	
Total	Frequency	90066	6416	96482
	Percent	93.35	6.65	100

Table G8. Cross-tabulation of miscarriage case identification methods among non-smoking women who have been pregnant.

		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	87533	0	87533
	Percent	90.72	0	90.72
	Row Pct	100	0	
	Col Pct	95.12	0	
Met case defintion	Frequency	4491	4458	8949
	Percent	4.65	4.62	9.28
	Row Pct	50.18	49.82	
	Col Pct	4.88	100	
Total	Frequency	92024	4458	96482
	Percent	95.38	4.62	100

Table G9. Cross-tabulation of ectopic pregnancy case identification methods among non-smoking women who have been pregnant.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	94123	0	94123
	Percent	97.55	0	97.55
	Row Pct	100	0	
	Col Pct	98.67	0	
Met case defintion	Frequency	1265	1094	2359
	Percent	1.31	1.13	2.45
	Row Pct	53.62	46.38	
	Col Pct	1.33	81.58	
Total	Frequency	95388	1094	96482
	Percent	98.87	1.13	100

Table G10. Cross-tabulation of nutritional anemia case identification methods among non-smoking adults.				
		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	776500	0	776500
	Percent	95.78	0	95.78
	Row Pct	100	0	
	Col Pct	98.1	0	
Met case defintion	Frequency	15003	19229	34232
	Percent	1.85	2.37	4.22
	Row Pct	43.83	56.17	
	Col Pct	1.9	100	
Total	Frequency	791503	19229	810732
	Percent	97.63	2.37	100

Table G11. Cross-tabulation of other acquired anemia case identification methods among non-smoking adults.

		Adapted CCW algorithm		
		Did not meet case definition	Diagnosed	Total
Less restrictive case definition				
Did not meet case definition	Frequency	760032	0	760032
	Percent	93.75	0	93.75
	Row Pct	100	0	
	Col Pct	97.39	0	
Met case defintion	Frequency	20399	30301	50700
	Percent	2.52	3.74	6.25
	Row Pct	40.23	59.77	
	Col Pct	2.61	100	
Total	Frequency	780431	30301	810732
	Percent	96.26	3.74	100

Appendix H. Number of diagnosis codes by individual among those with at least one diagnosis code

These bar charts show how many people were diagnosed with each health condition, and how often those diagnoses appeared in their medical records between 2017 and 2019. Each bar groups people based on how many times the condition was recorded during any type of healthcare visit (such as hospital stays or doctor visits). For most conditions, people had the diagnosis recorded more than once, and many had it recorded multiple times.

Figure H1. Number of diagnosis codes by individual, nutritional anemia

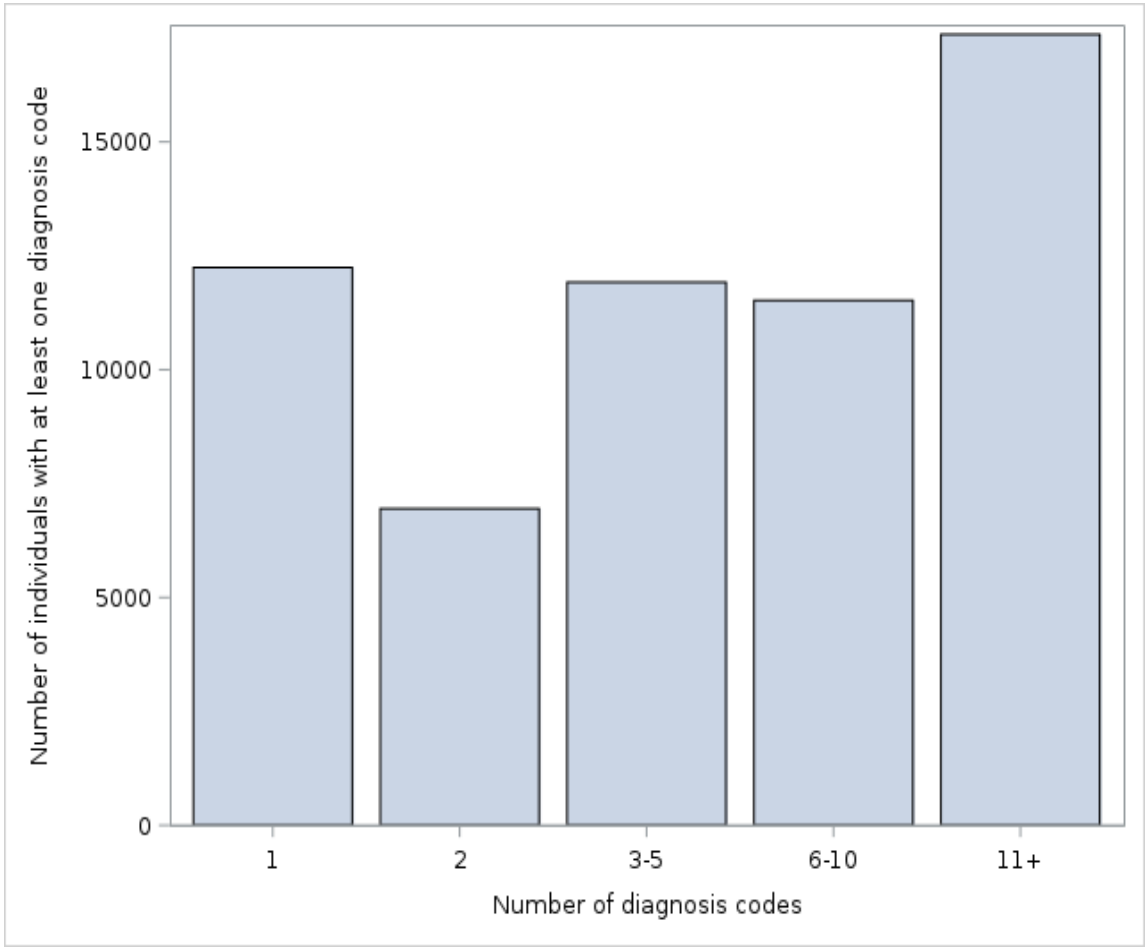


Figure H2. Number of diagnosis codes by individual, other acquired anemia

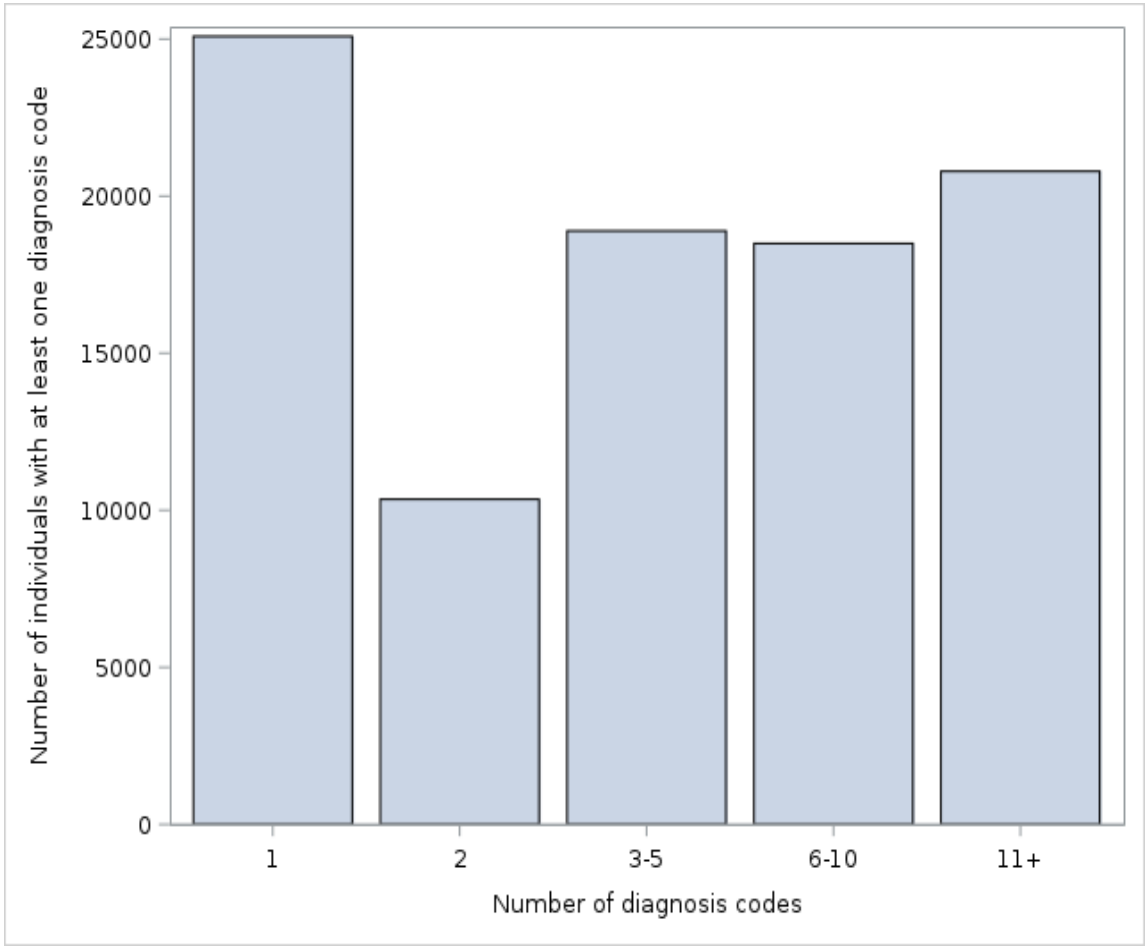


Figure H3. Number of diagnosis codes by individual, lupus

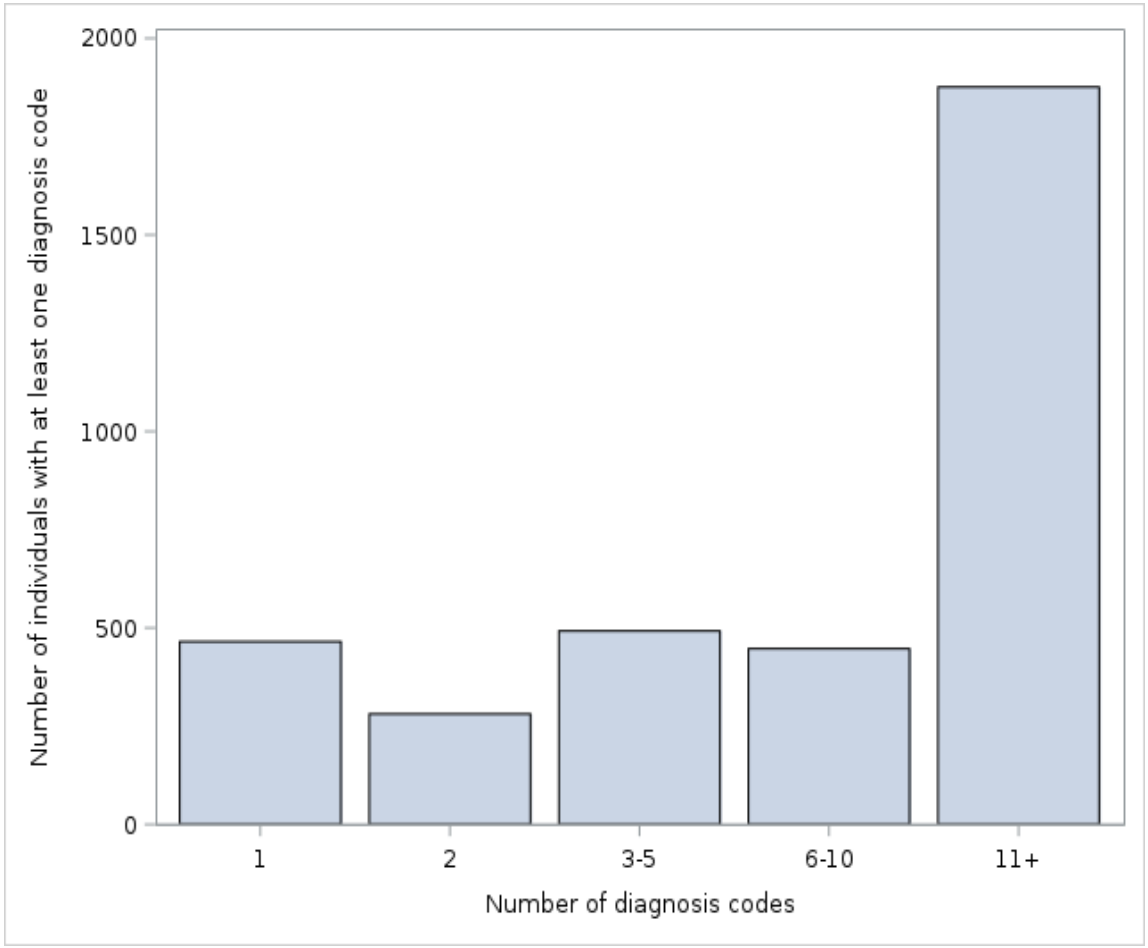


Figure H4. Number of diagnosis codes by individual, multiple sclerosis

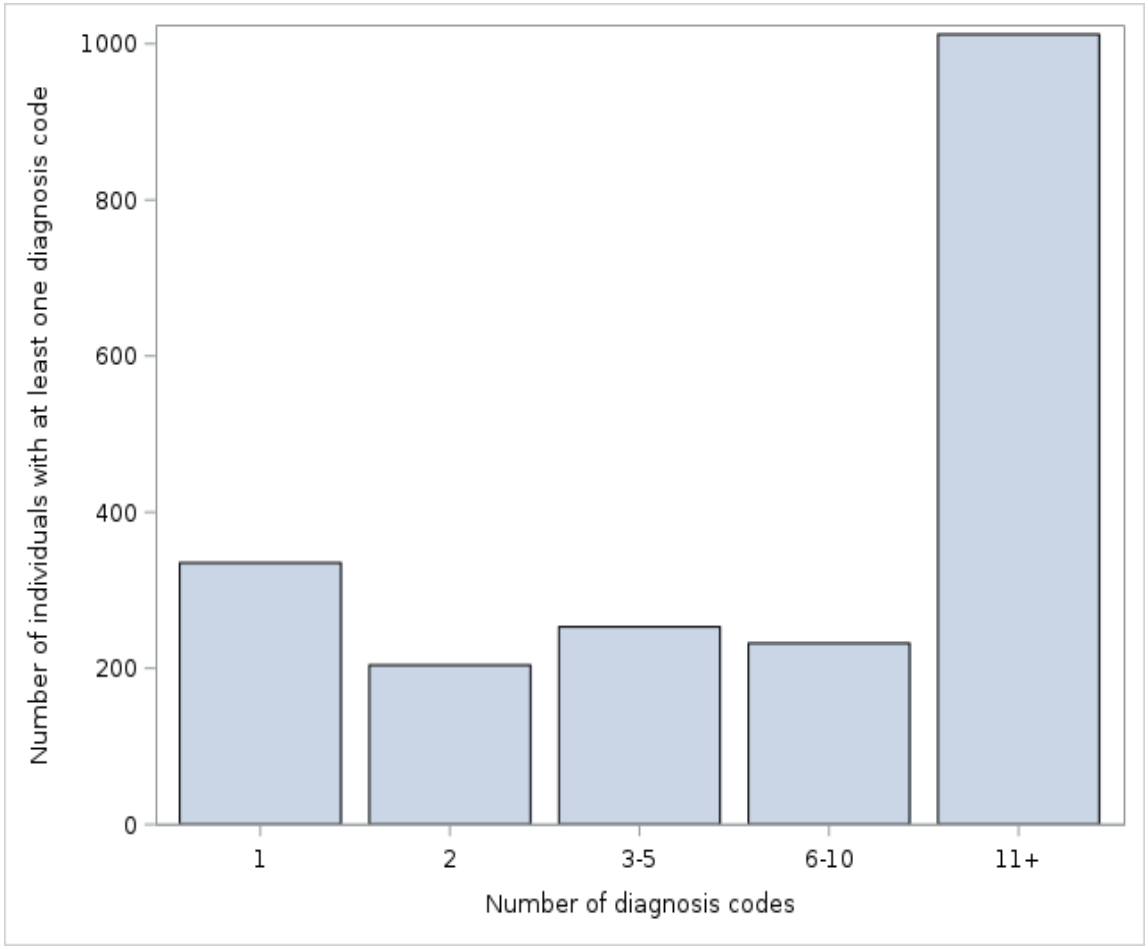


Figure H5. Number of diagnosis codes by individual, systemic sclerosis

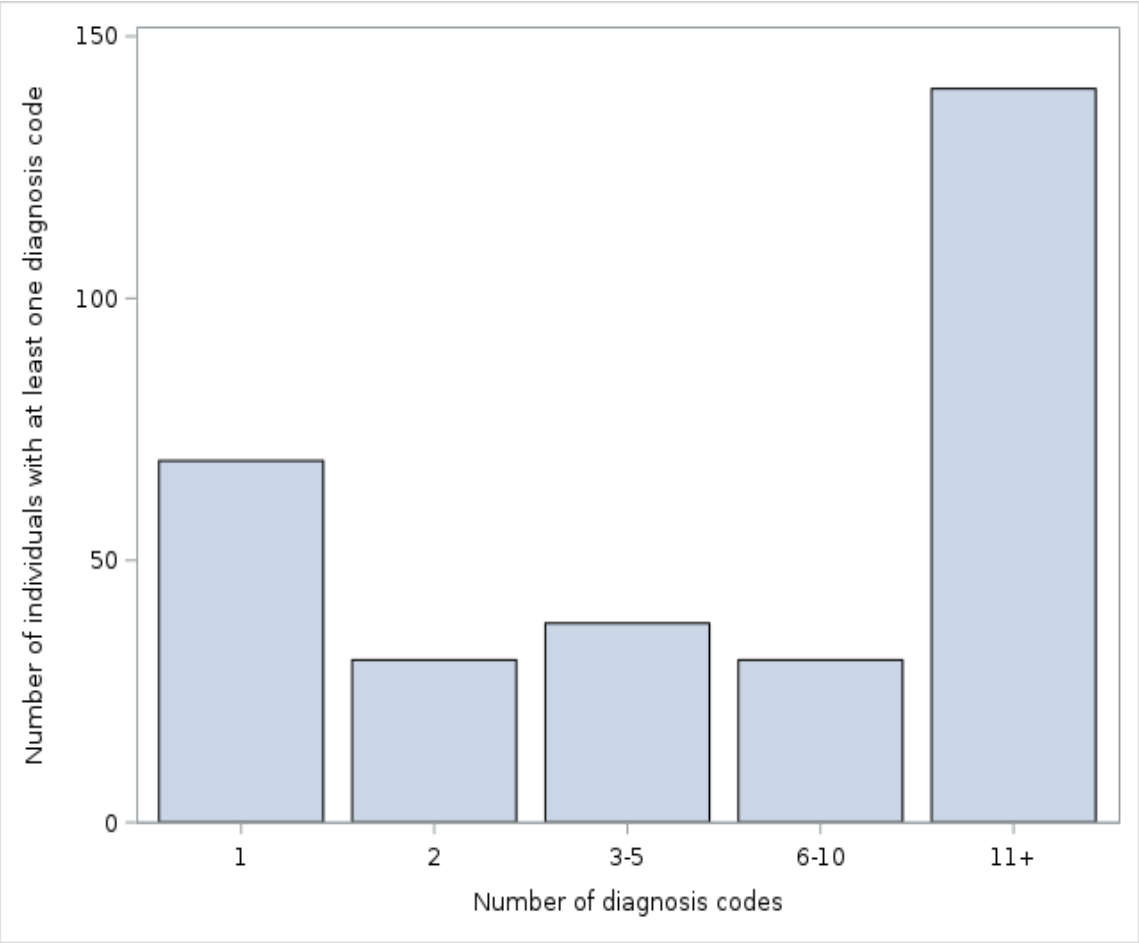


Figure H6. Number of diagnosis codes by individual, myositis

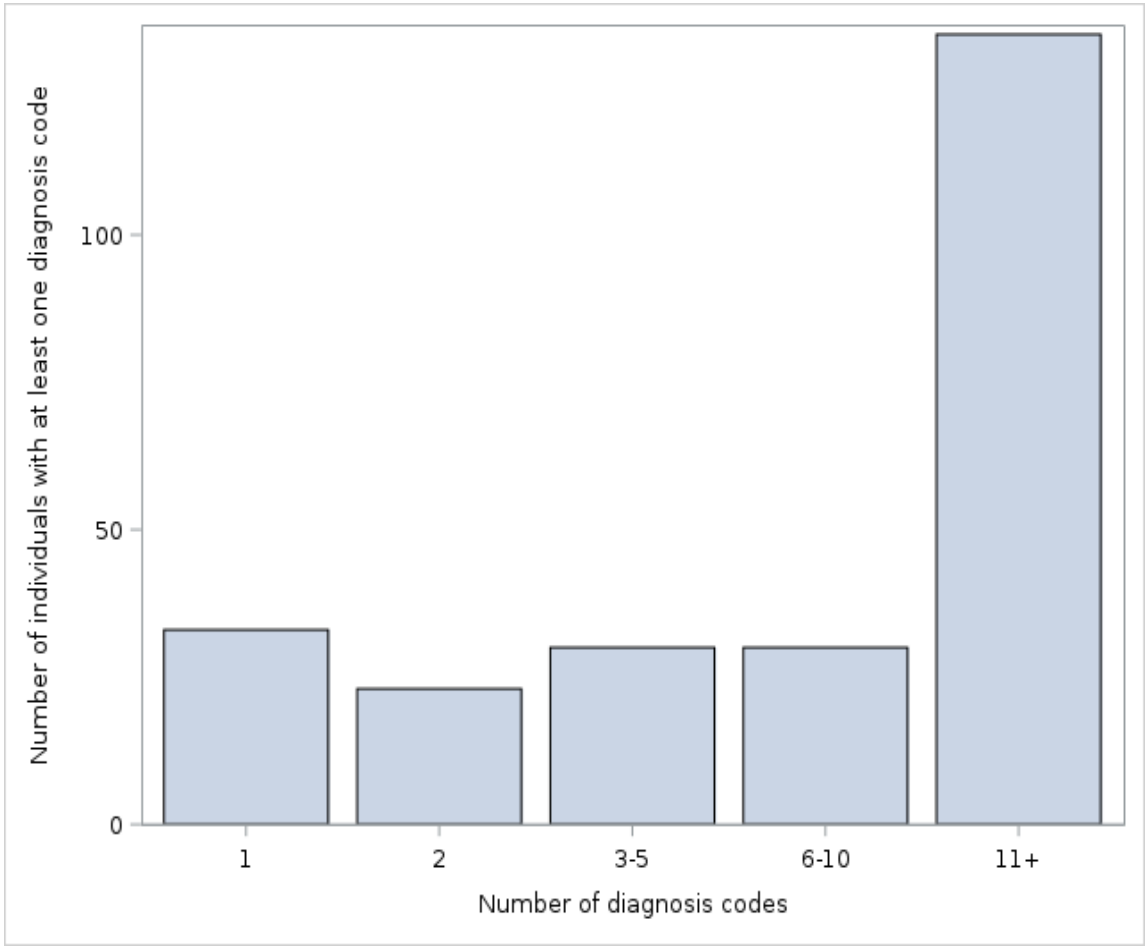


Figure H7. Number of diagnosis codes by individual, other connective tissue disorders

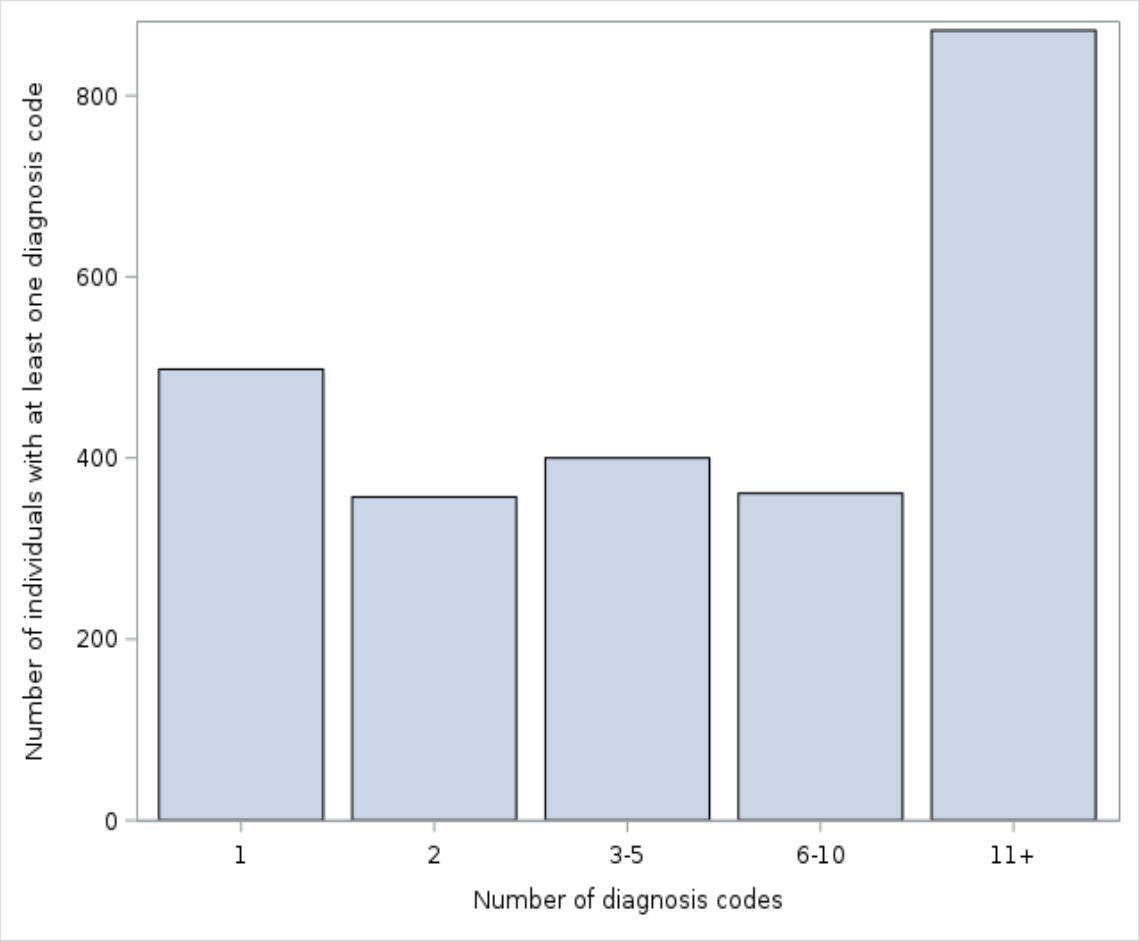


Figure H8. Number of diagnosis codes by individual, low birth weight

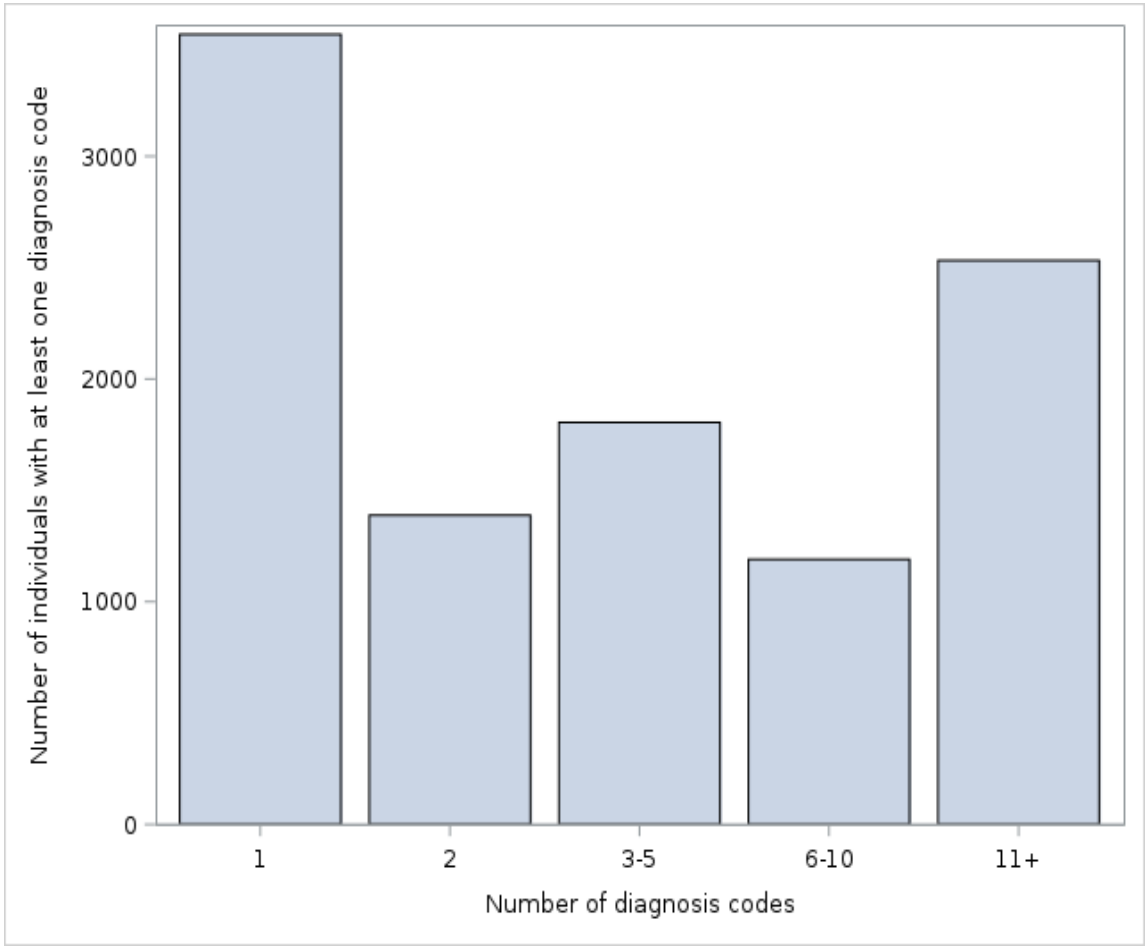


Figure H9. Number of diagnosis codes by individual, preterm birth

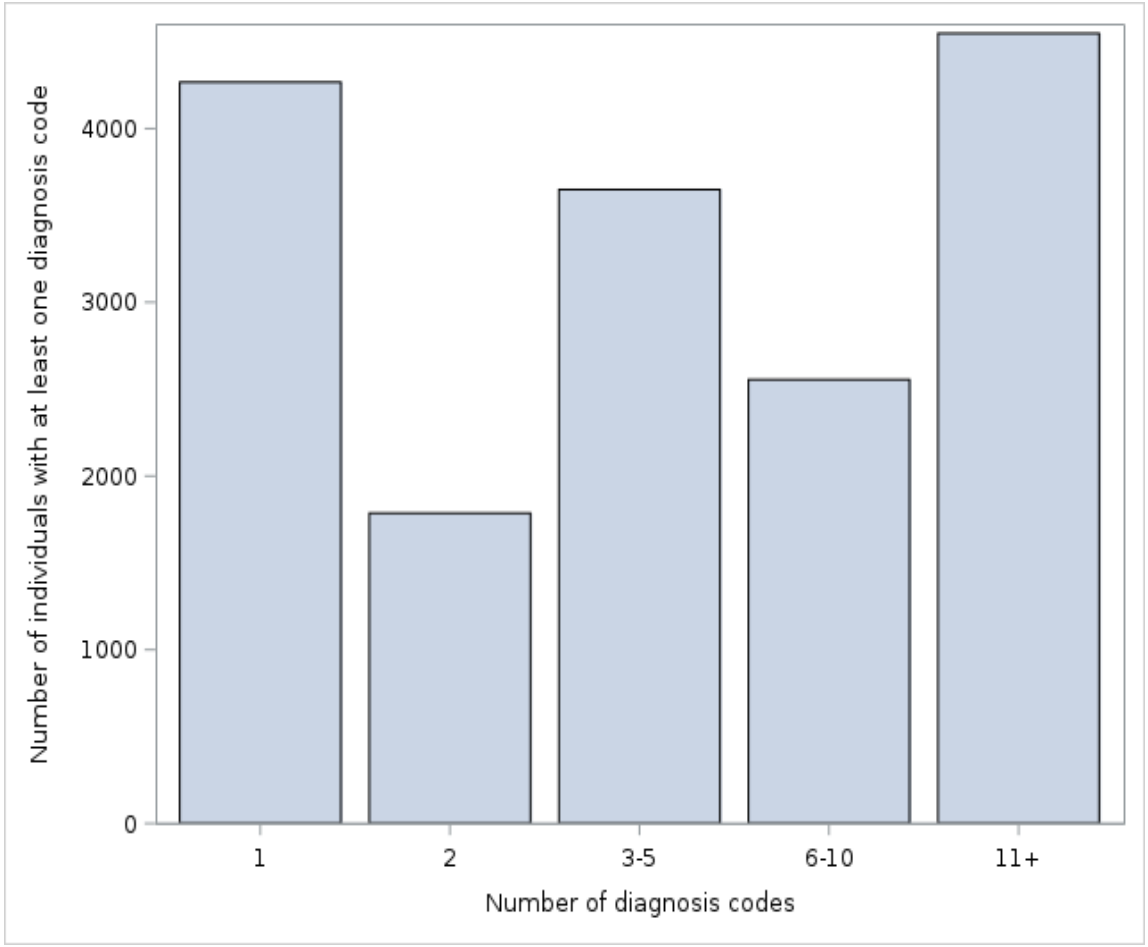


Figure H10. Number of diagnosis codes by individual, birth defects

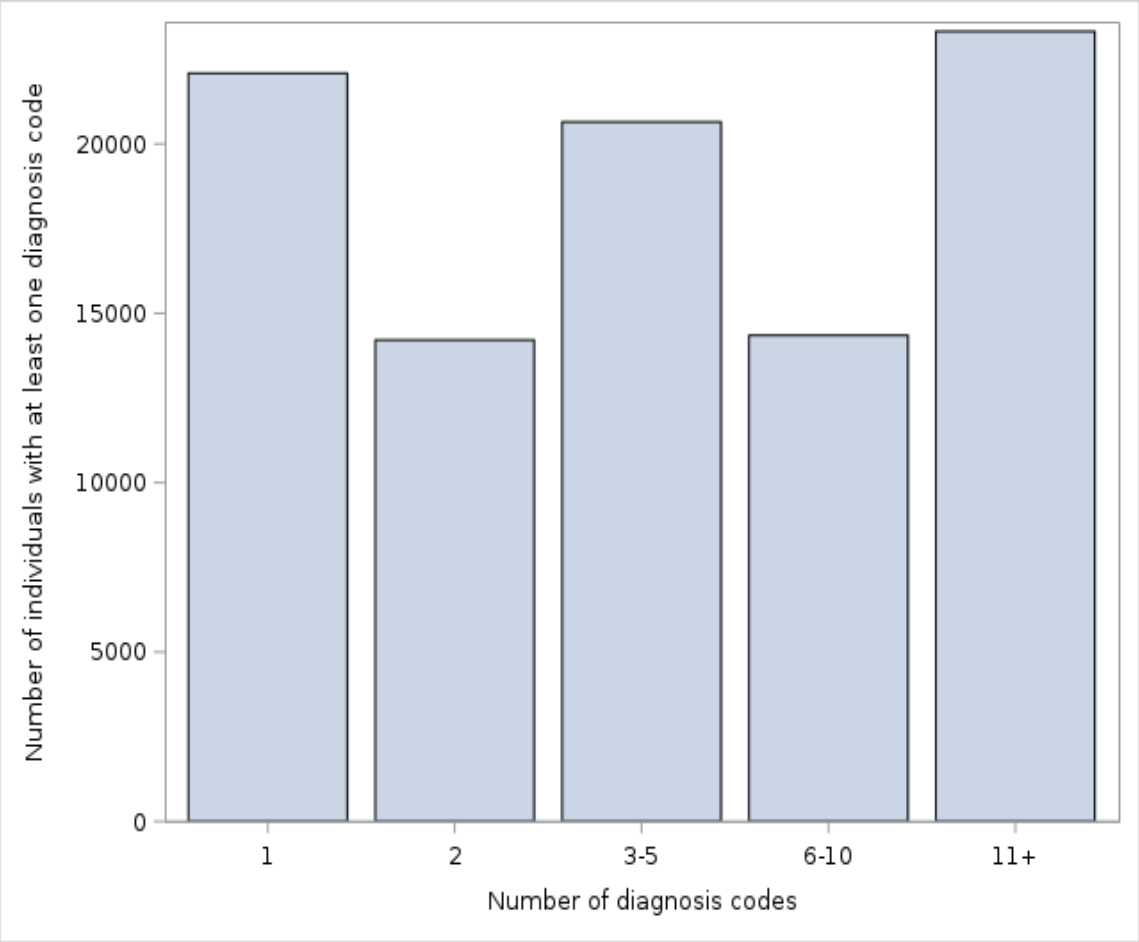


Figure H11. Number of diagnosis codes by individual, breast cancer

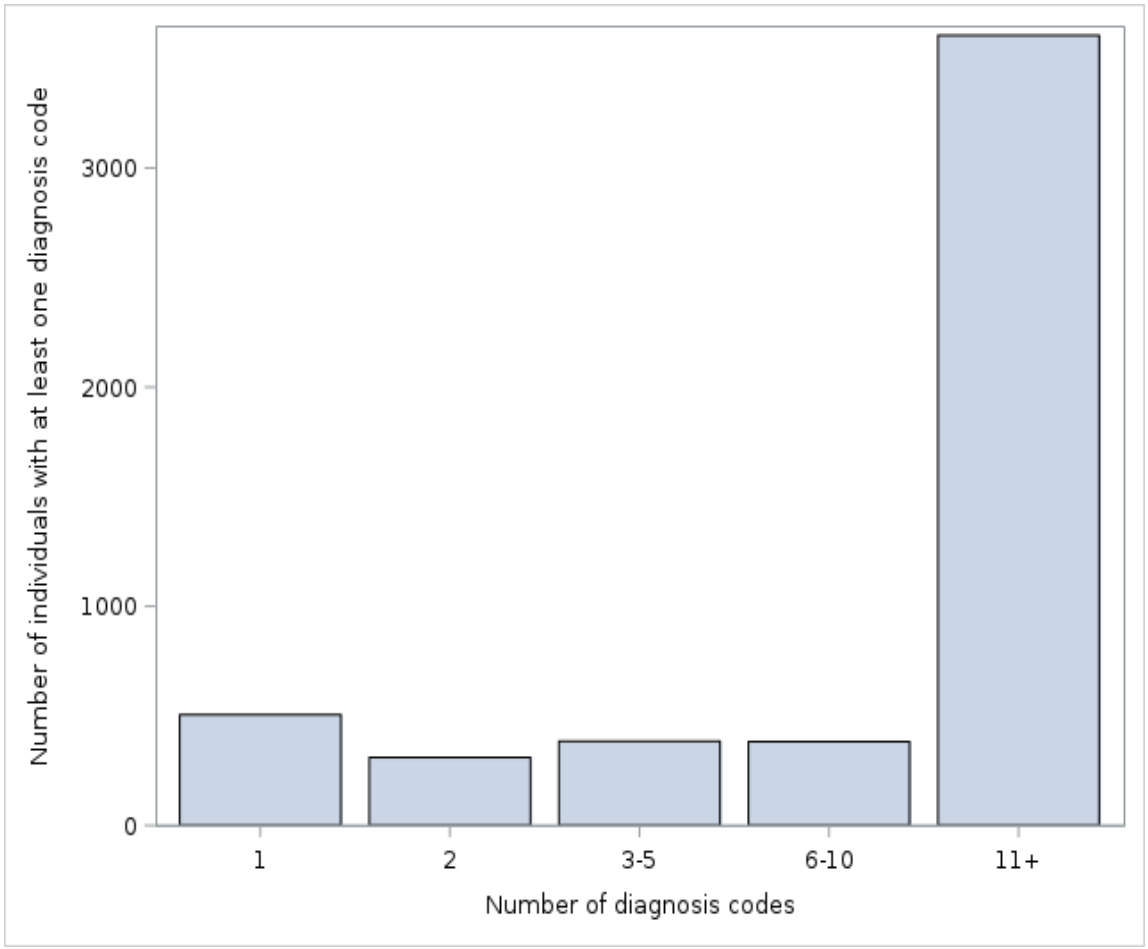


Figure H12. Number of diagnosis codes by individual, colorectal cancer

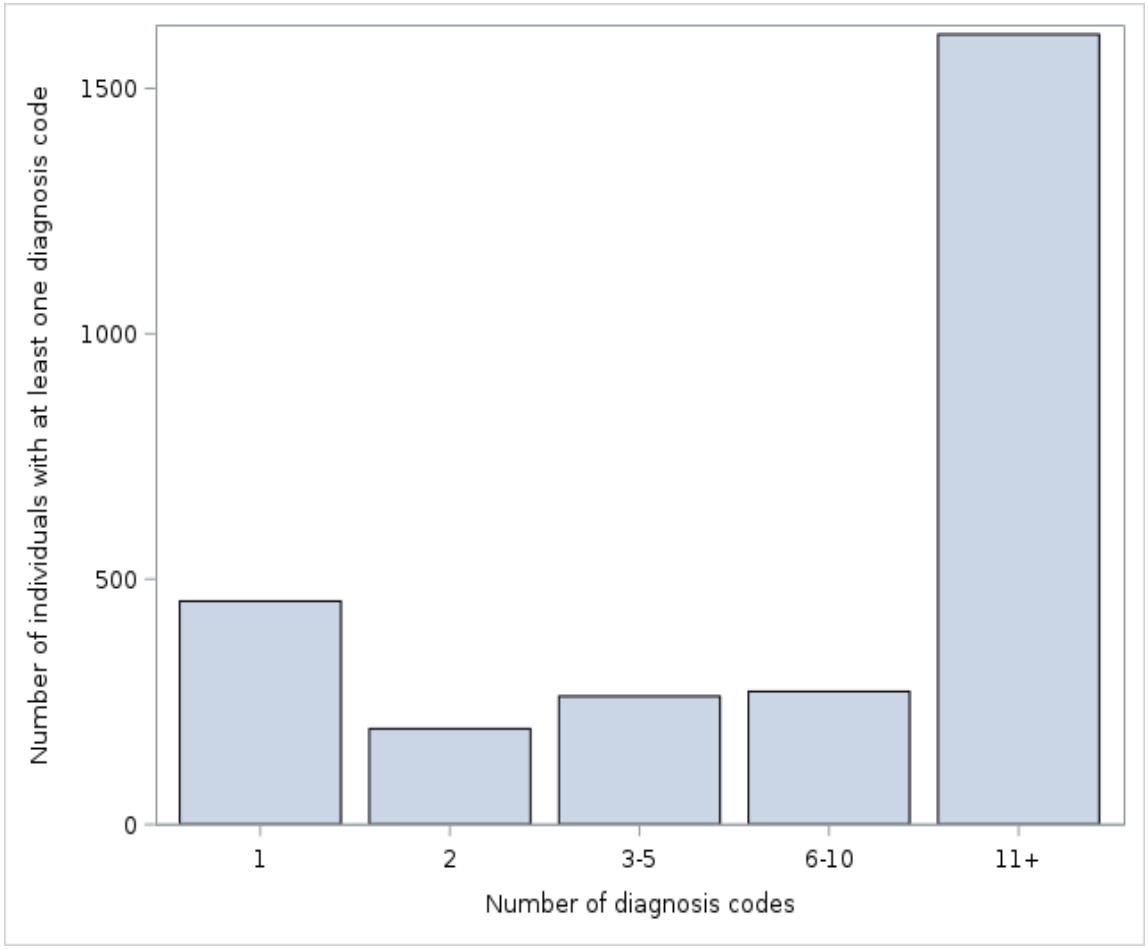


Figure H13. Number of diagnosis codes by individual, lung cancer

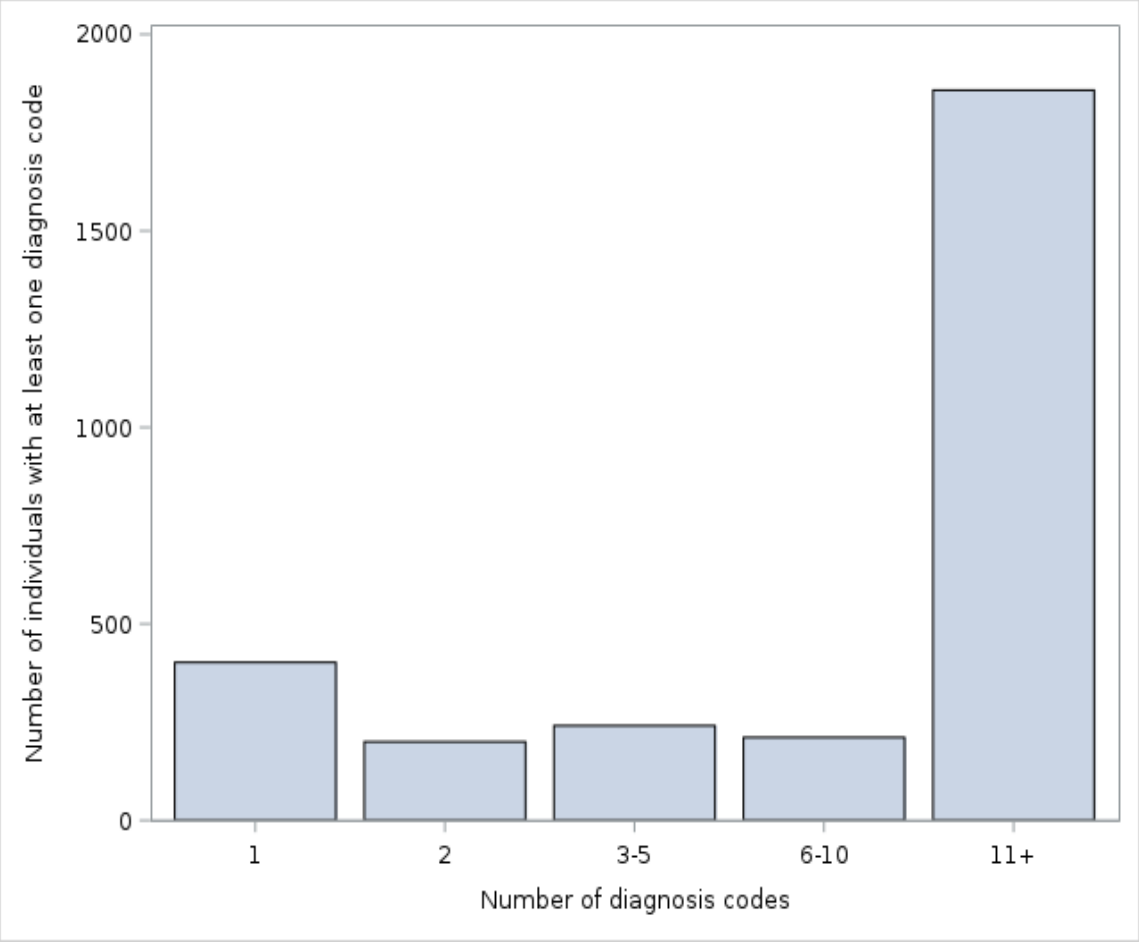


Figure H14. Number of diagnosis codes by individual, prostate cancer

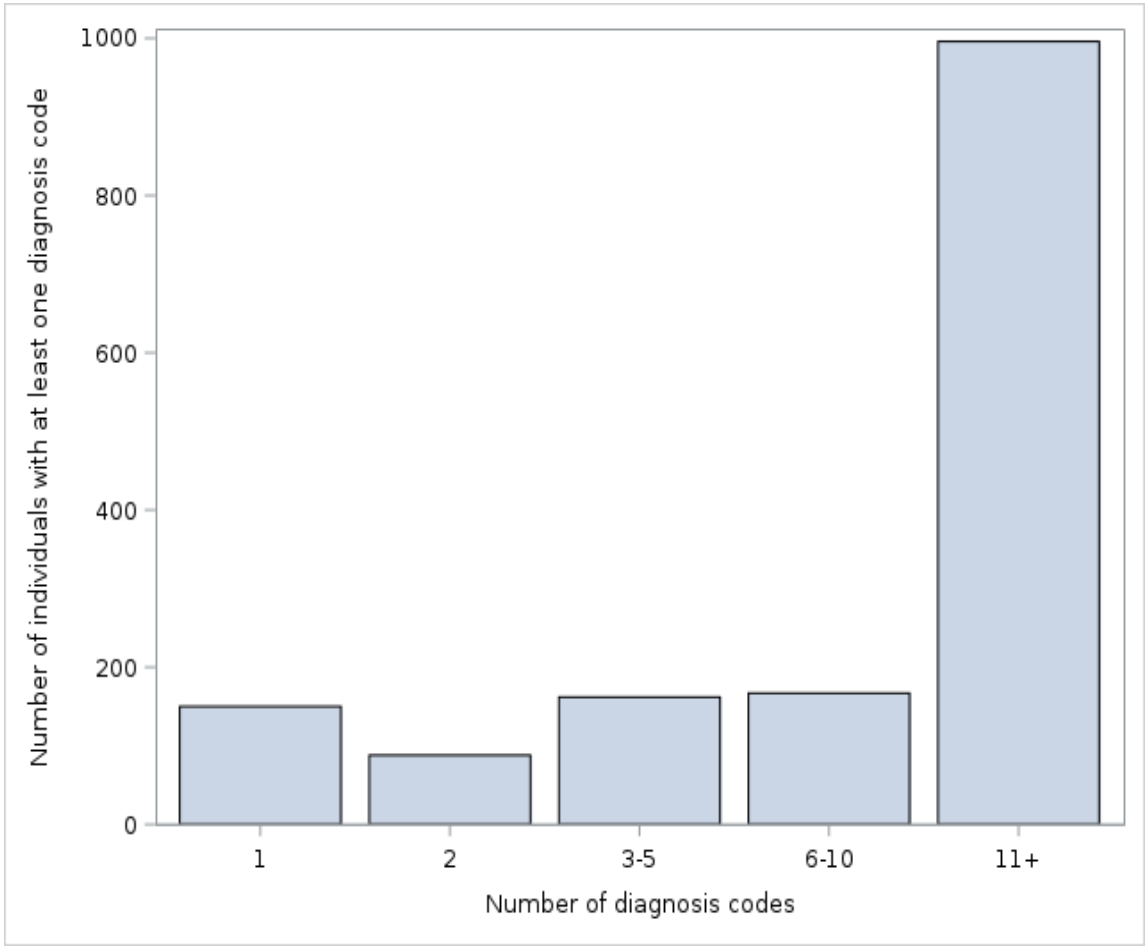


Figure H15. Number of diagnosis codes by individual, urologic cancer

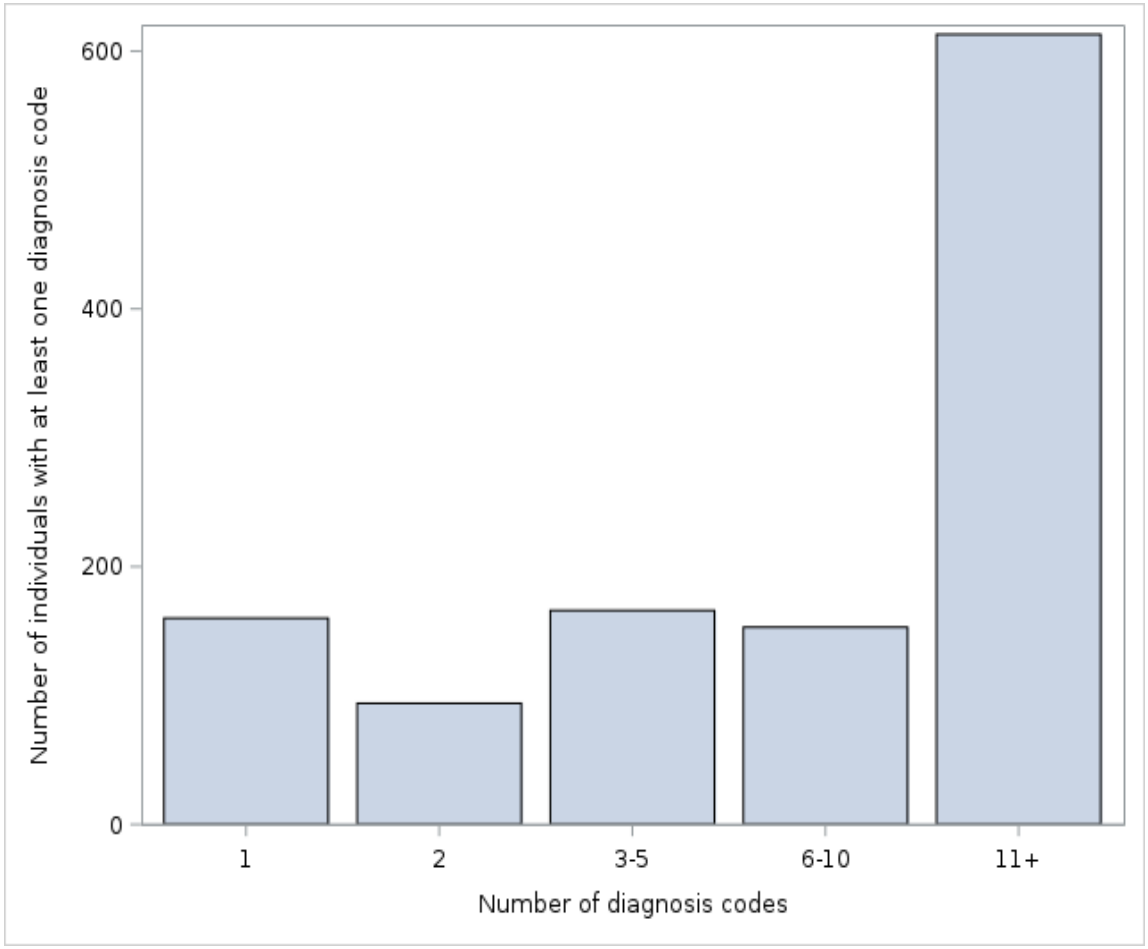


Figure H16. Number of diagnosis codes by individual, leukemia and lymphoma

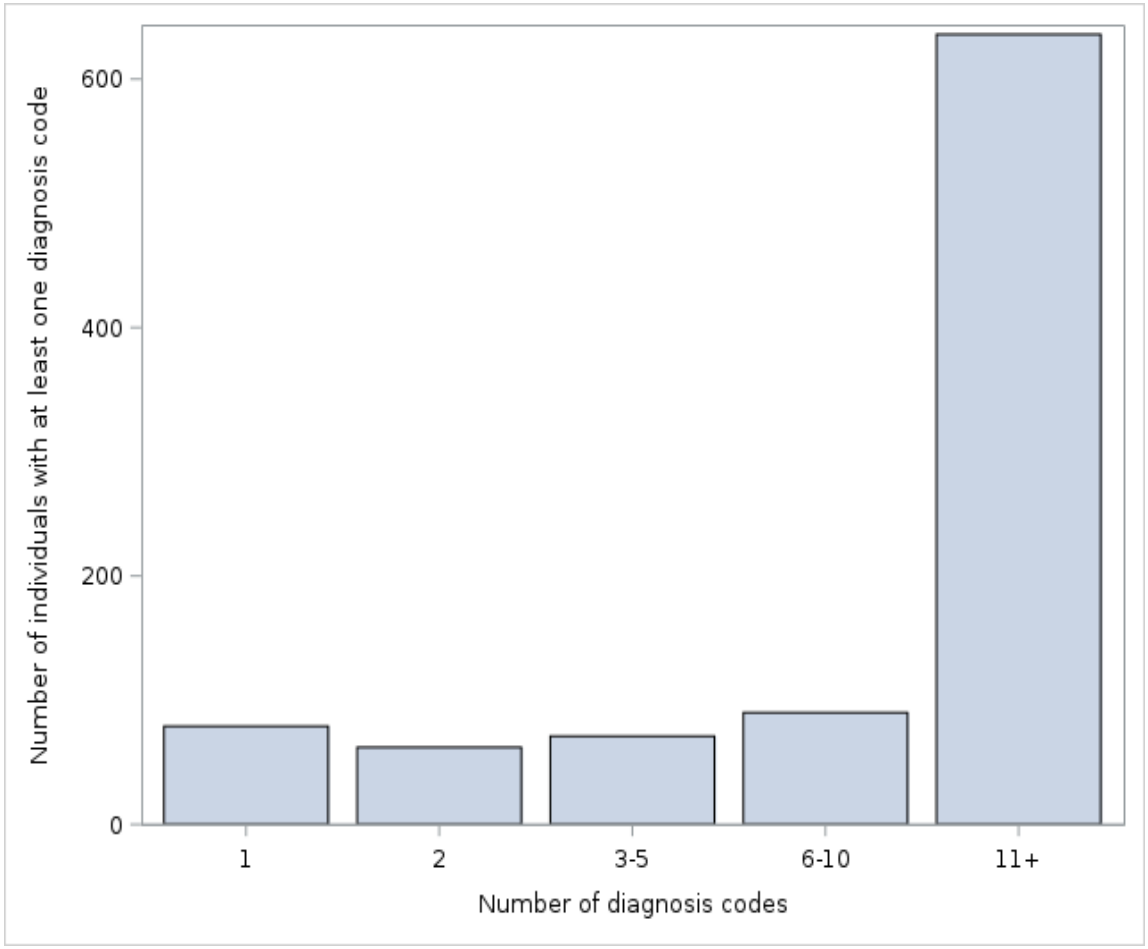


Figure H17. Number of diagnosis codes by individual, female reproductive cancers

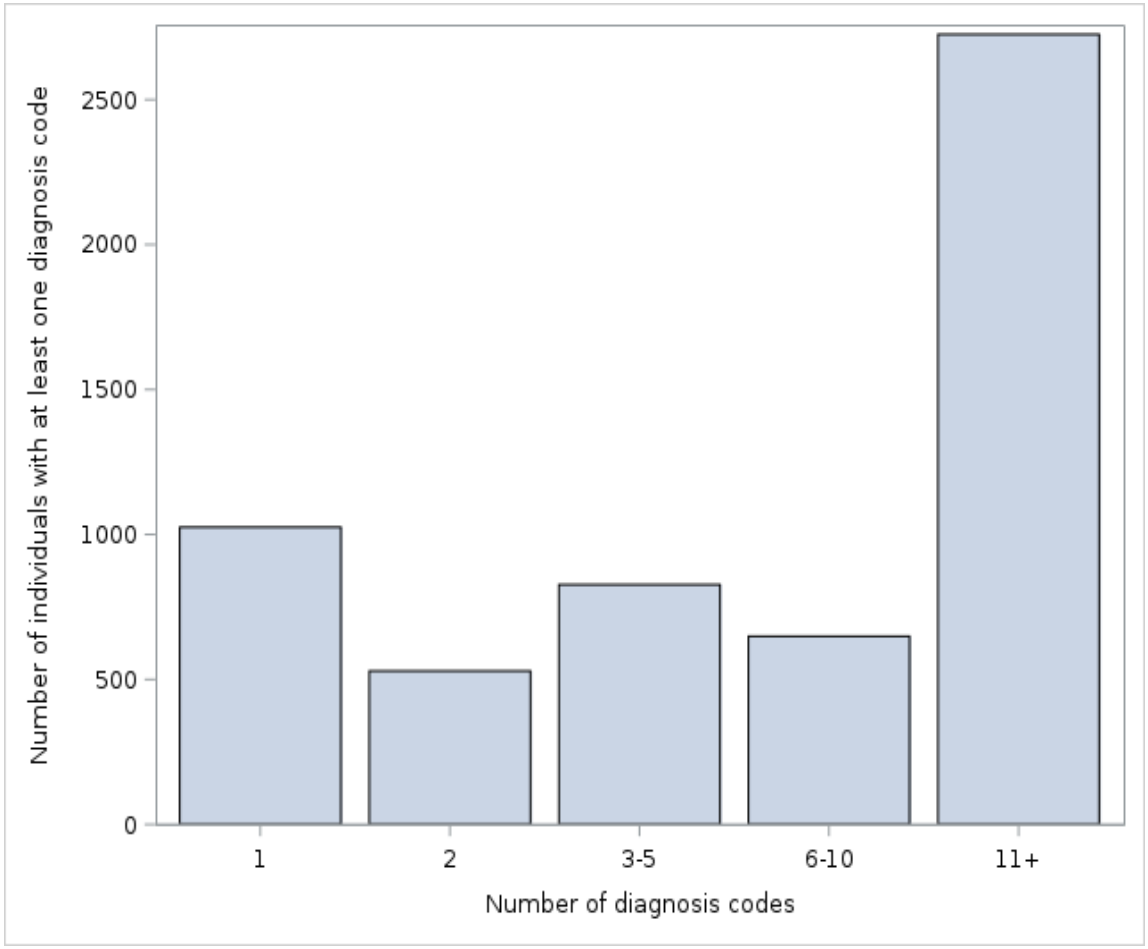


Figure H18. Number of diagnosis codes by individual, all cancers

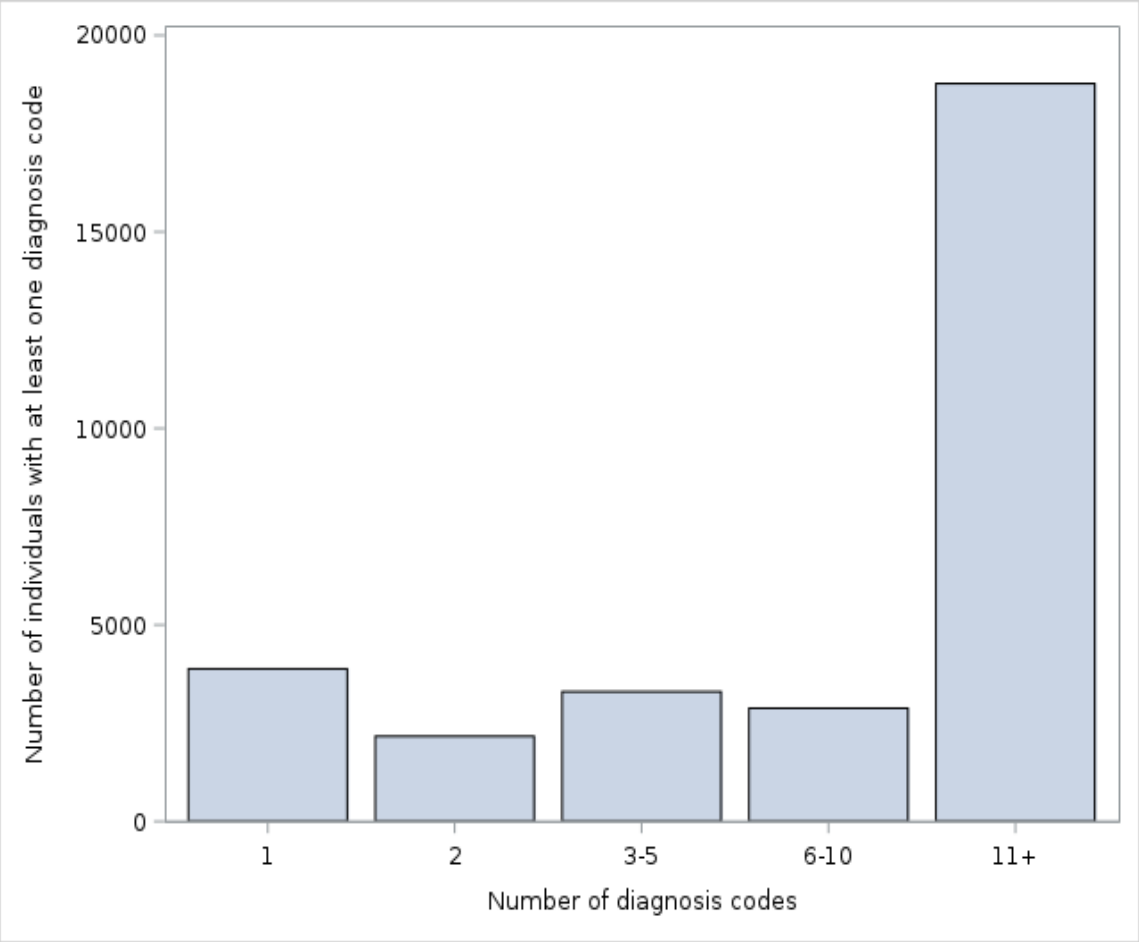


Figure H19. Number of diagnosis codes by individual, thyroid cancer

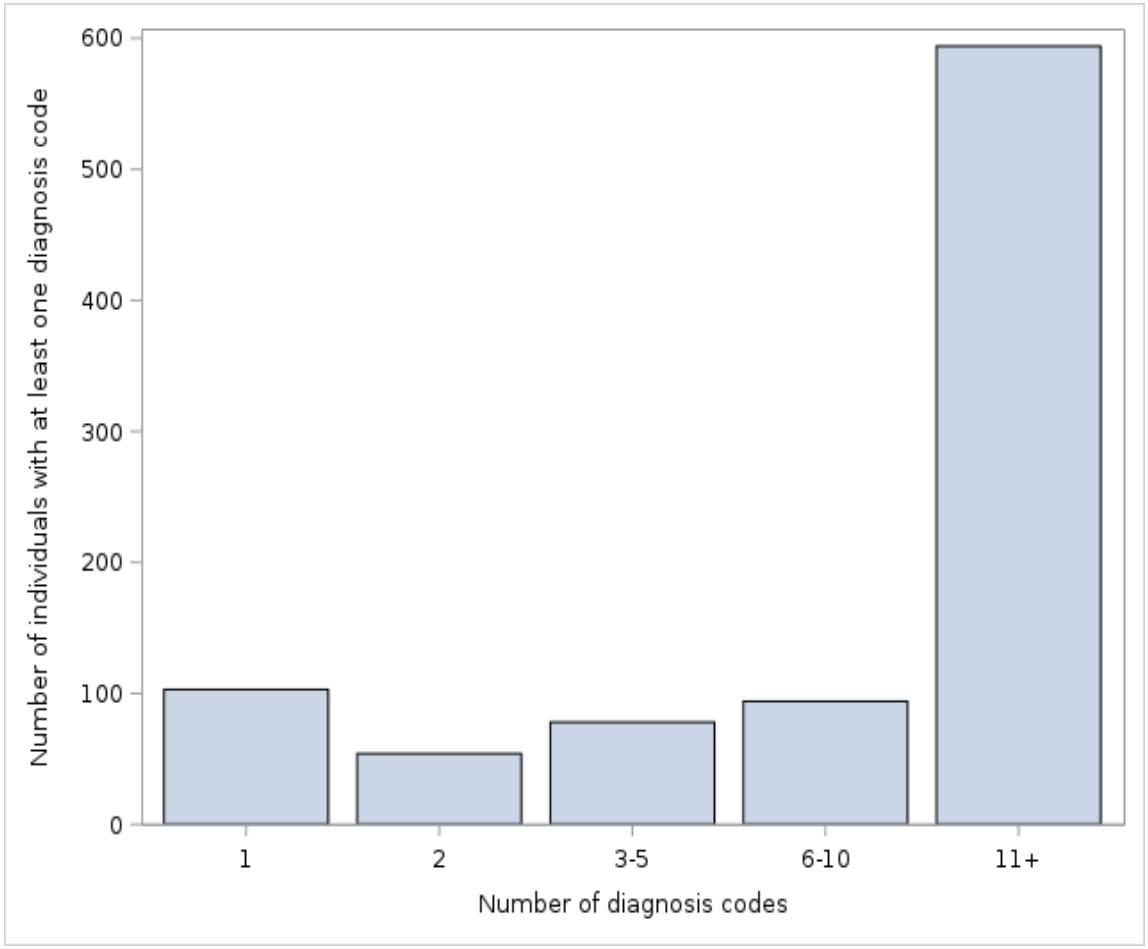


Figure H20. Number of diagnosis codes by individual, heart attack

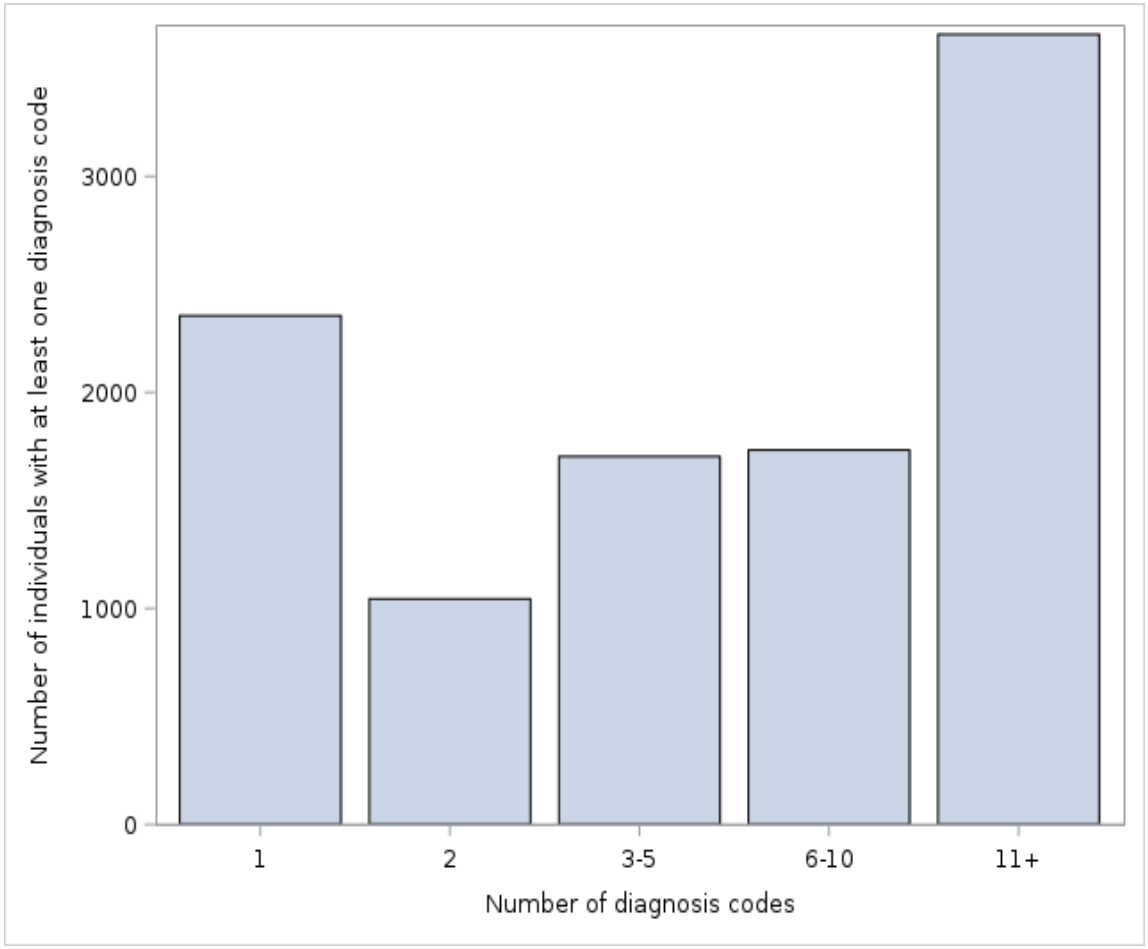


Figure H21. Number of diagnosis codes by individual, atrial fibrillation

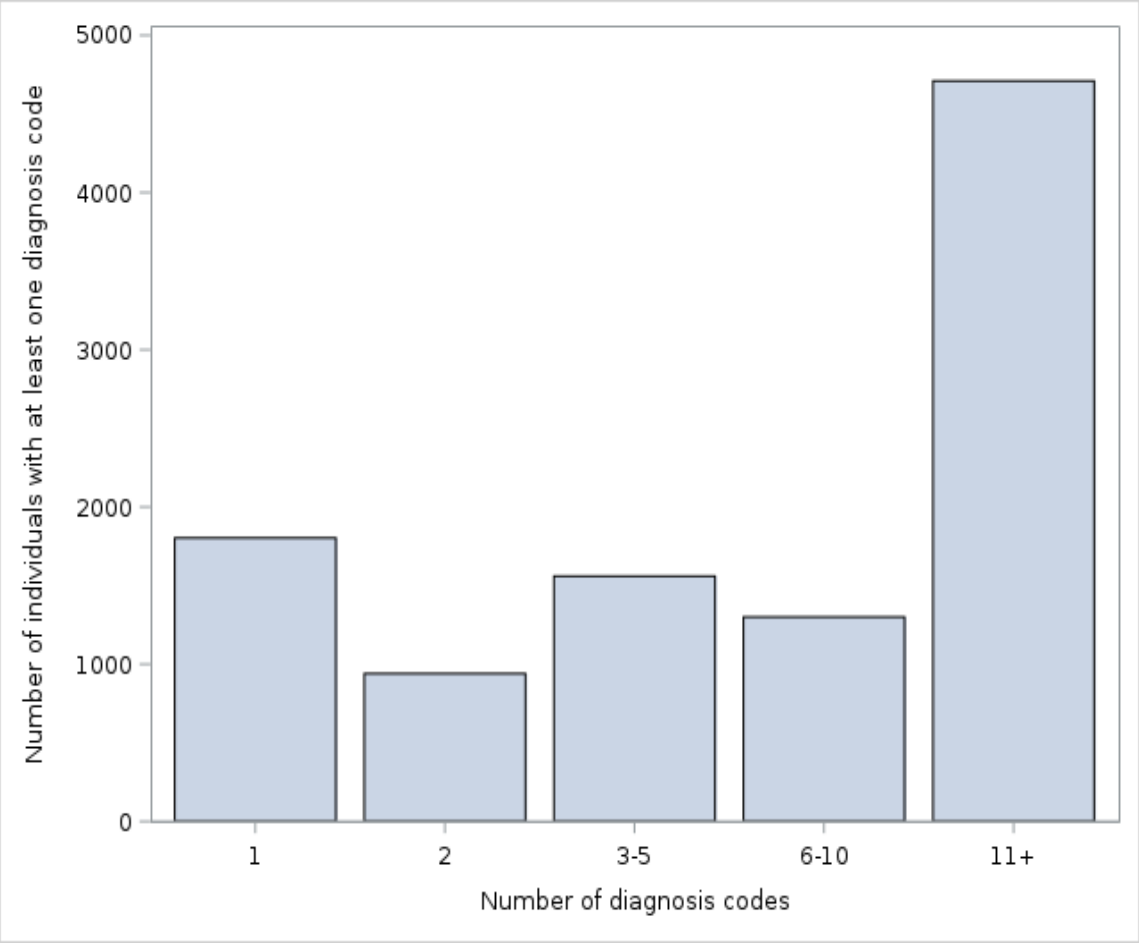


Figure H22. Number of diagnosis codes by individual, heart failure and non-ischemic heart disease

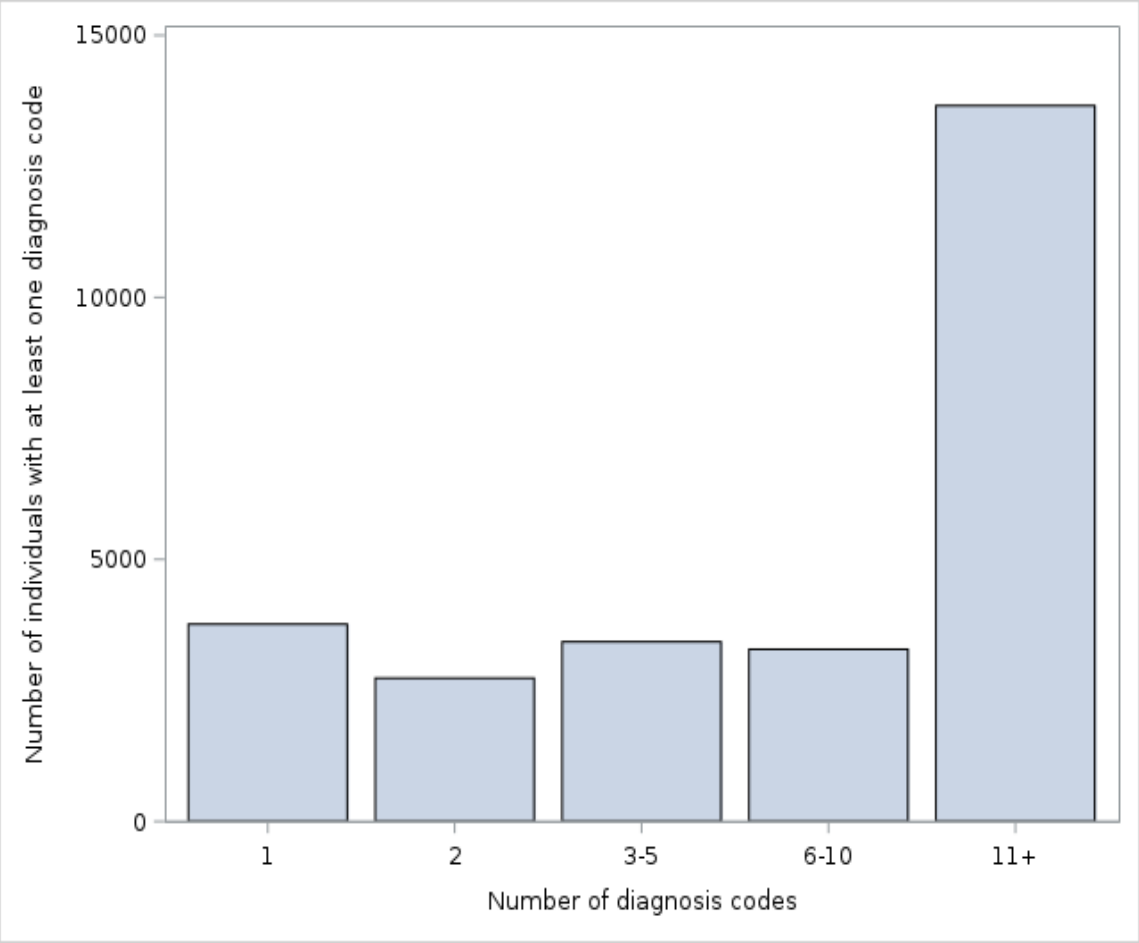


Figure H23. Number of diagnosis codes by individual, ischemic heart disease

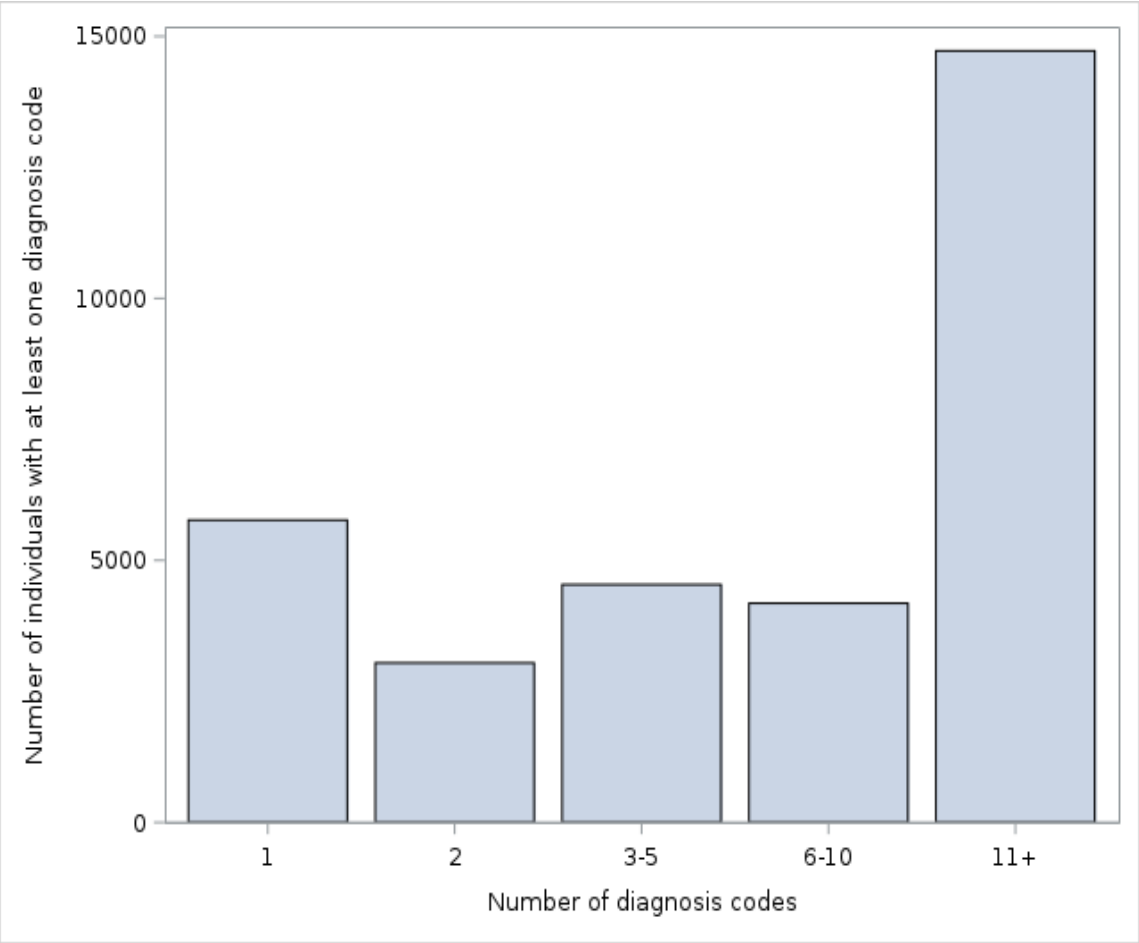


Figure H24. Number of diagnosis codes by individual, peripheral vascular disease

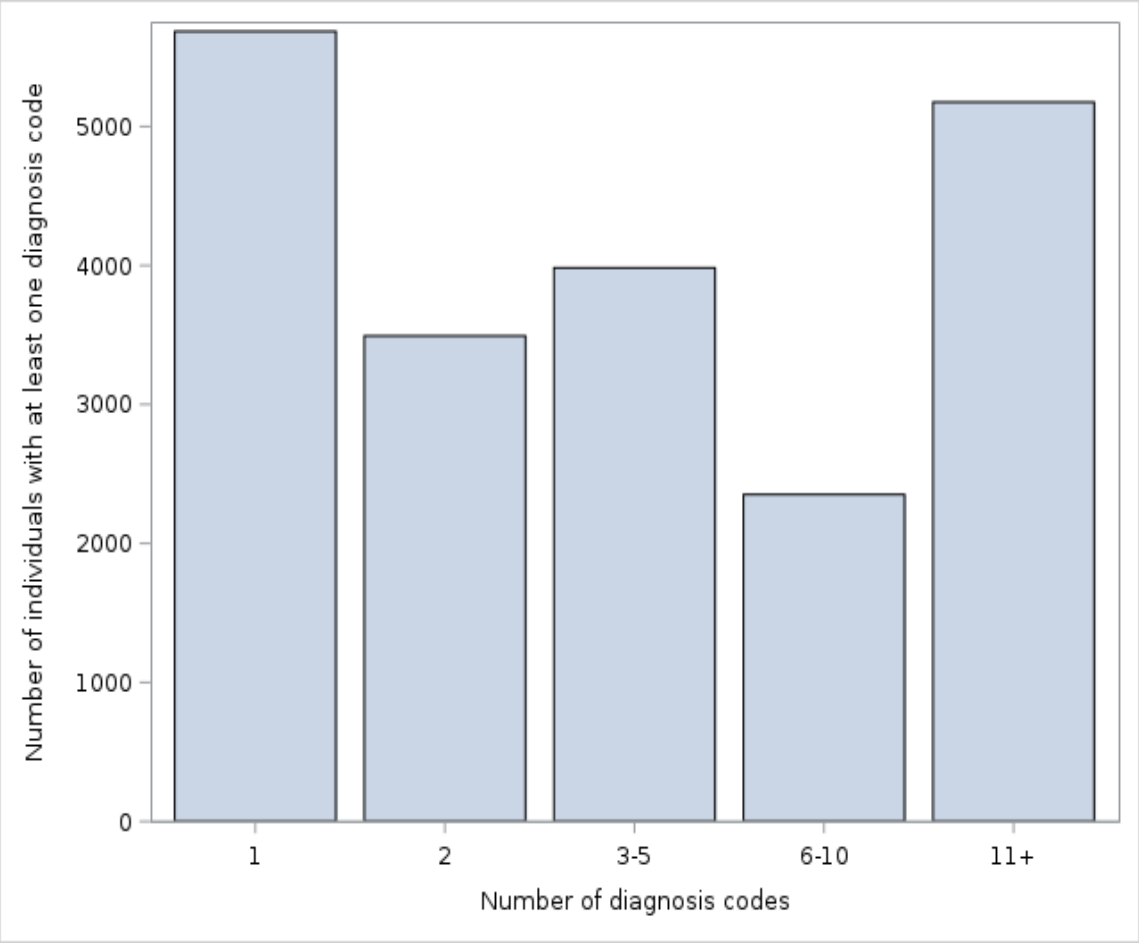


Figure H25. Number of diagnosis codes by individual, hypertension

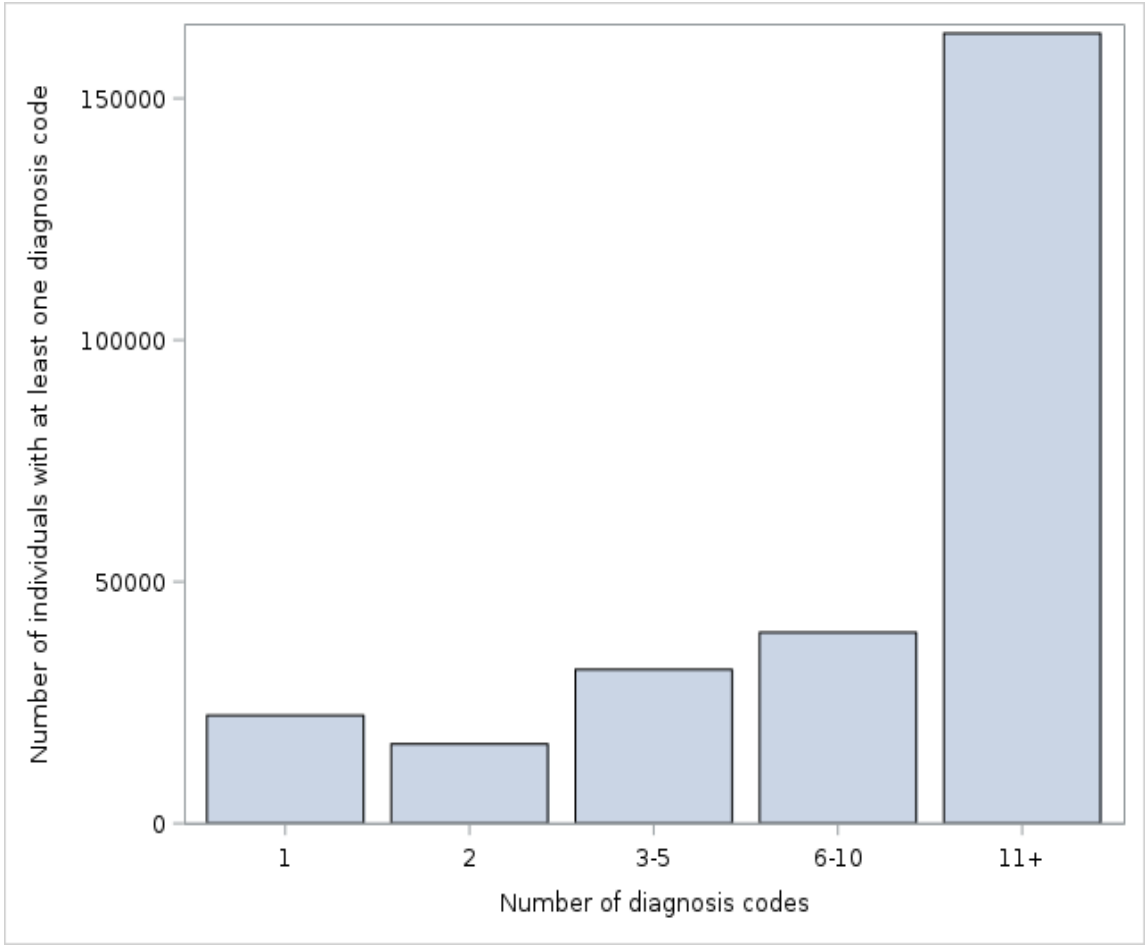


Figure H26. Number of diagnosis codes by individual, stroke

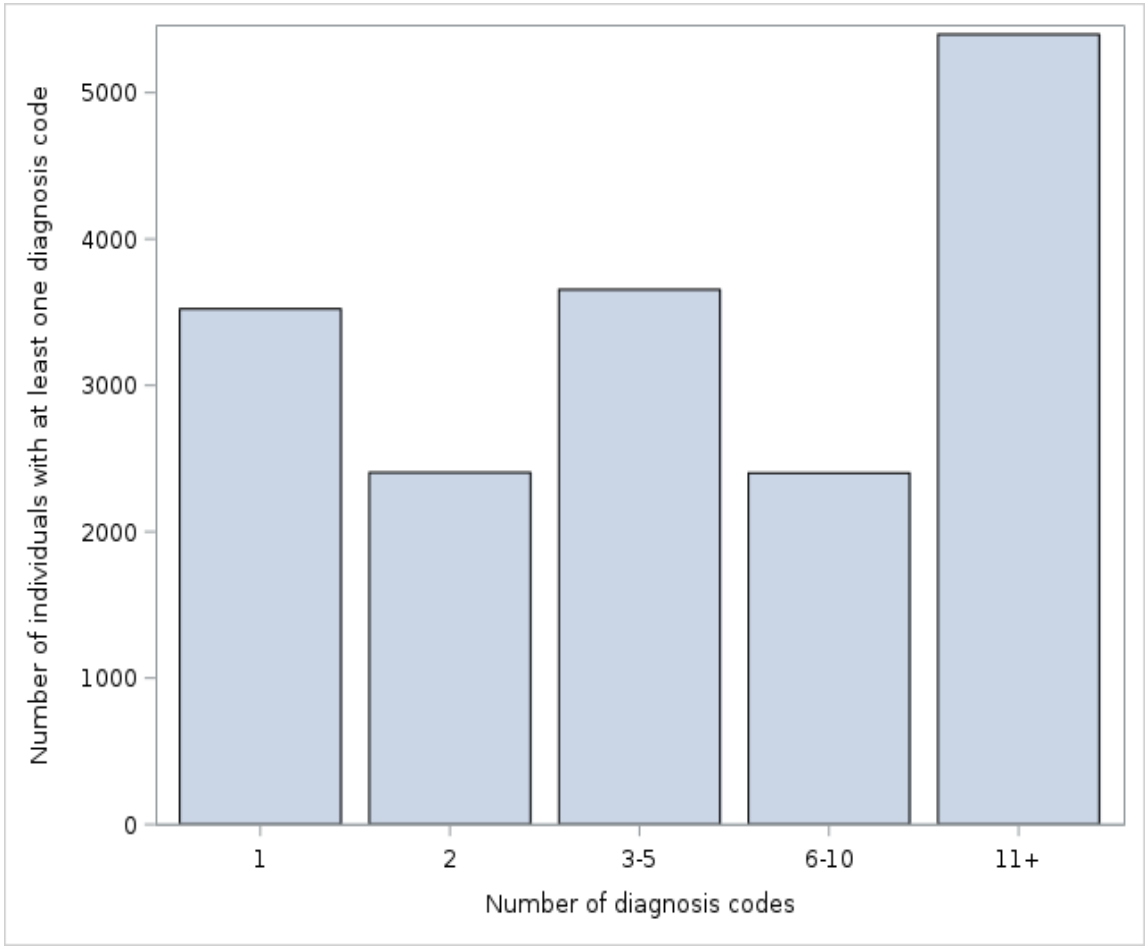


Figure H27. Number of diagnosis codes by individual, hypothyroidism

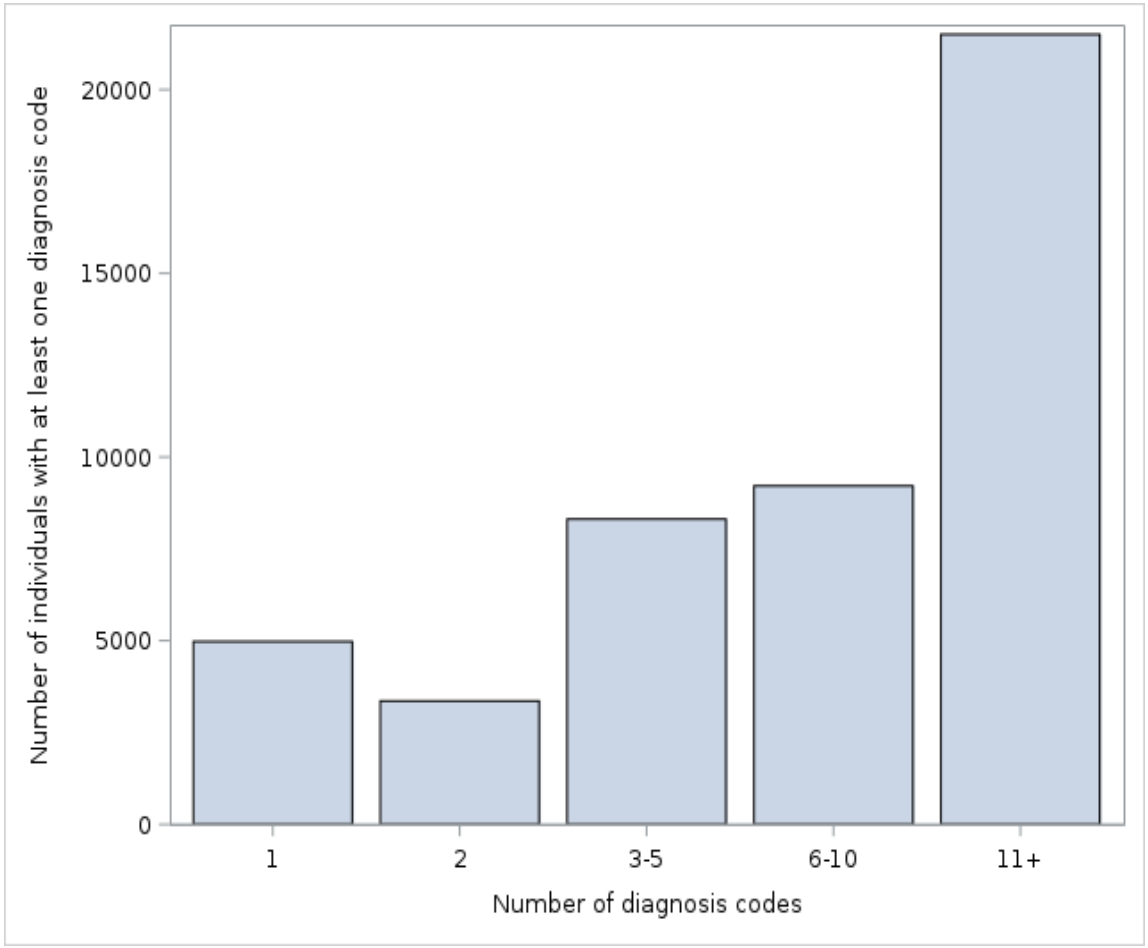


Figure H28. Number of diagnosis codes by individual, type 1 diabetes

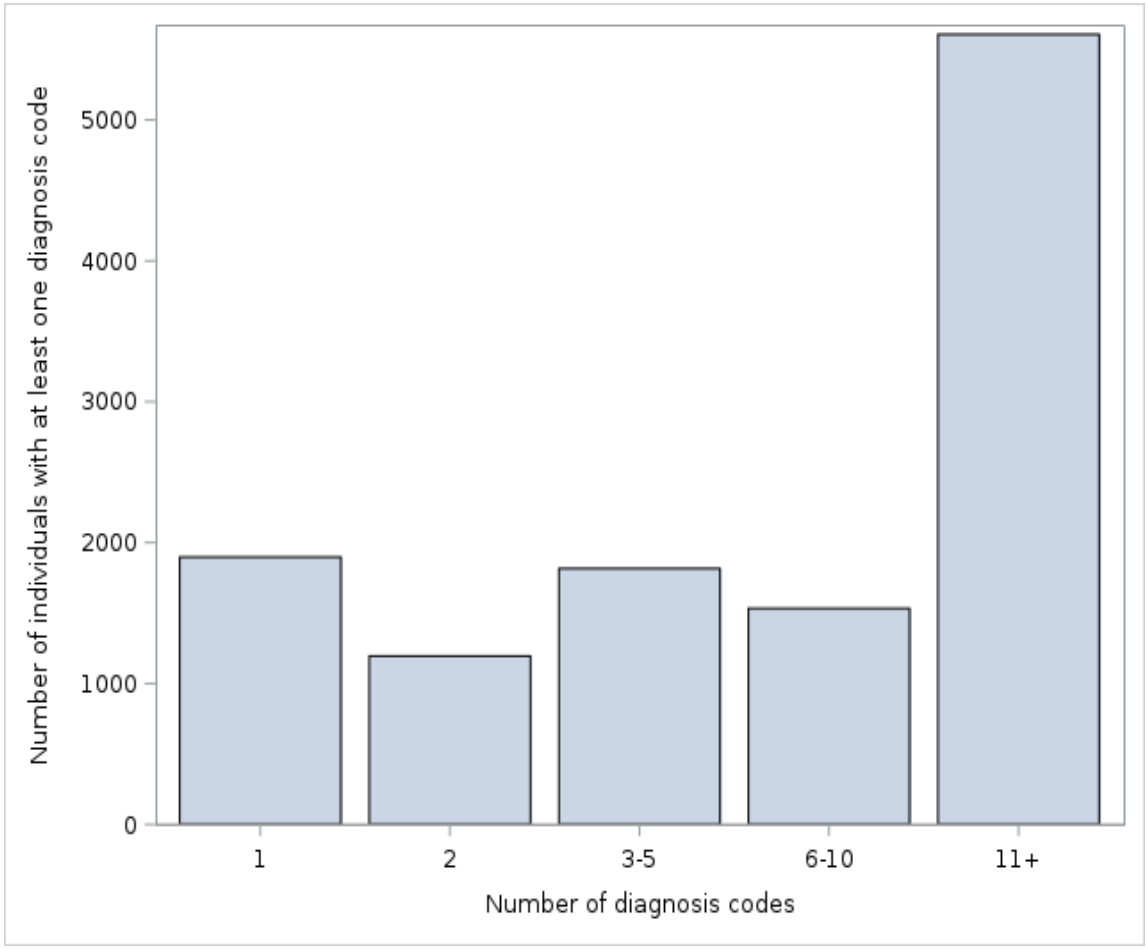


Figure H29. Number of diagnosis codes by individual, type 2 diabetes

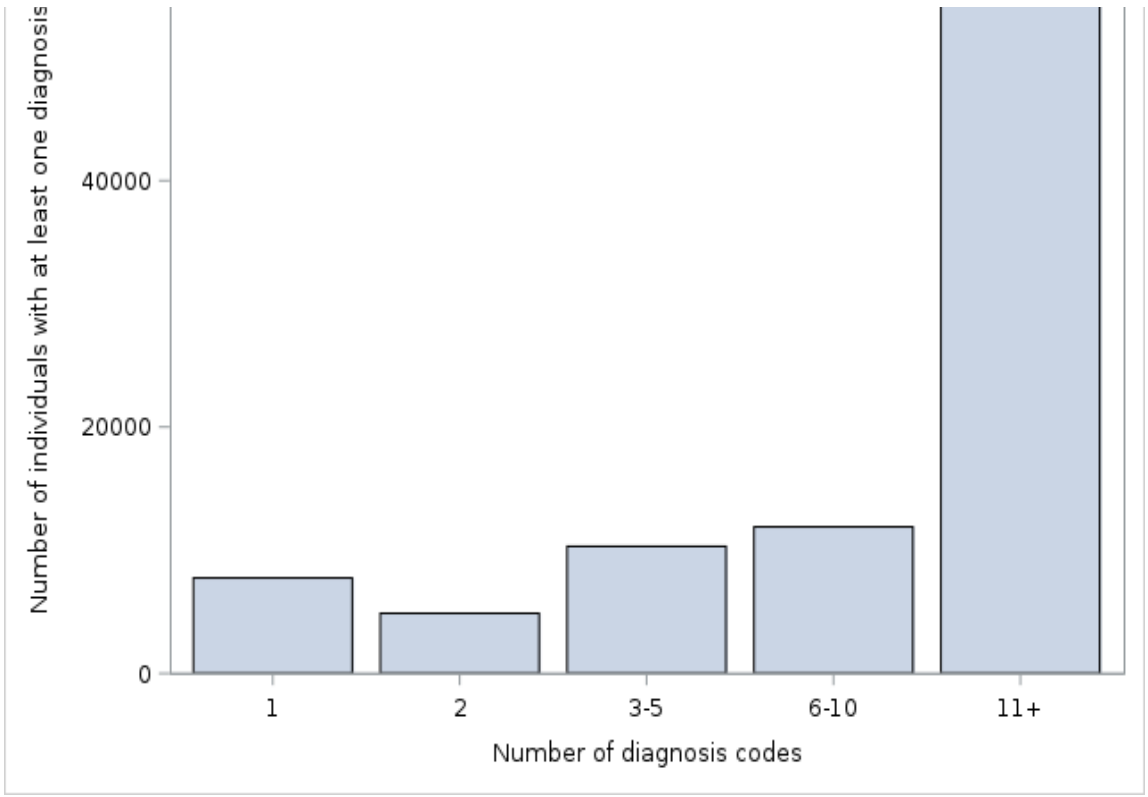

Number of individuals with at least one diagnosis
40000
20000
0
1
2
3-5
6-10
11+
Number of diagnosis codes

Figure H30. Number of diagnosis codes by individual, thyroid gland disorders

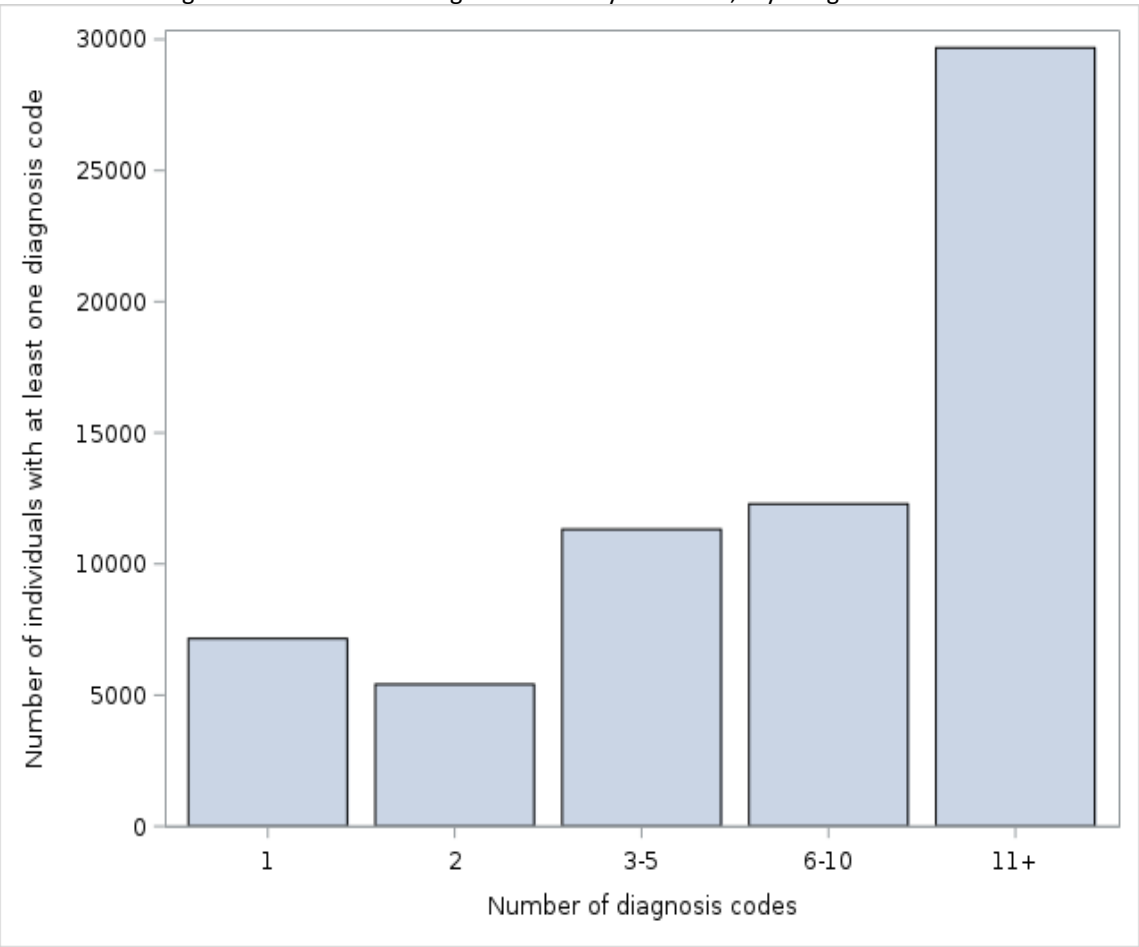


Figure H31. Number of diagnosis codes by individual, other glucose disorders

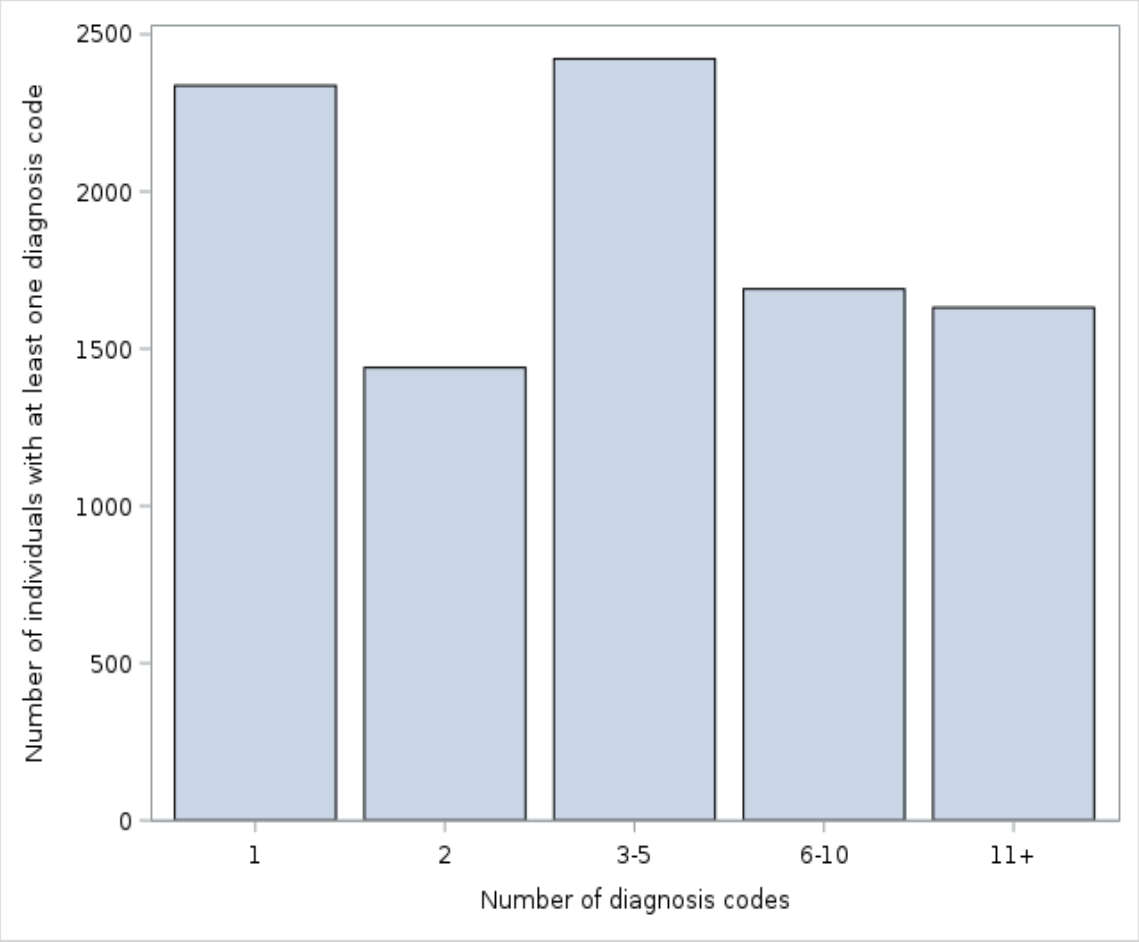


Figure H32. Number of diagnosis codes by individual, irritable bowel syndrome

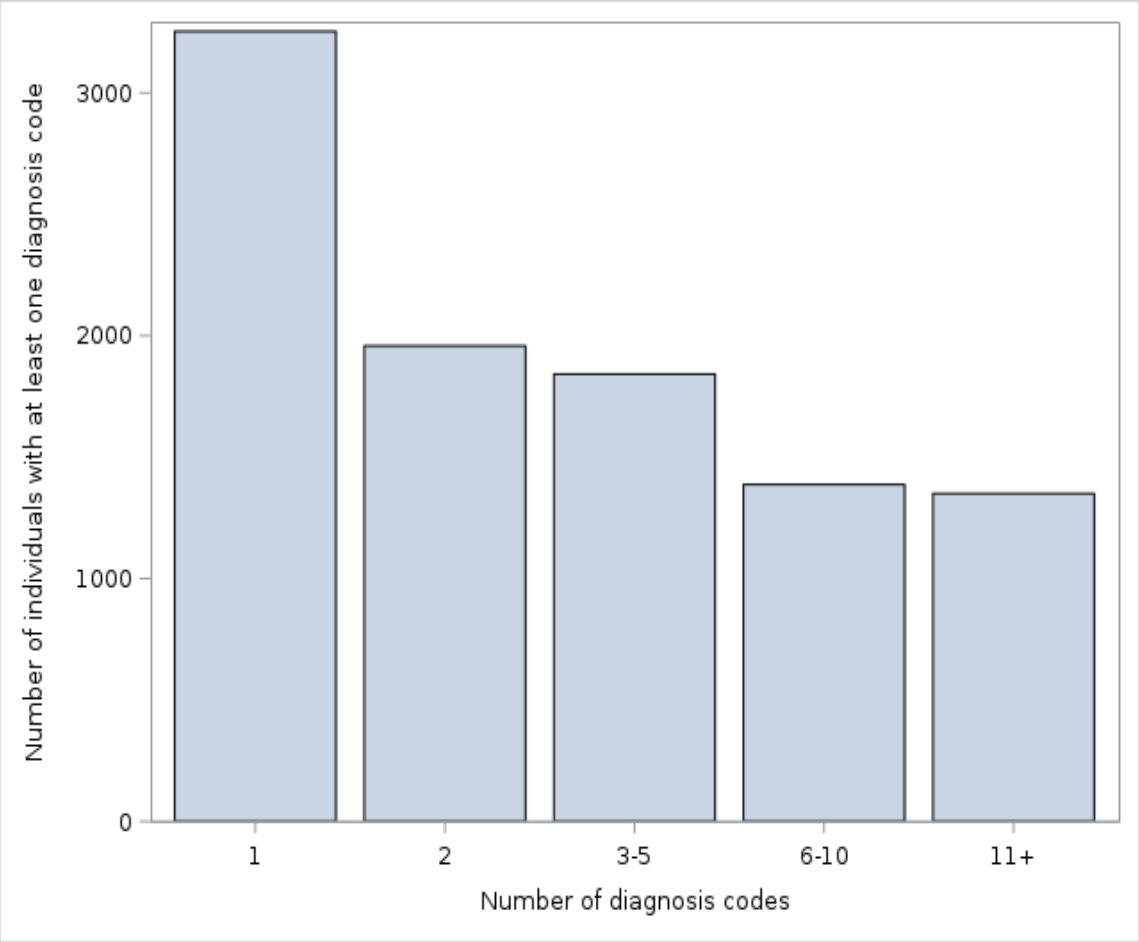


Figure H33. Number of diagnosis codes by individual, gastrointestinal reflux disease

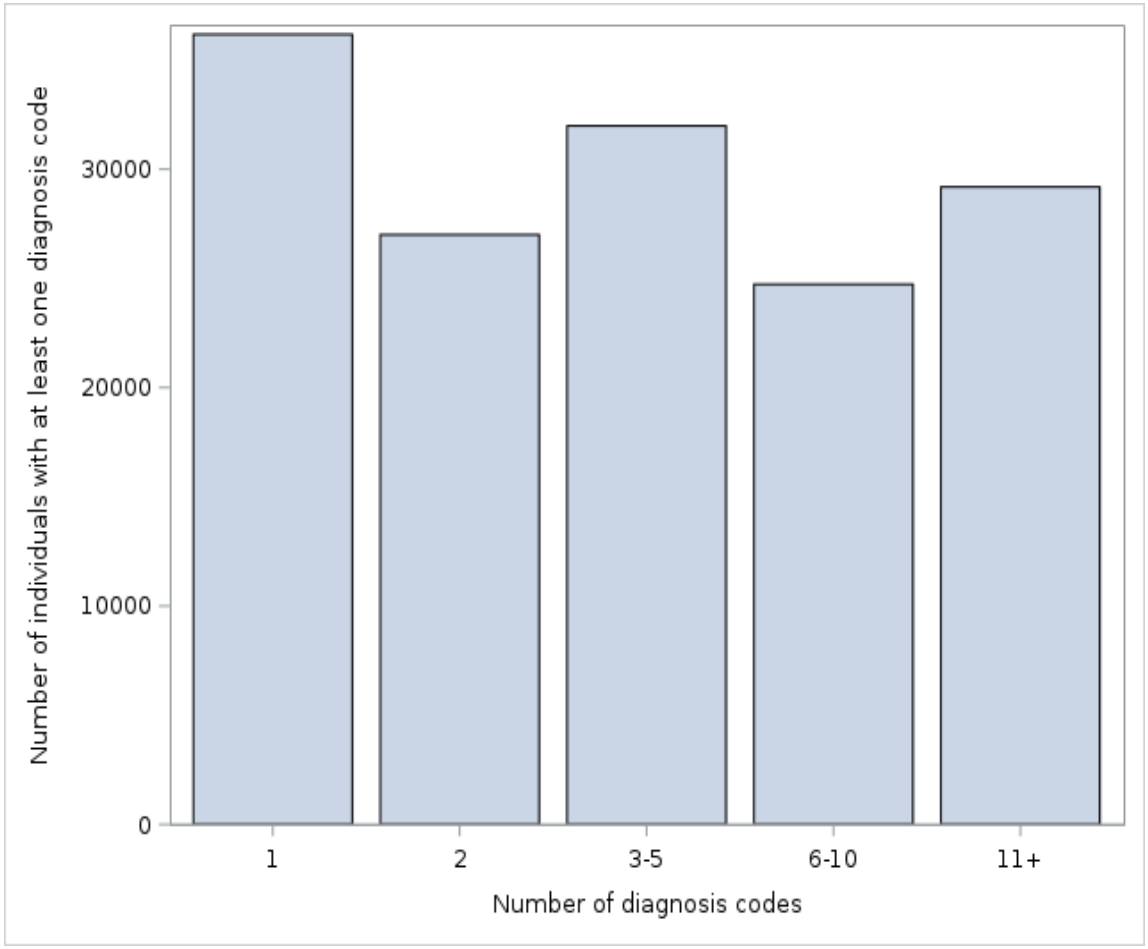


Figure H34. Number of diagnosis codes by individual, inflammatory bowel diseases

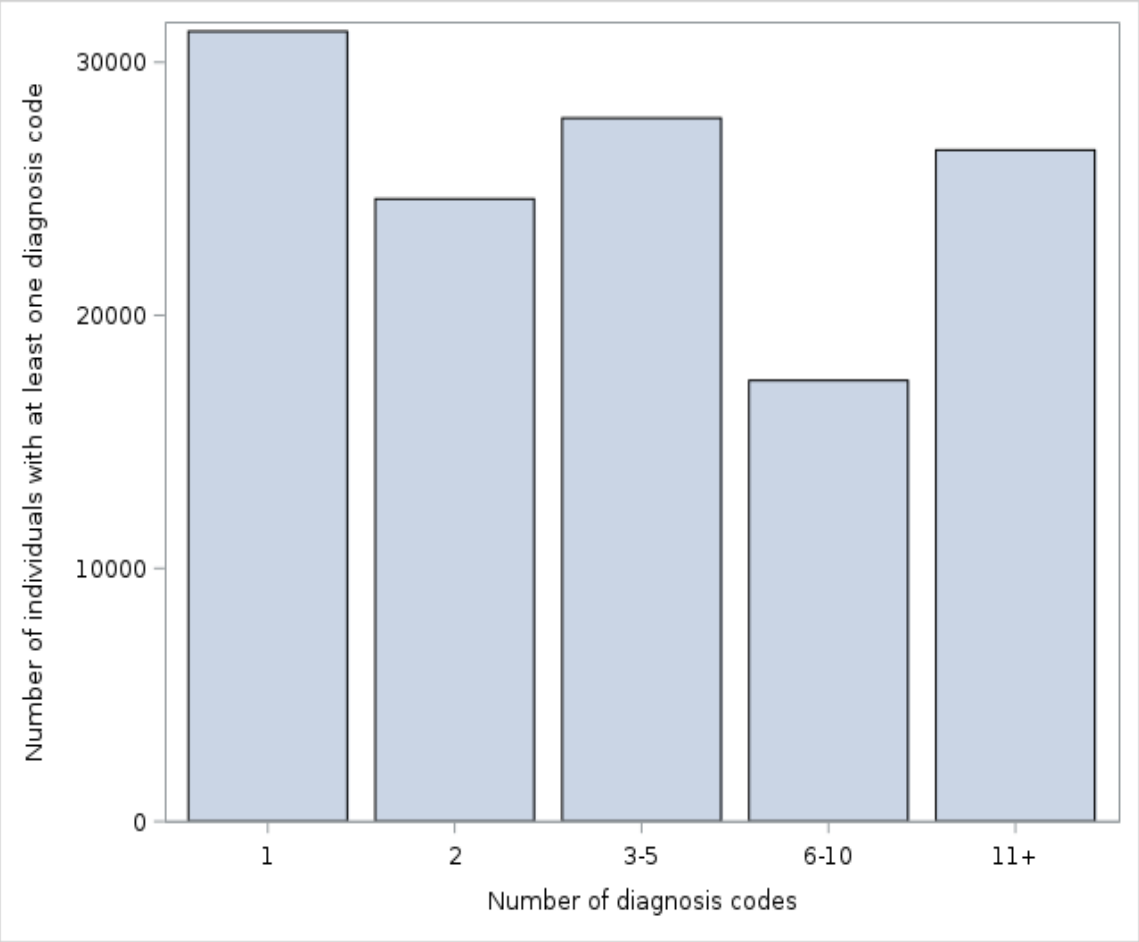


Figure H35. Number of diagnosis codes by individual, anxiety disorders

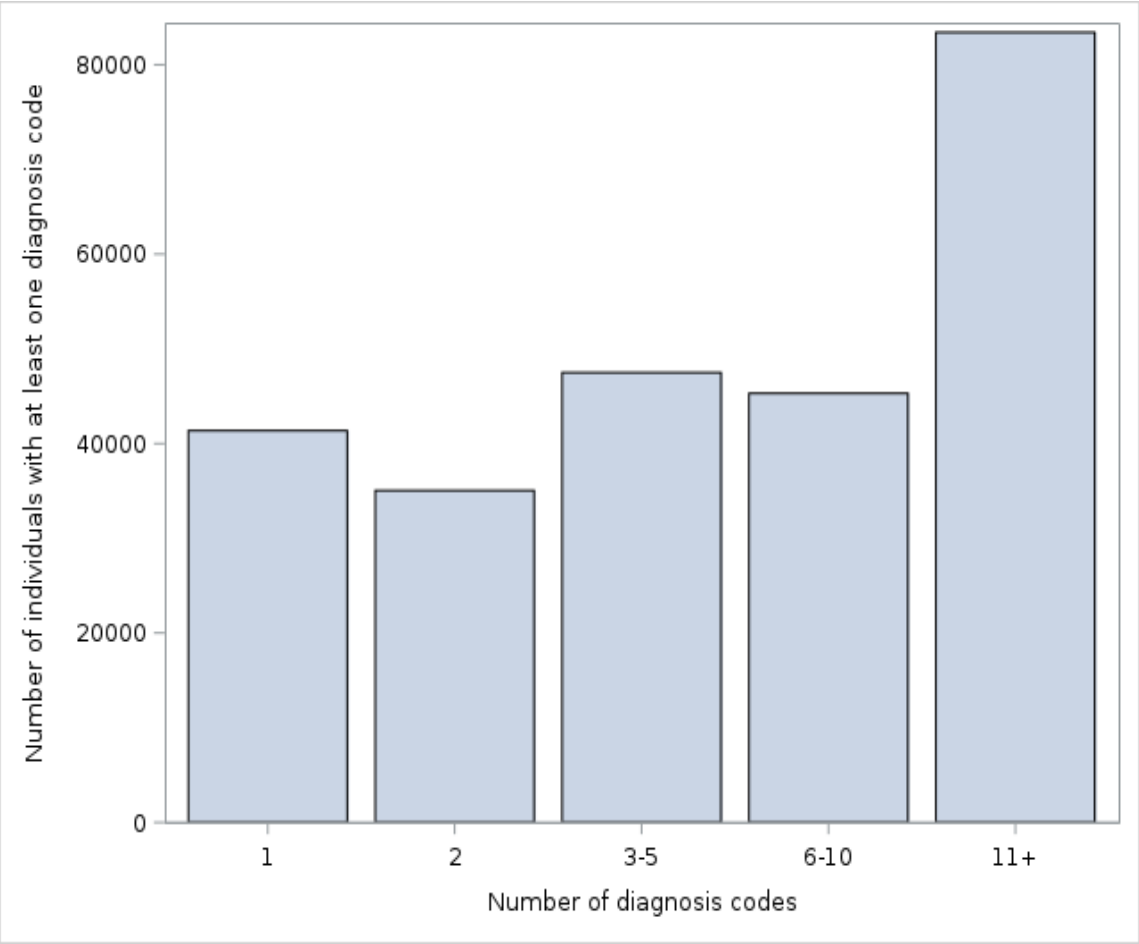


Figure H36. Number of diagnosis codes by individual, bipolar disorder

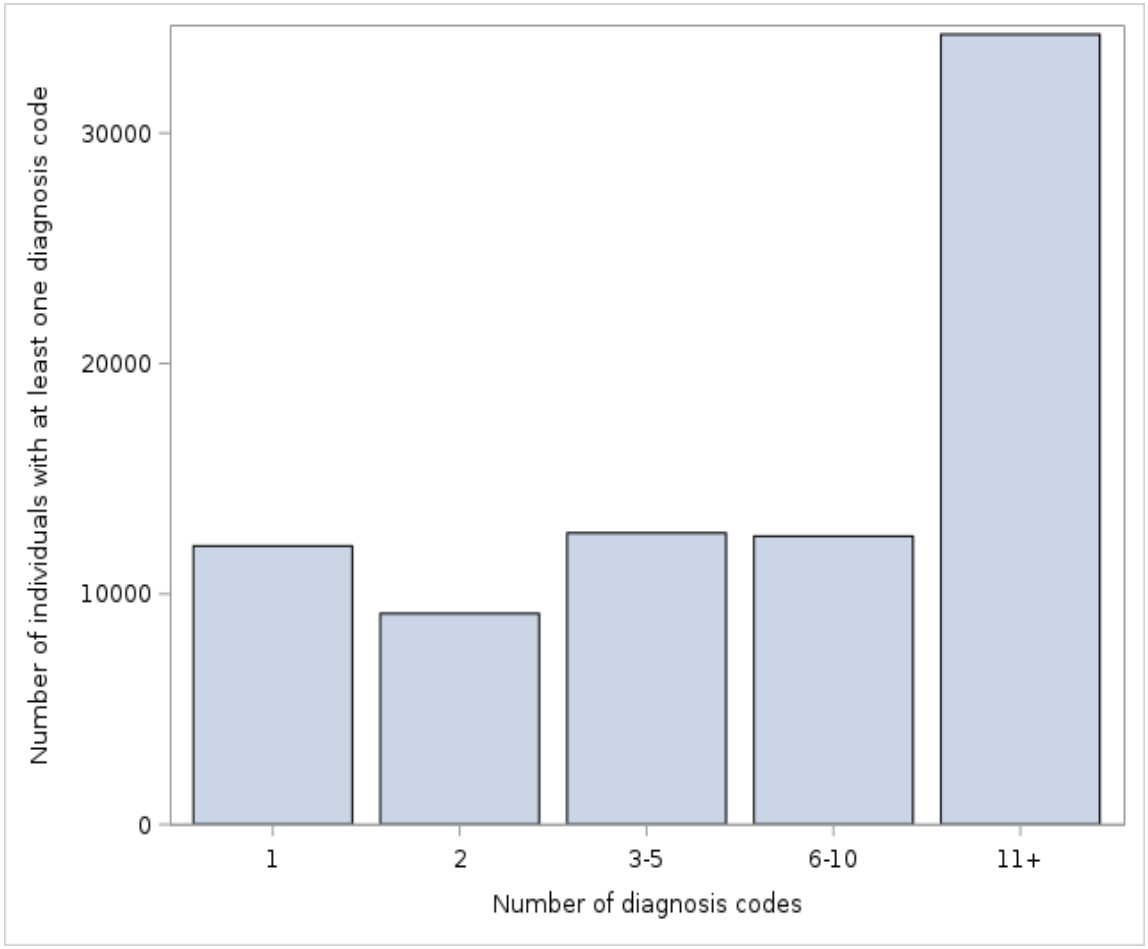


Figure H37. Number of diagnosis codes by individual, depressive disorders

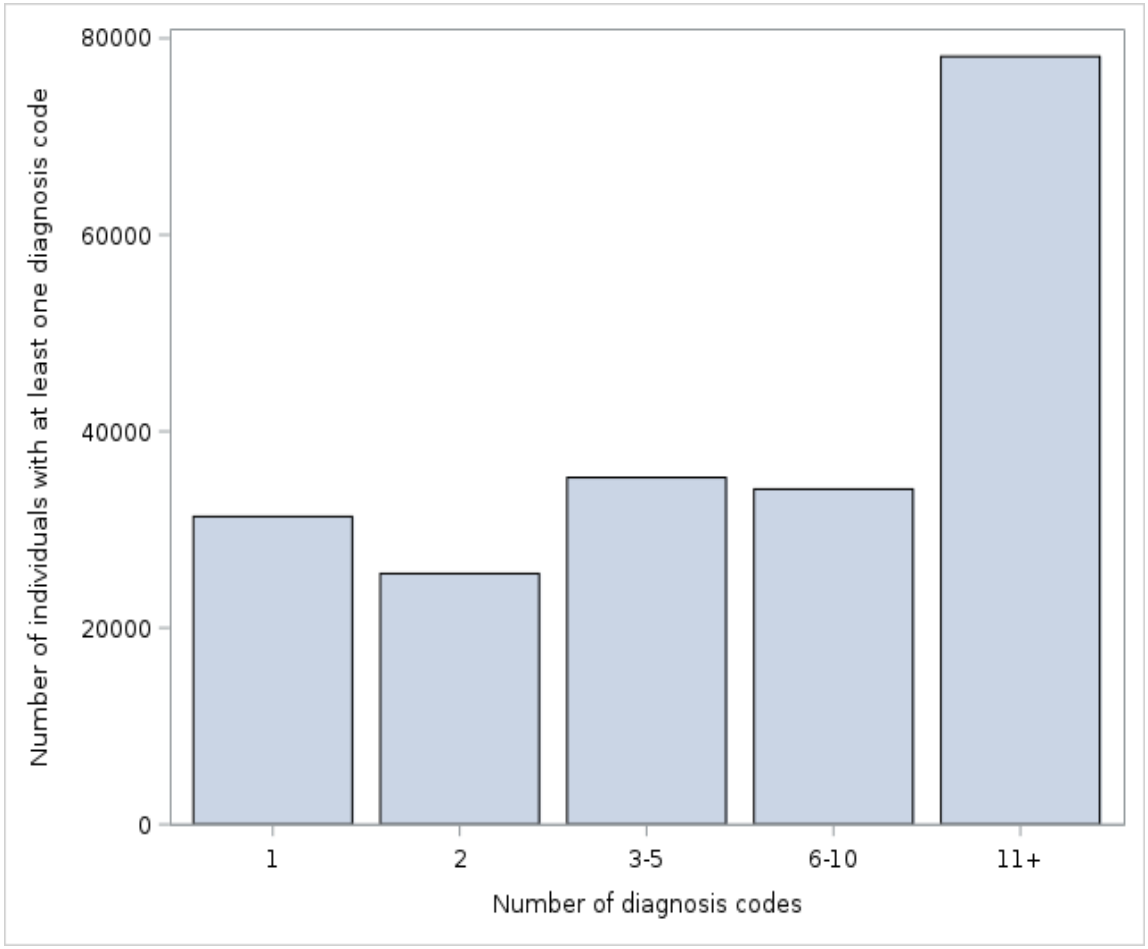


Figure H38. Number of diagnosis codes by individual, arthritis

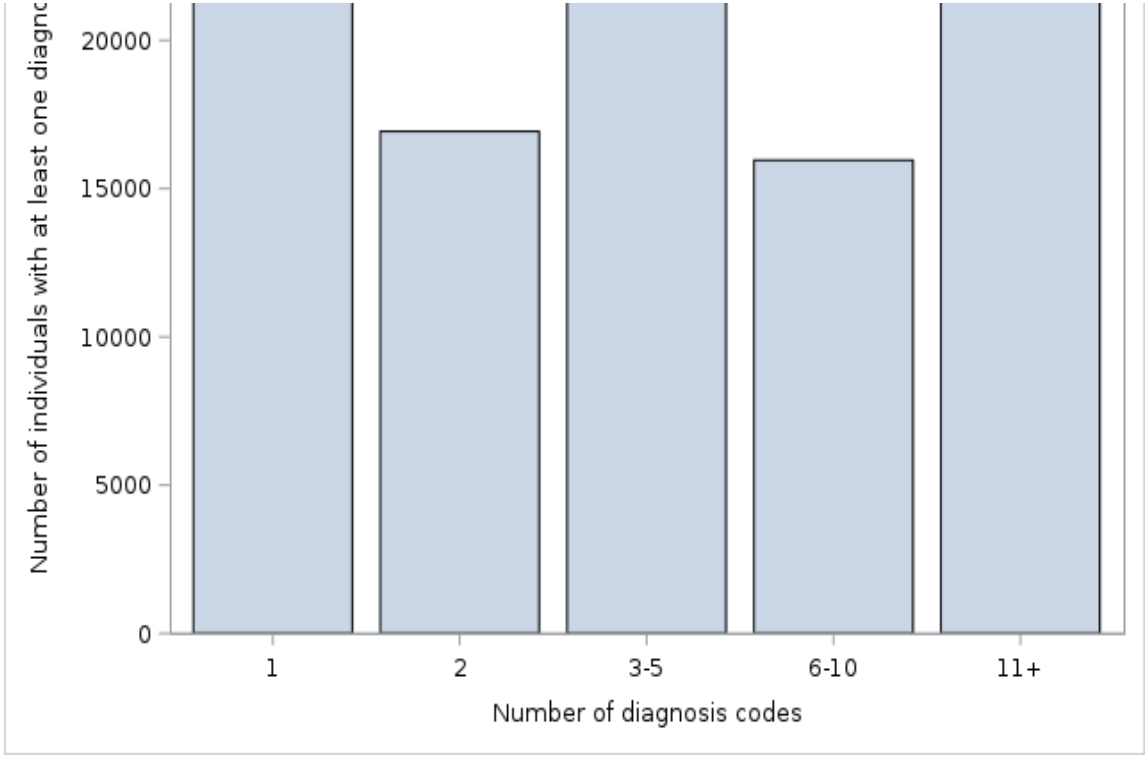

Number of individuals with at least one diagnc
20000
15000
10000
5000
0
1
2
3-5
6-10
11+
Number of diagnosis codes

Figure H39. Number of diagnosis codes by individual, osteoporosis

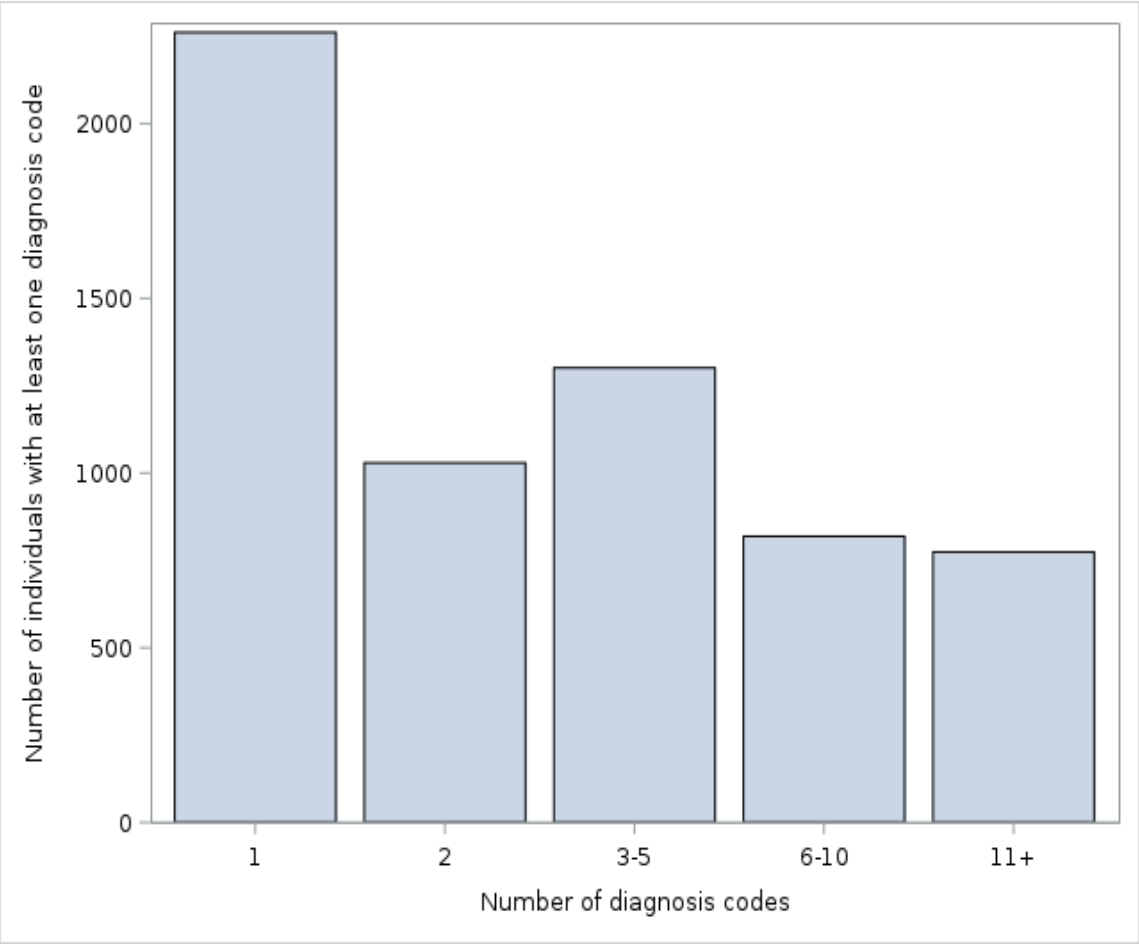


Figure H40. Number of diagnosis codes by individual, Alzheimer's disease

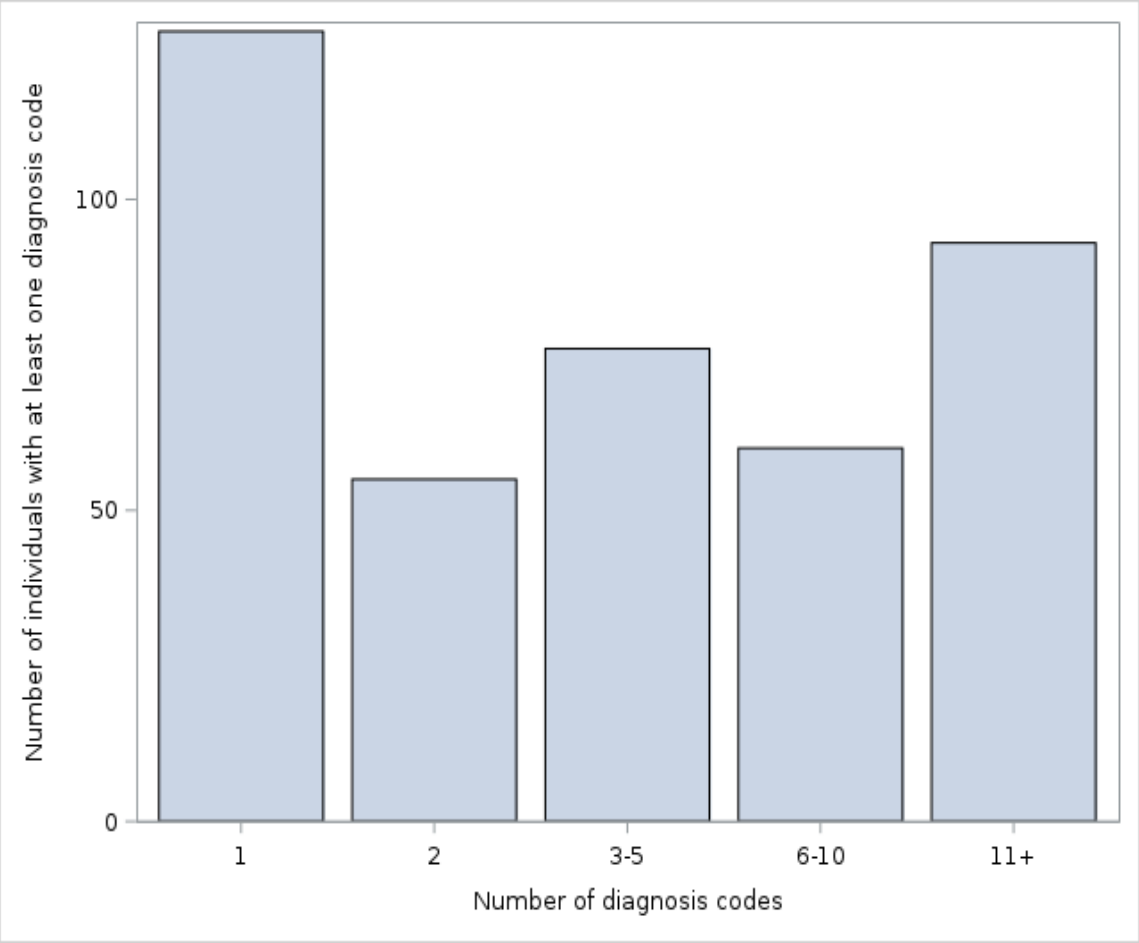


Figure H41. Number of diagnosis codes by individual, dementia

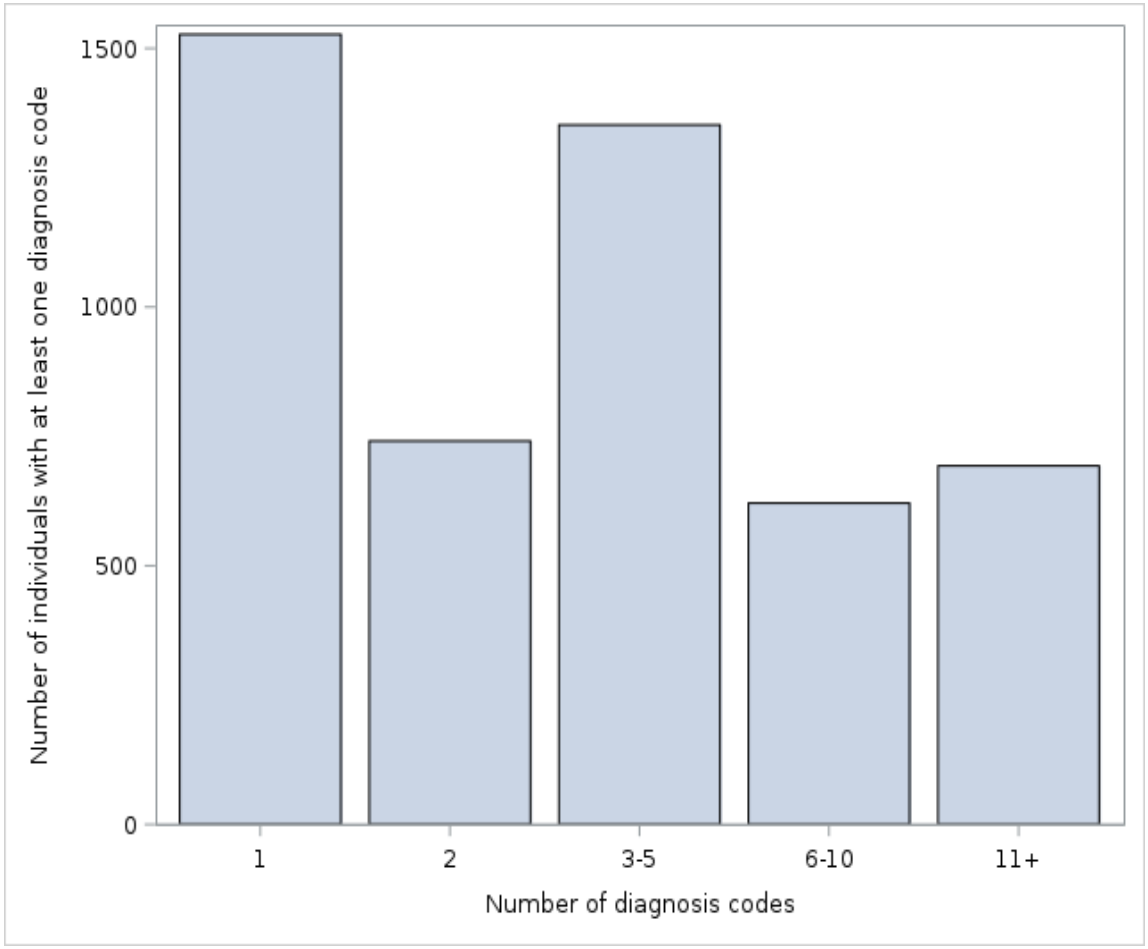


Figure H42. Number of diagnosis codes by individual, Parkinson's disease and secondary Parkinsonism

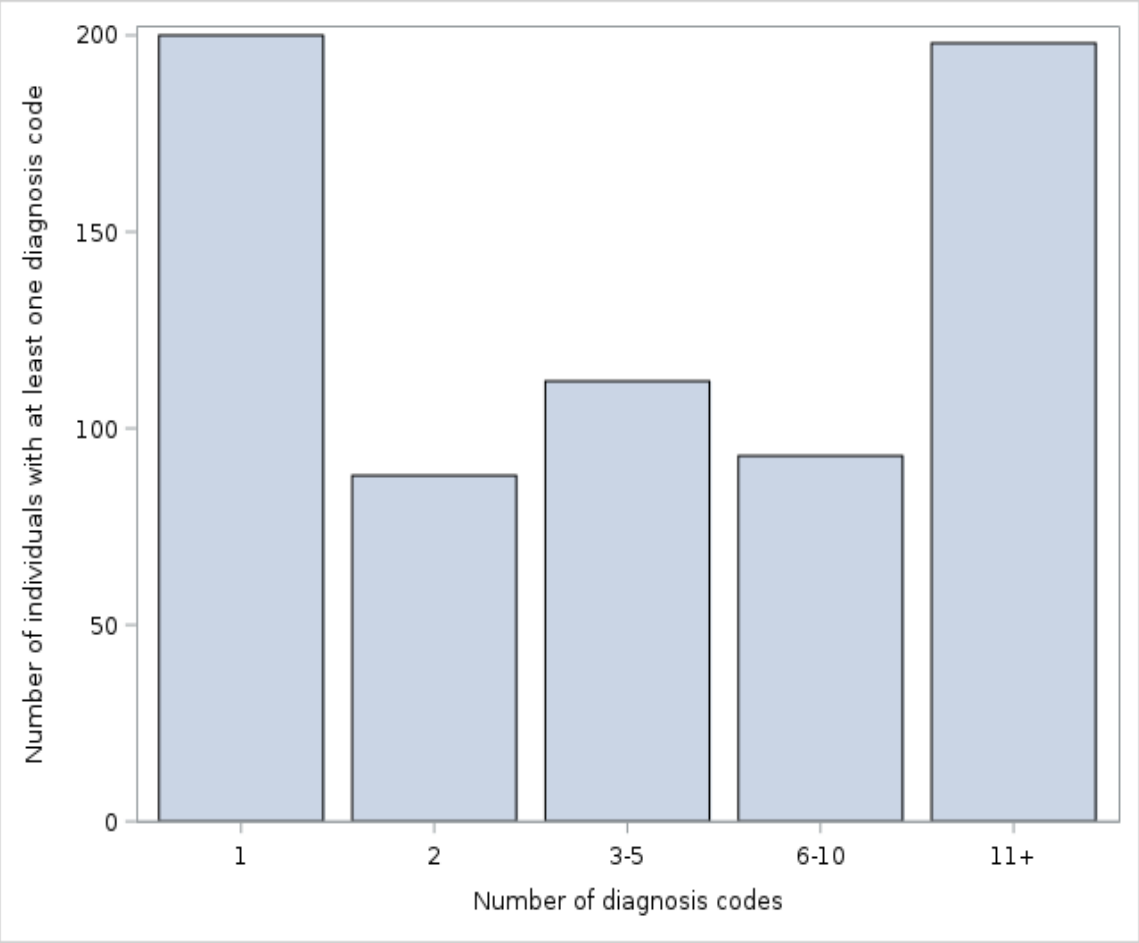


Figure H43. Number of diagnosis codes by individual, ADHD

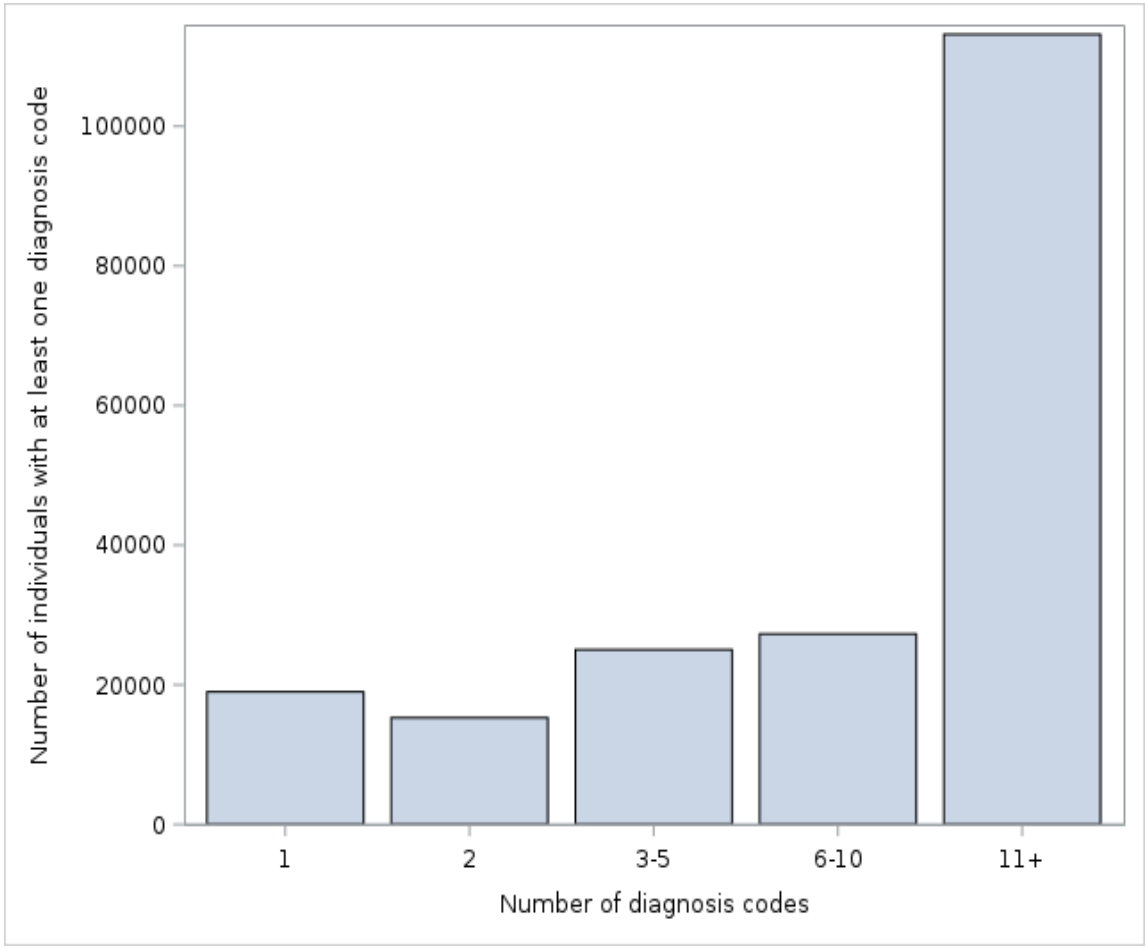


Figure H44. Number of diagnosis codes by individual, autism spectrum disorders

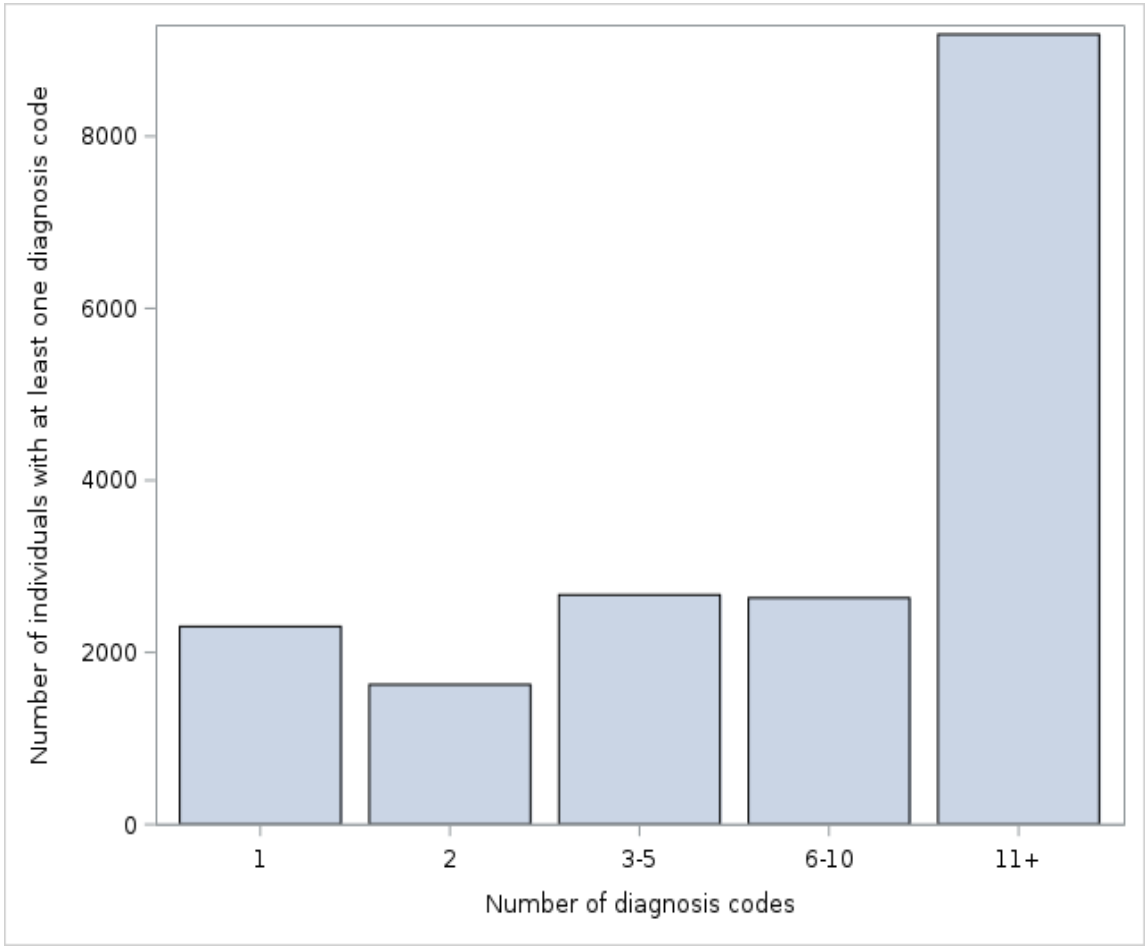


Figure H45. Number of diagnosis codes by individual, intellectual disabilities

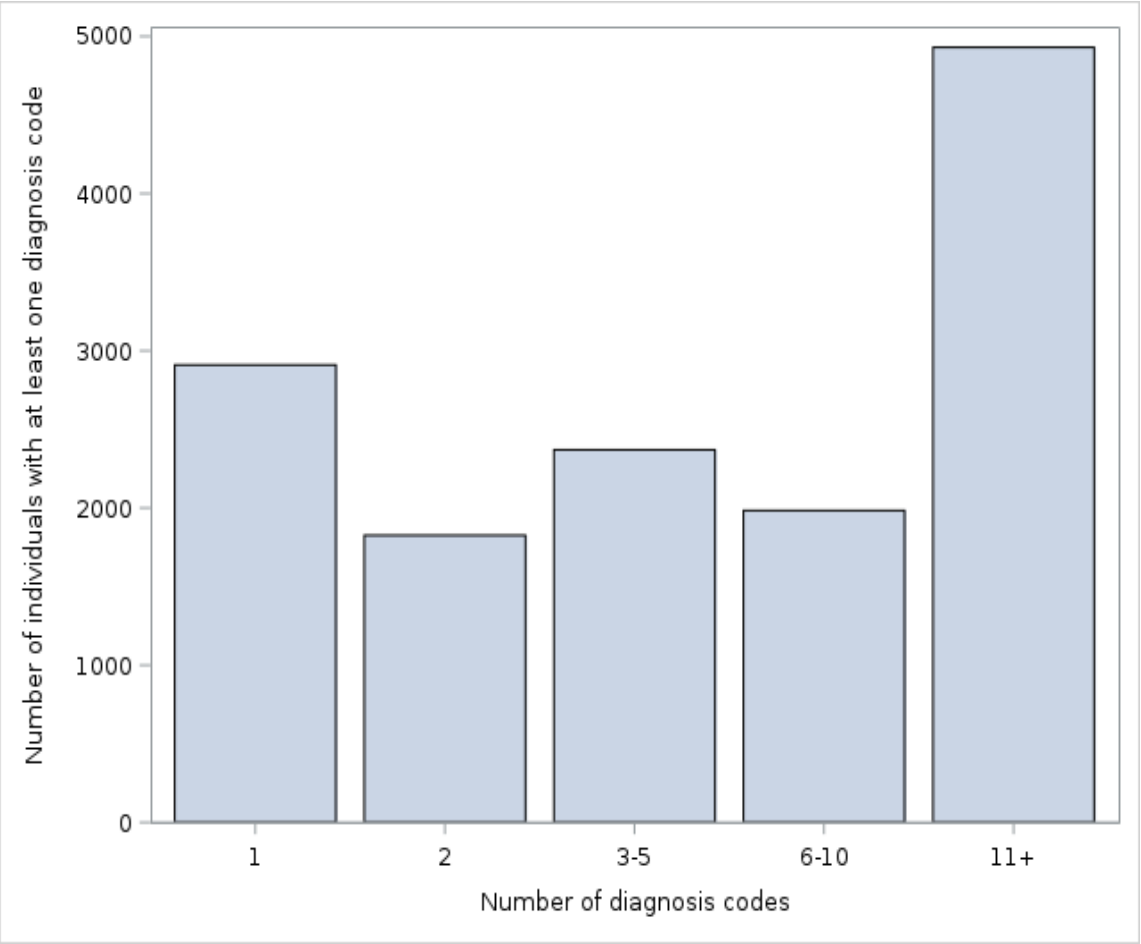


Figure H46. Number of diagnosis codes by individual, learning disabilities

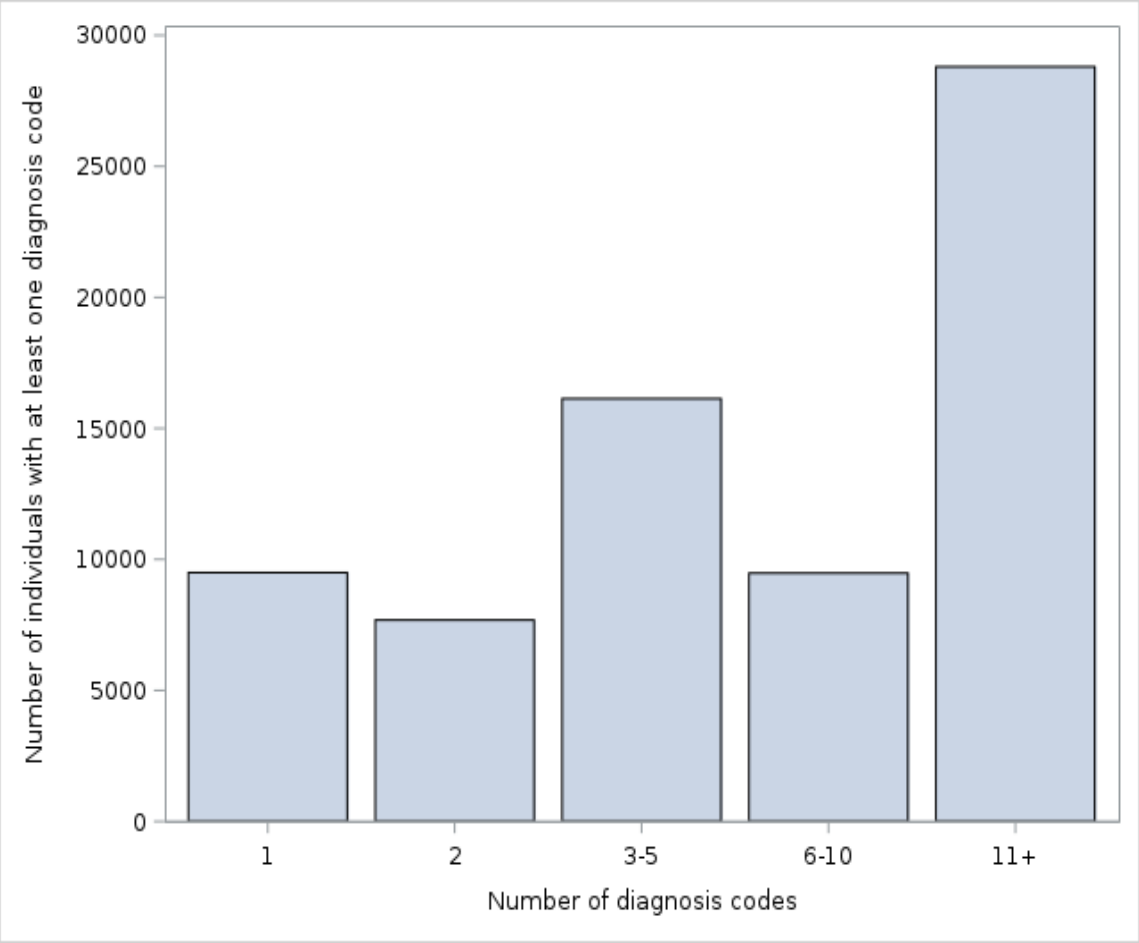


Figure H47. Number of diagnosis codes by individual, other developmental delays

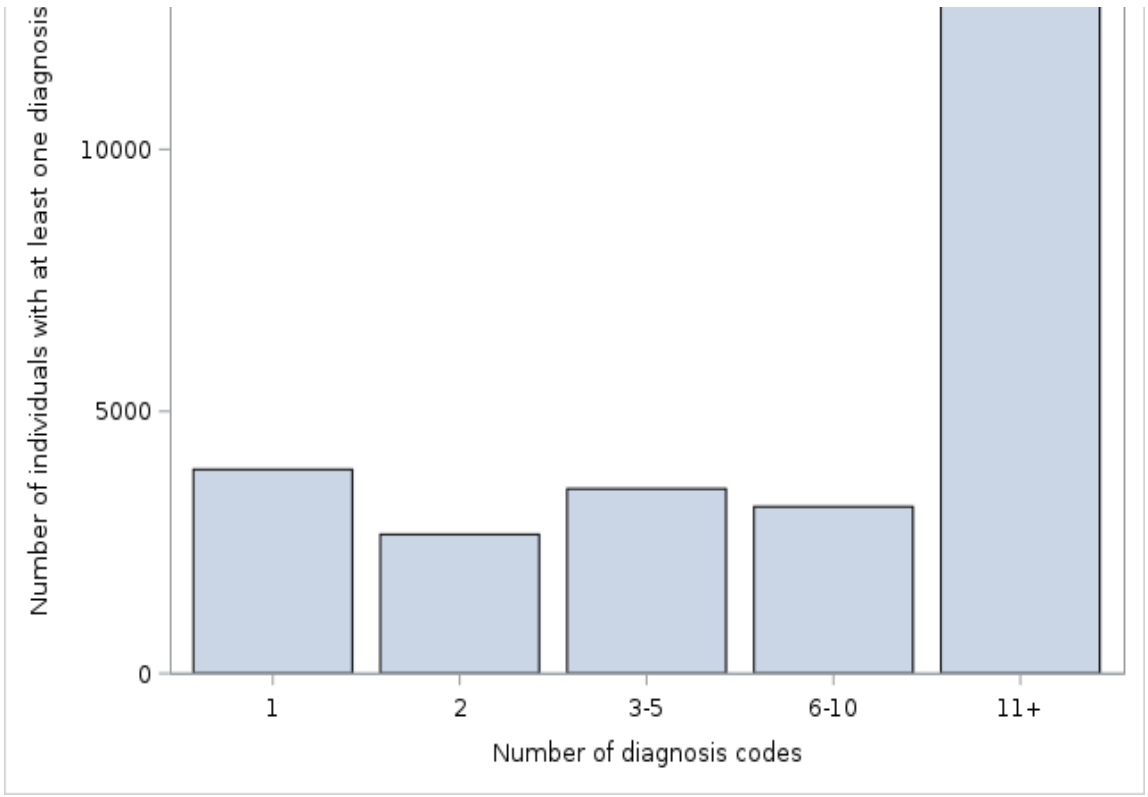

Number of individuals with at least one diagnosis
10000
5000
0
1
2
3-5
6-10
11+
Number of diagnosis codes

Figure H48. Number of diagnosis codes by individual, chronic kidney disease

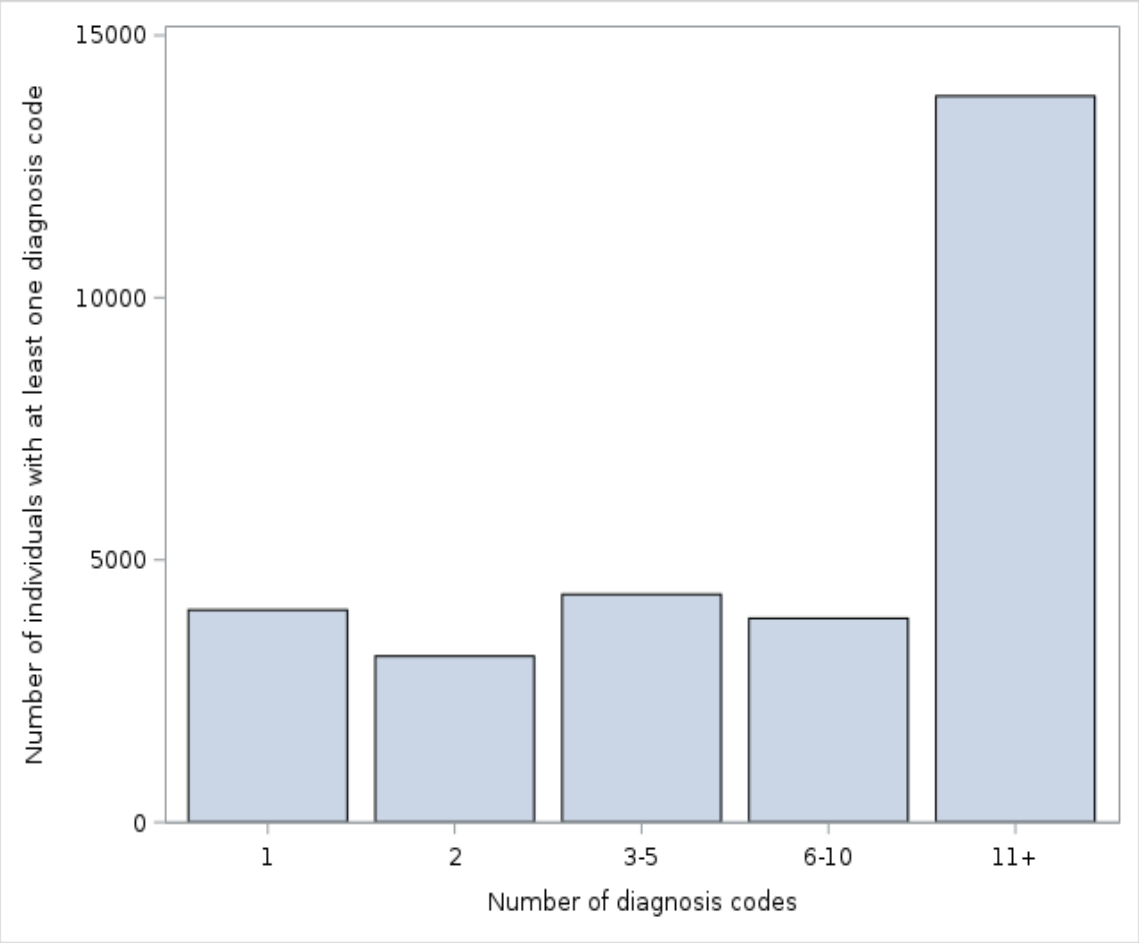


Figure H49. Number of diagnosis codes by individual, liver disease

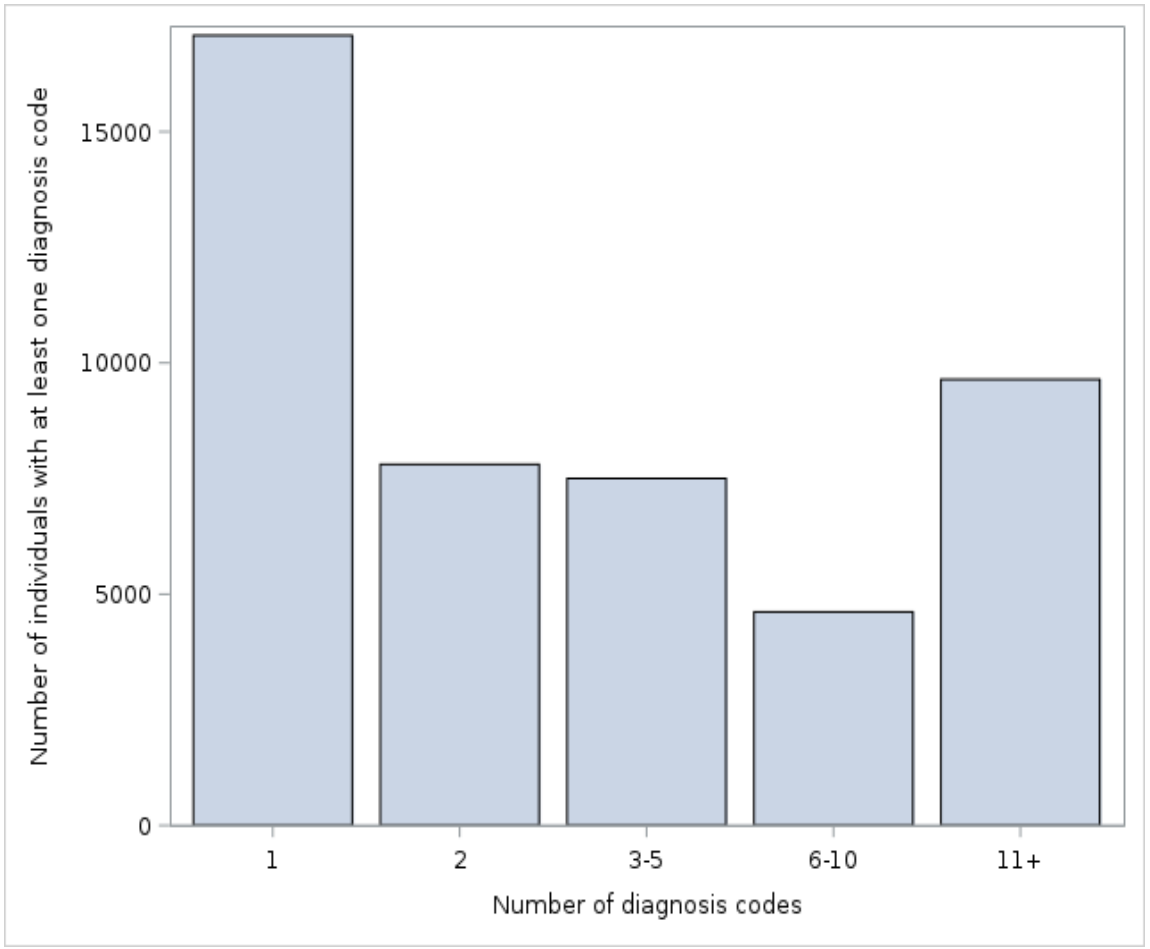


Figure H50. Number of diagnosis codes by individual, migraine

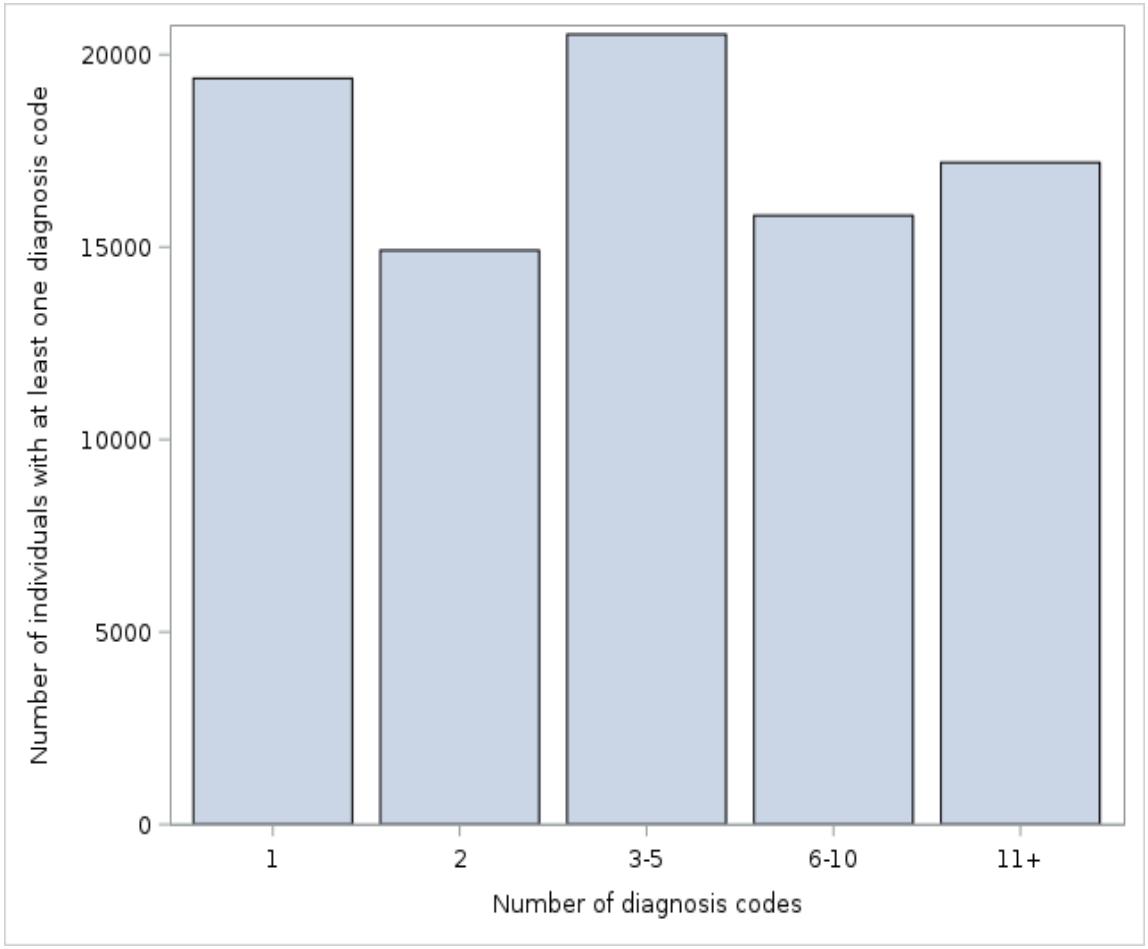


Figure H51. Number of diagnosis codes by individual, epilepsy

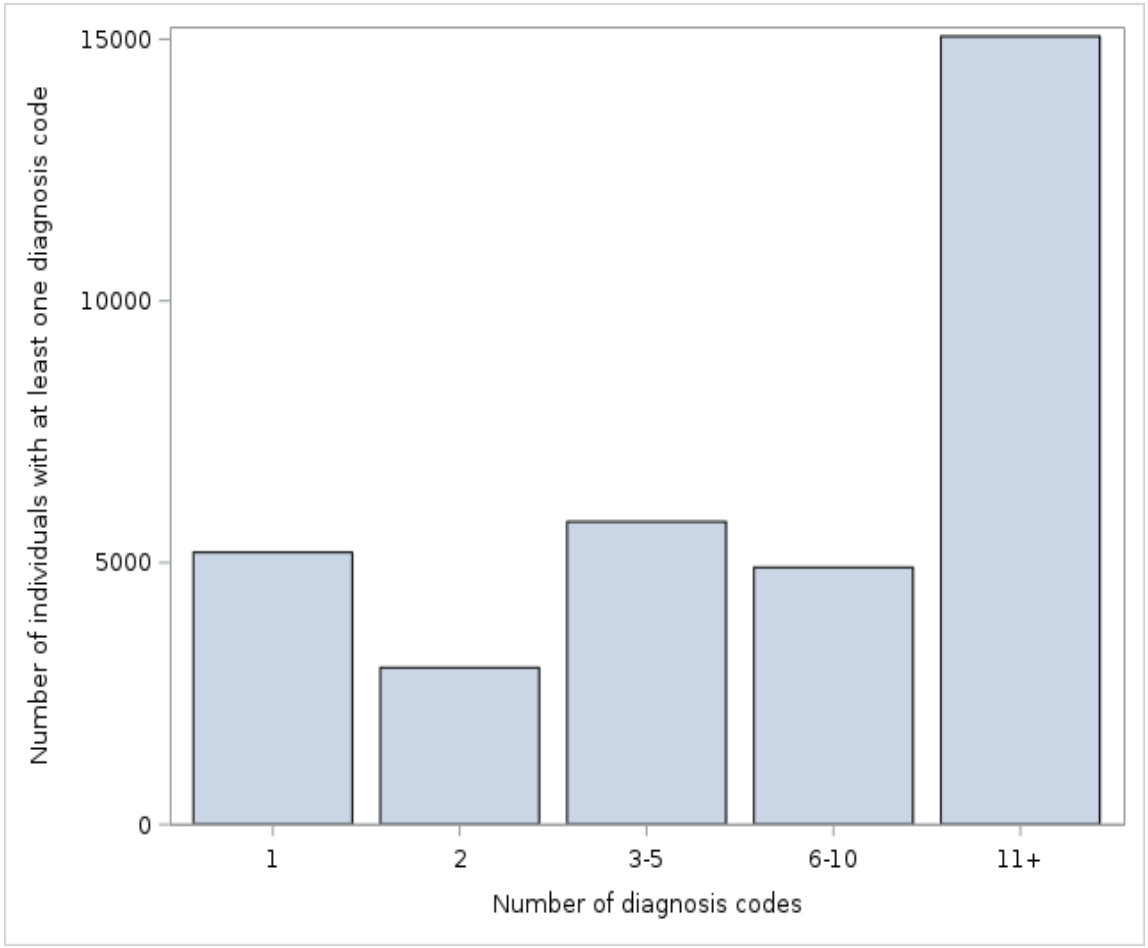


Figure H52. Number of diagnosis codes by individual, obesity

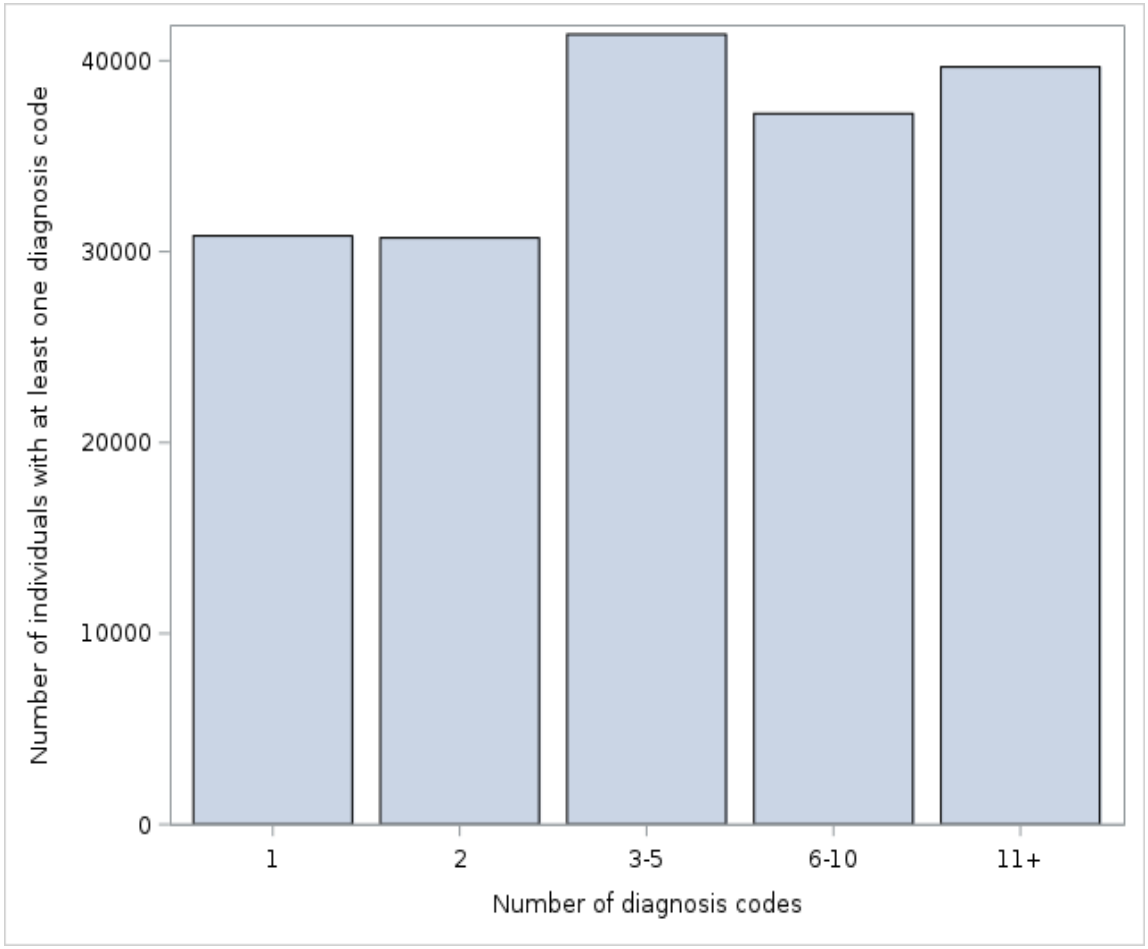


Figure H53. Number of diagnosis codes by individual, preeclampsia and eclampsia

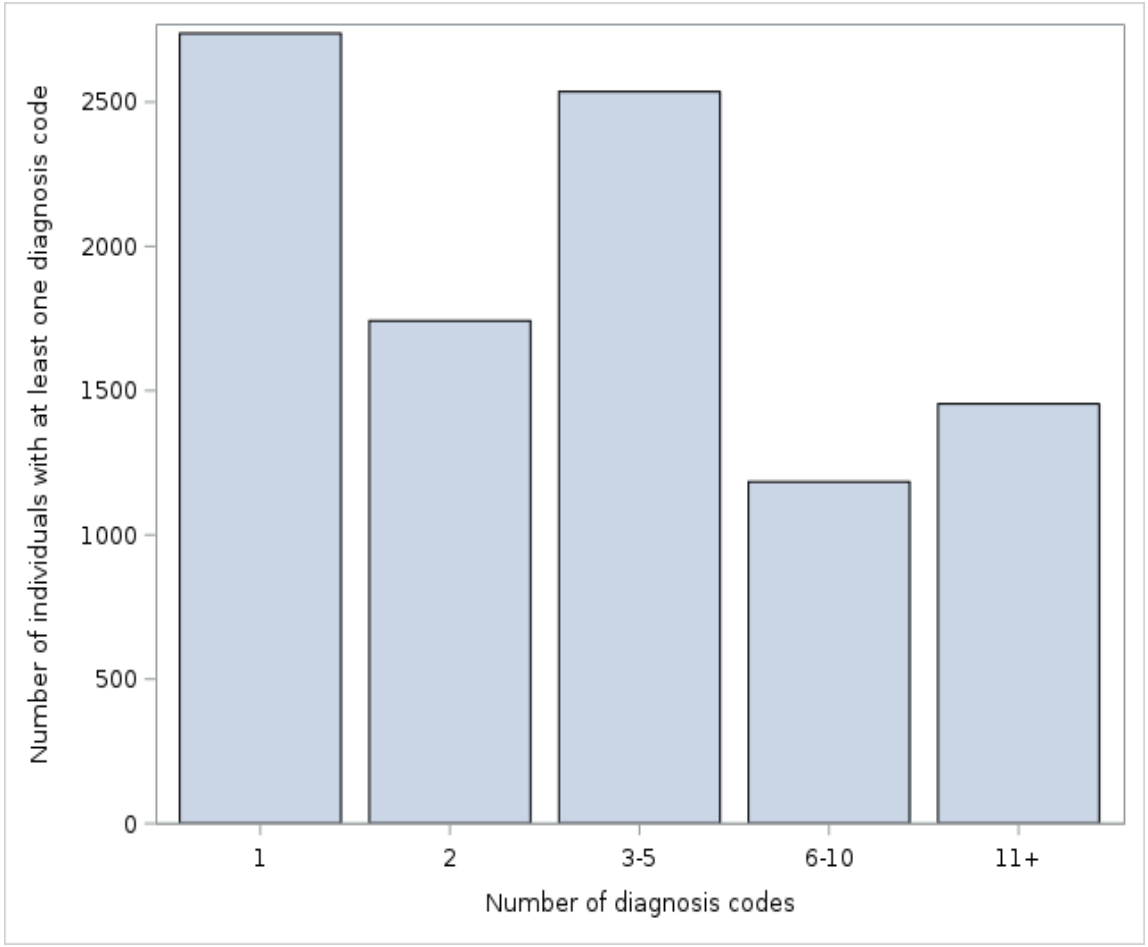


Figure H54. Number of diagnosis codes by individual, ectopic pregnancy

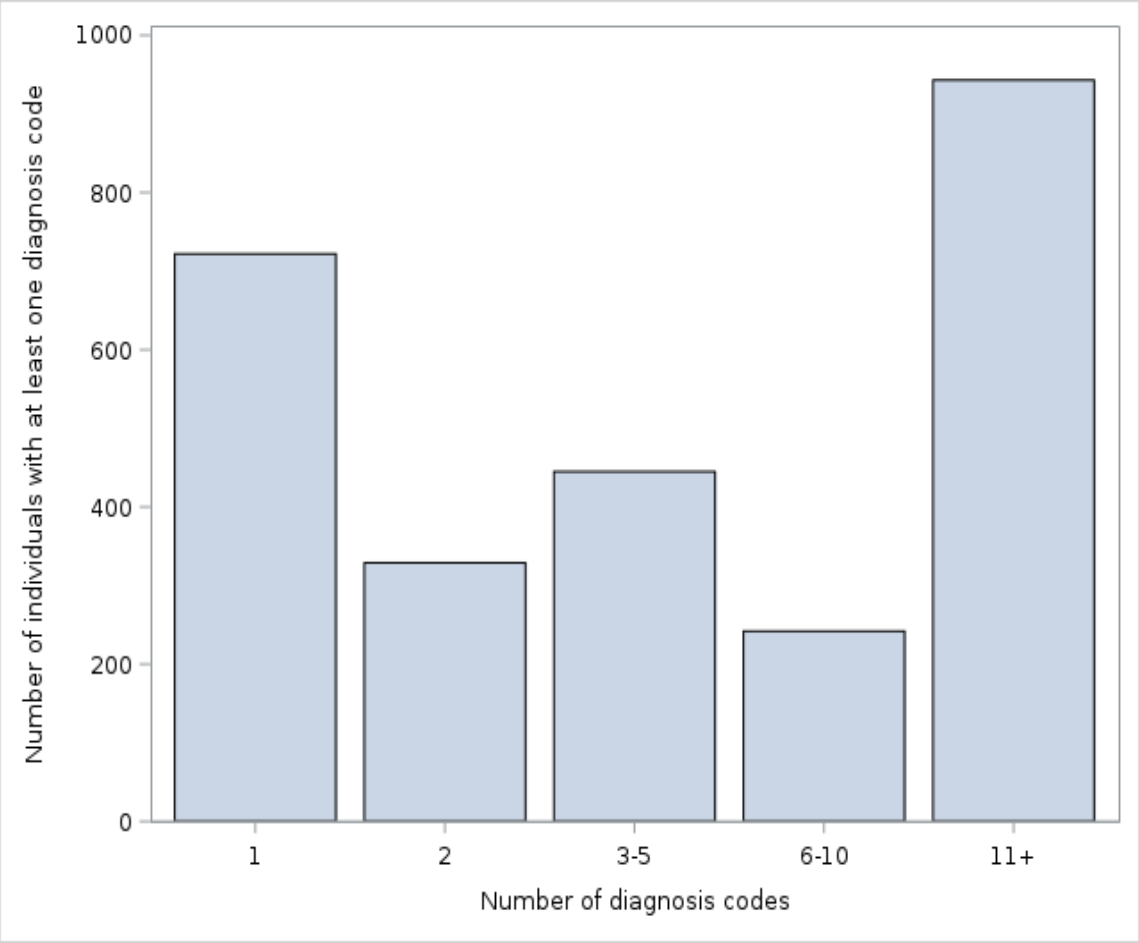


Figure H55. Number of diagnosis codes by individual, miscarriage

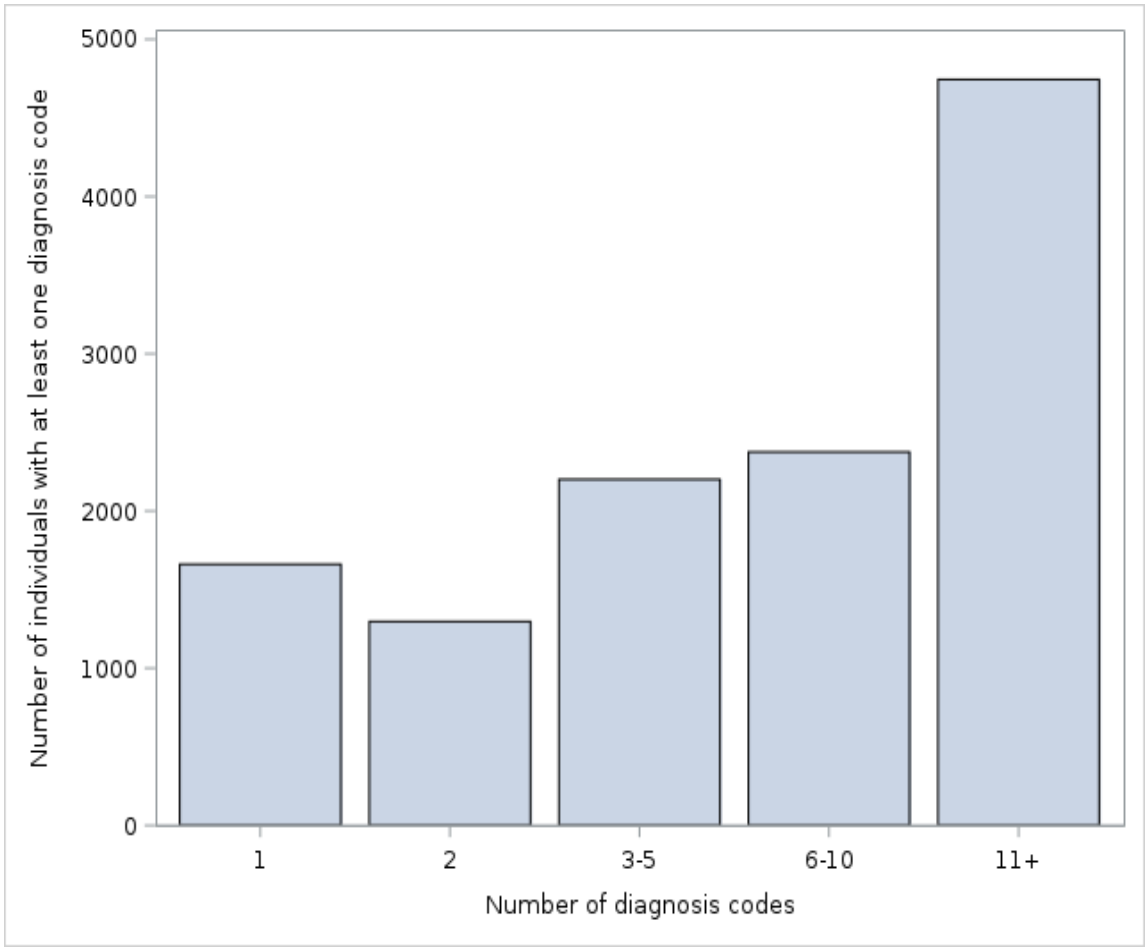


Figure H56. Number of diagnosis codes by individual, gestational diabetes

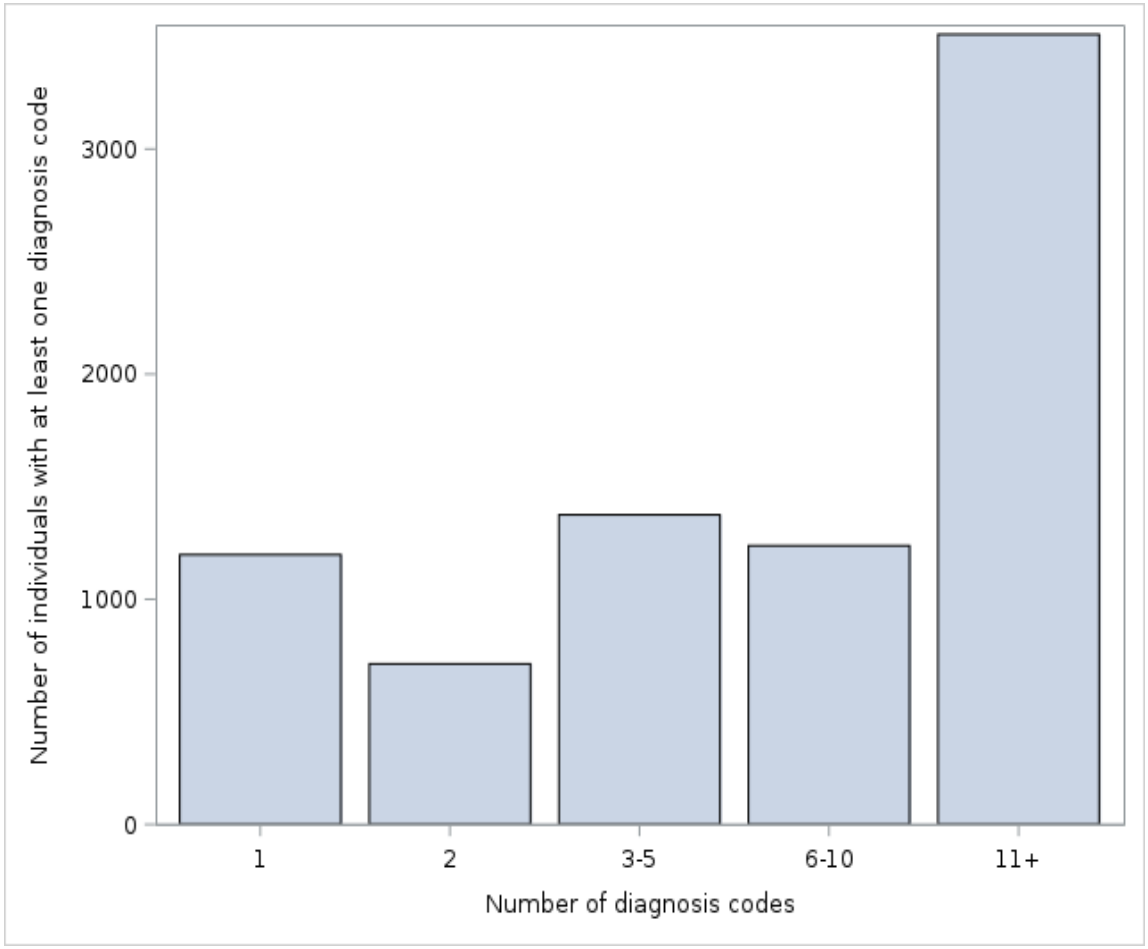


Figure H57. Number of diagnosis codes by individual, other maternal hypertension

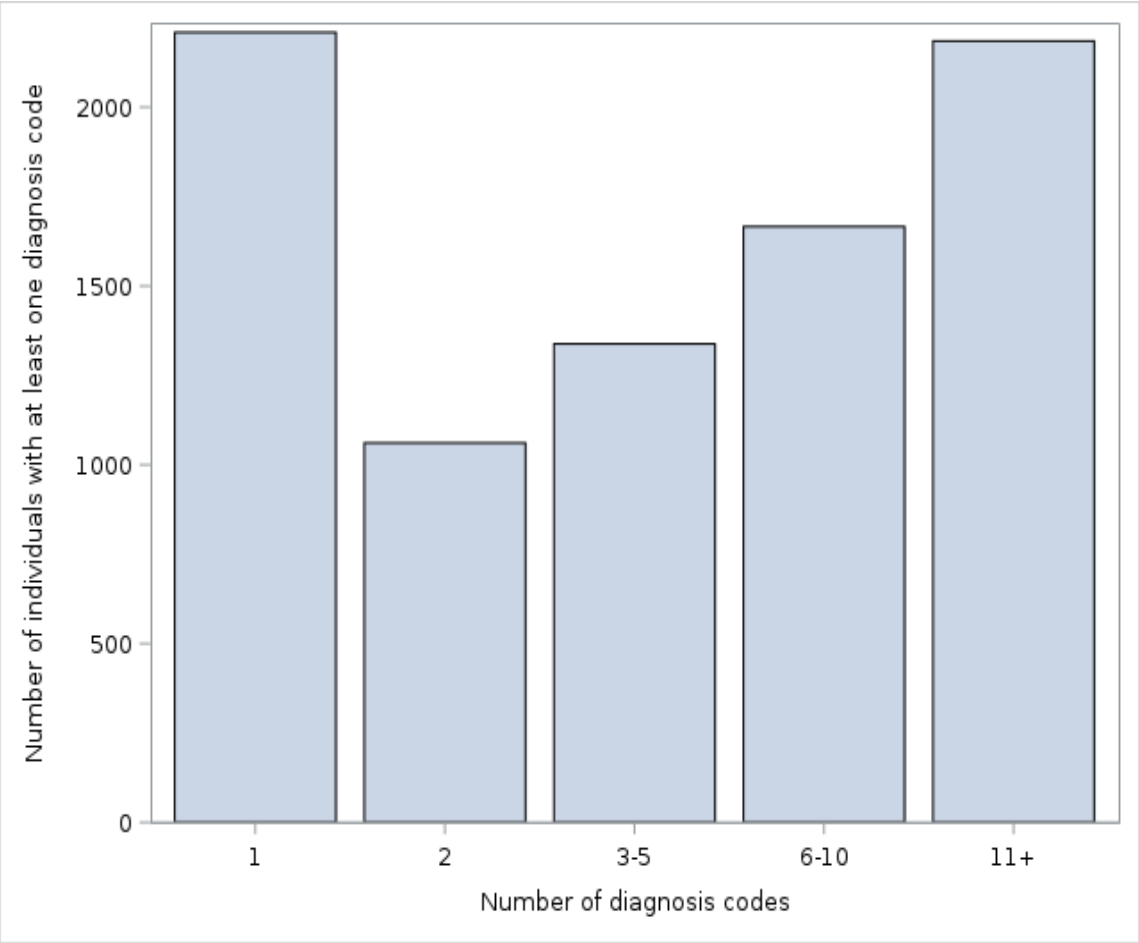


Figure H58. Number of diagnosis codes by individual, male infertility

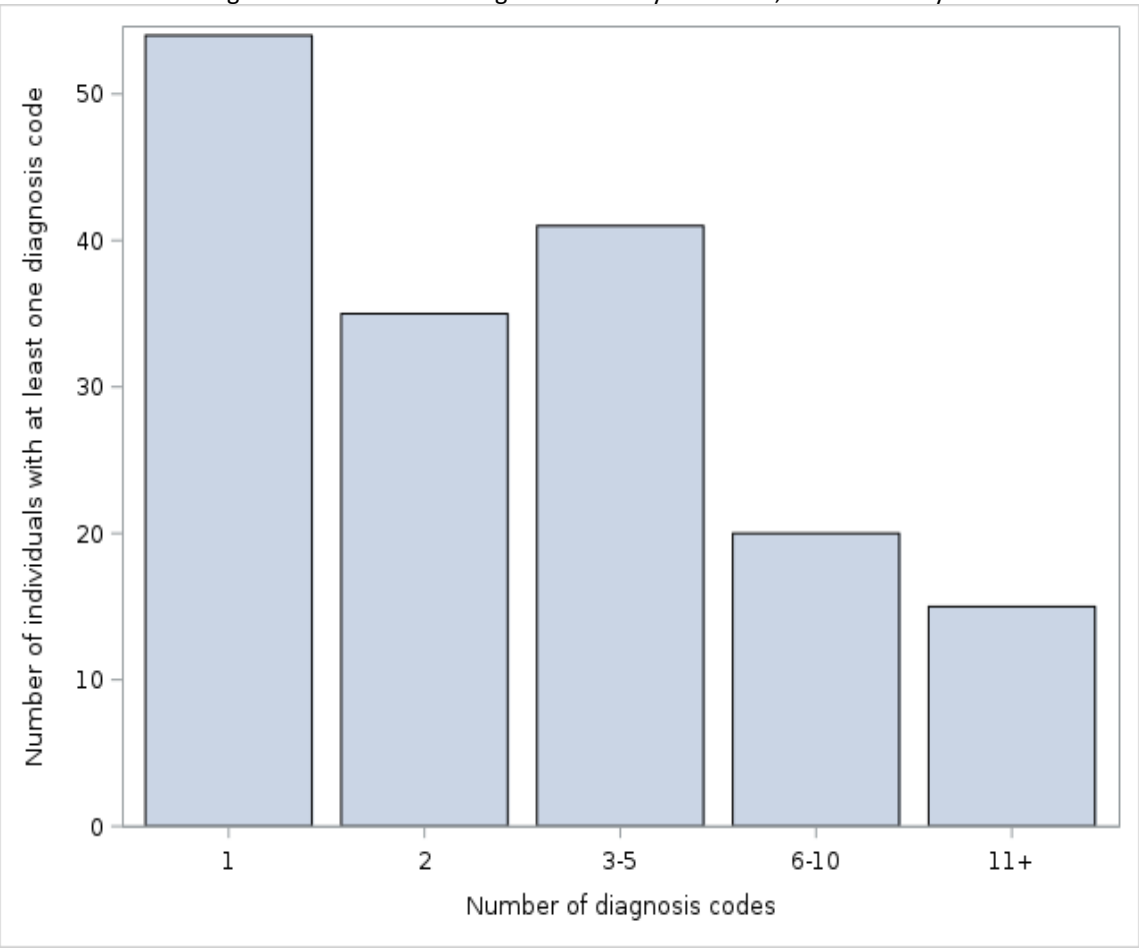


Figure H59. Number of diagnosis codes by individual, female infertility

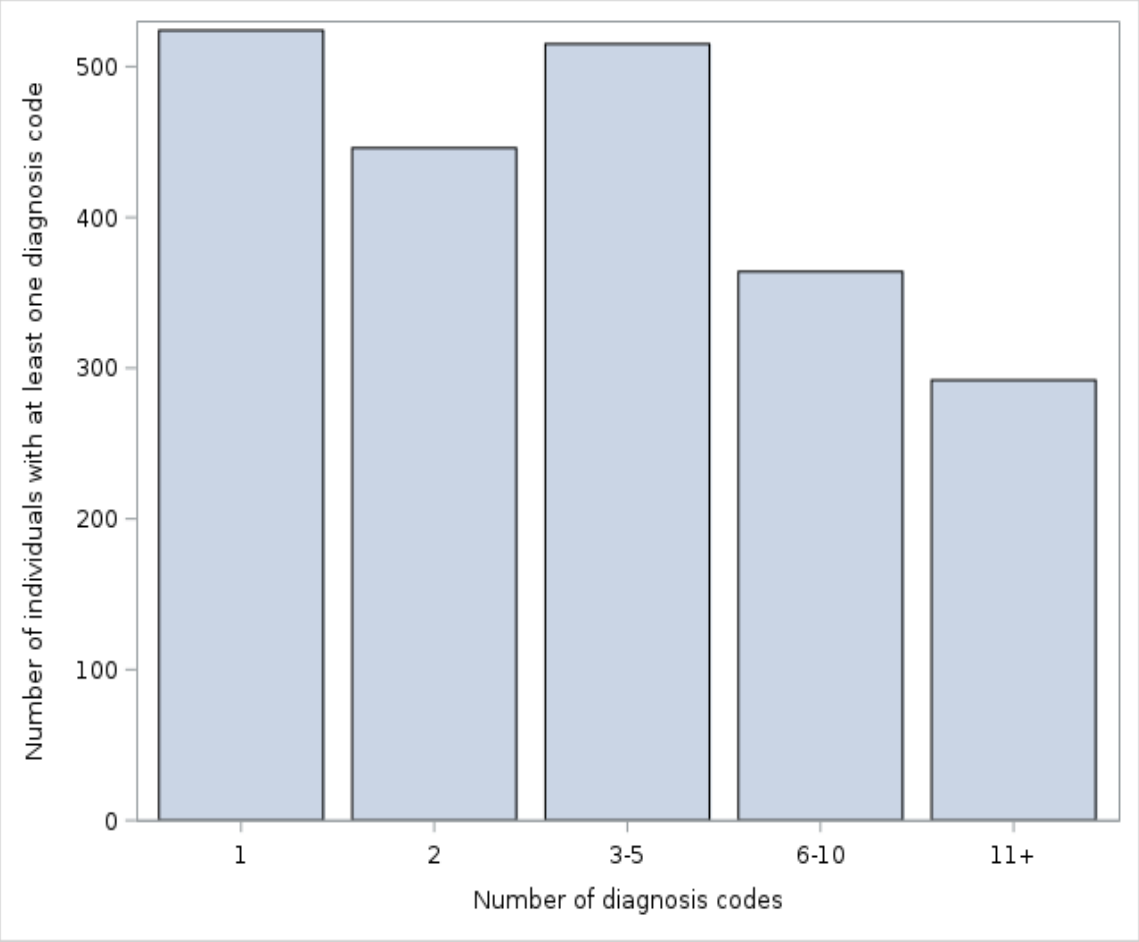


Figure H60. Number of diagnosis codes by individual, uterine fibroids

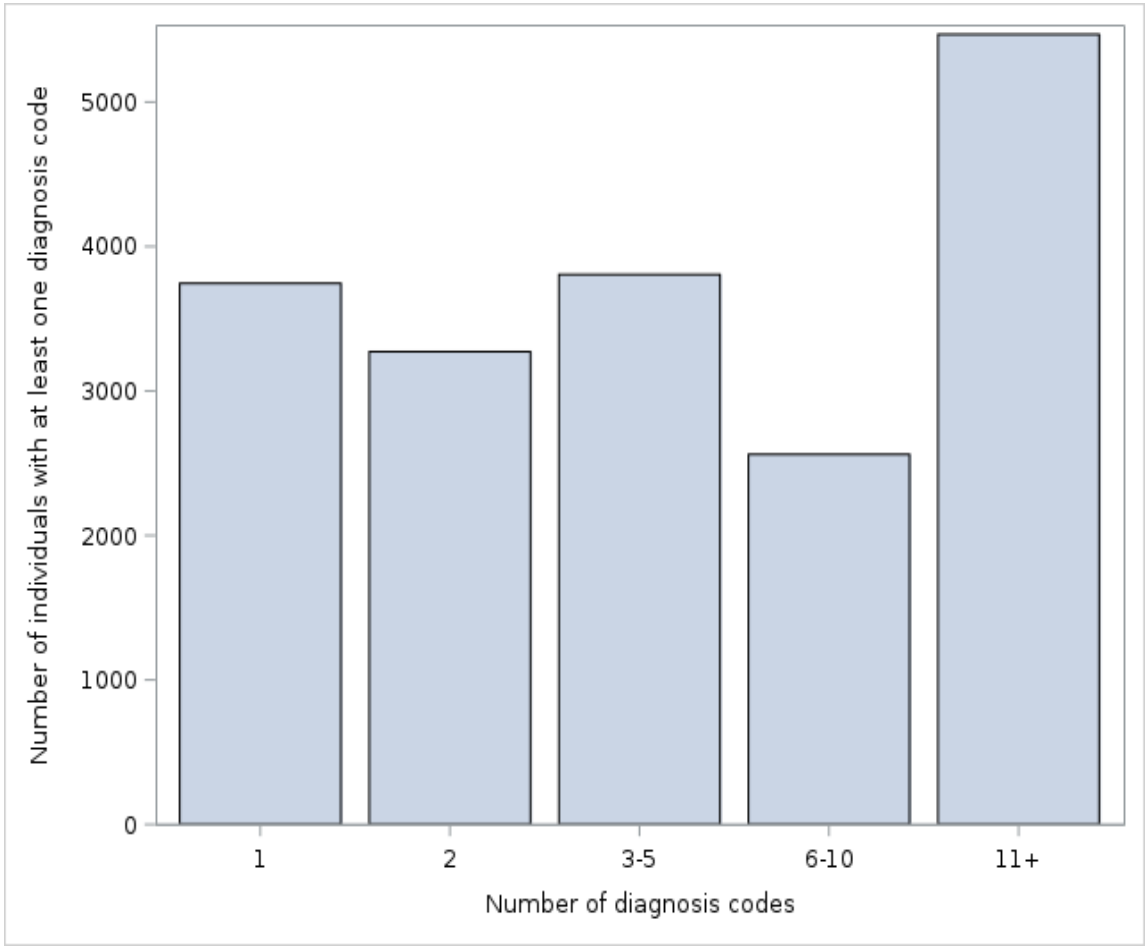


Figure H61. Number of diagnosis codes by individual, early puberty

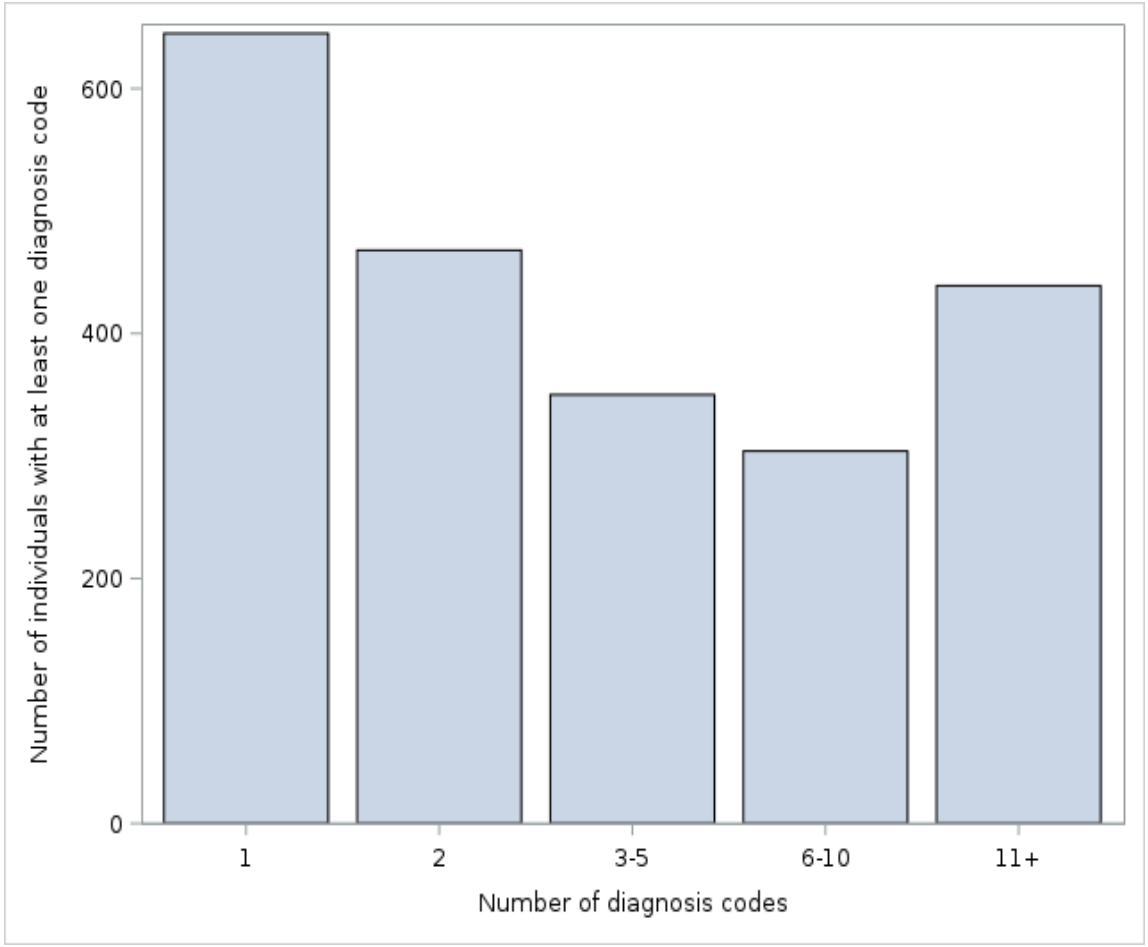


Figure H62. Number of diagnosis codes by individual, menopause

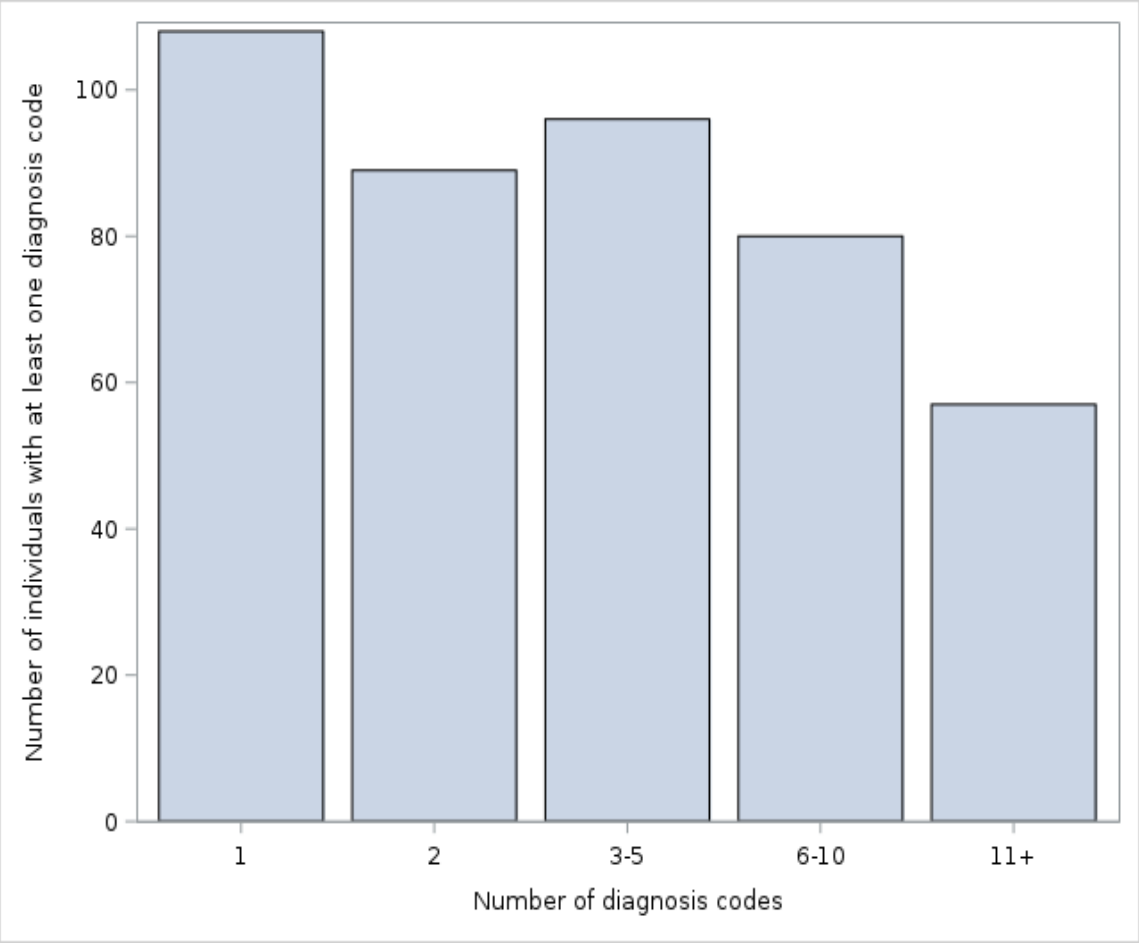


Figure H63. Number of diagnosis codes by individual, asthma

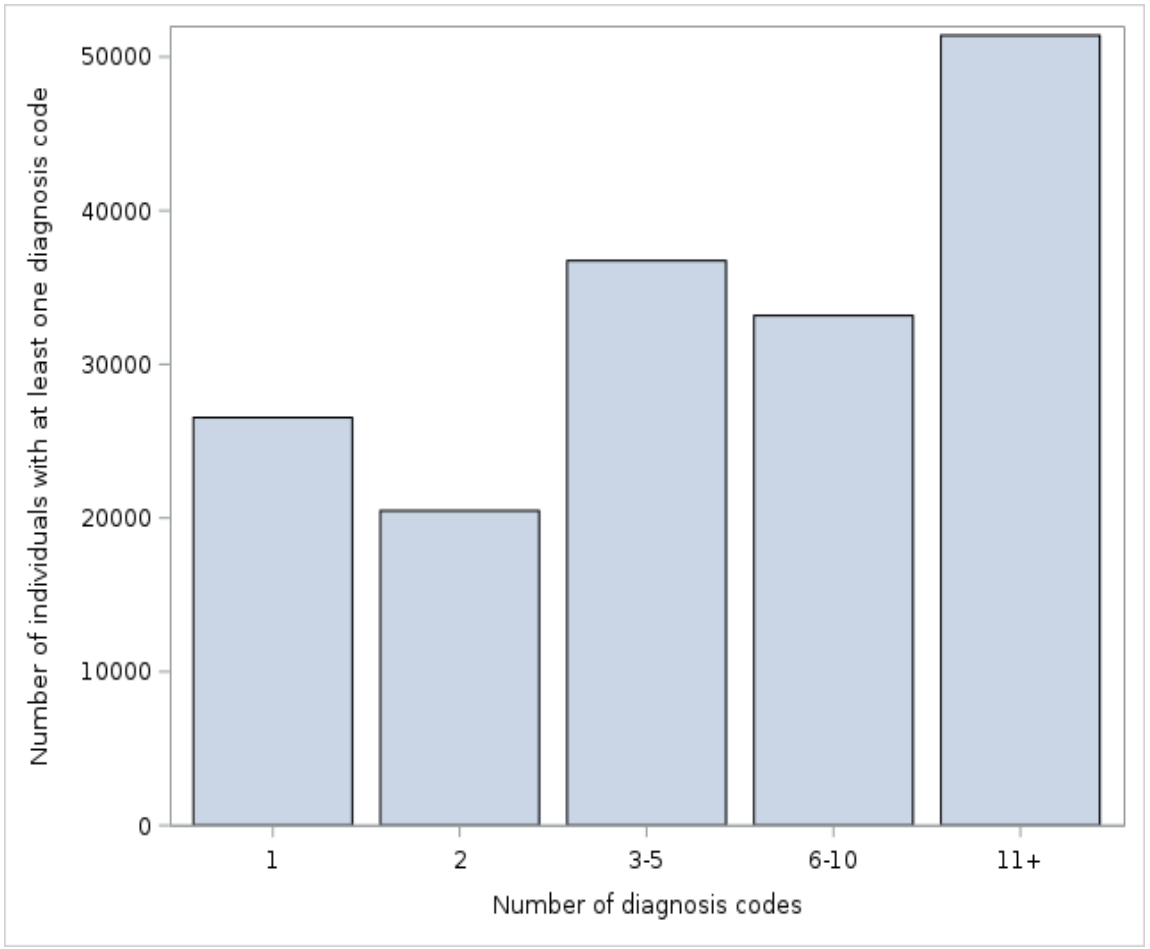


Figure H64. Number of diagnosis codes by individual, chronic obstructive pulmonary disorder (COPD)

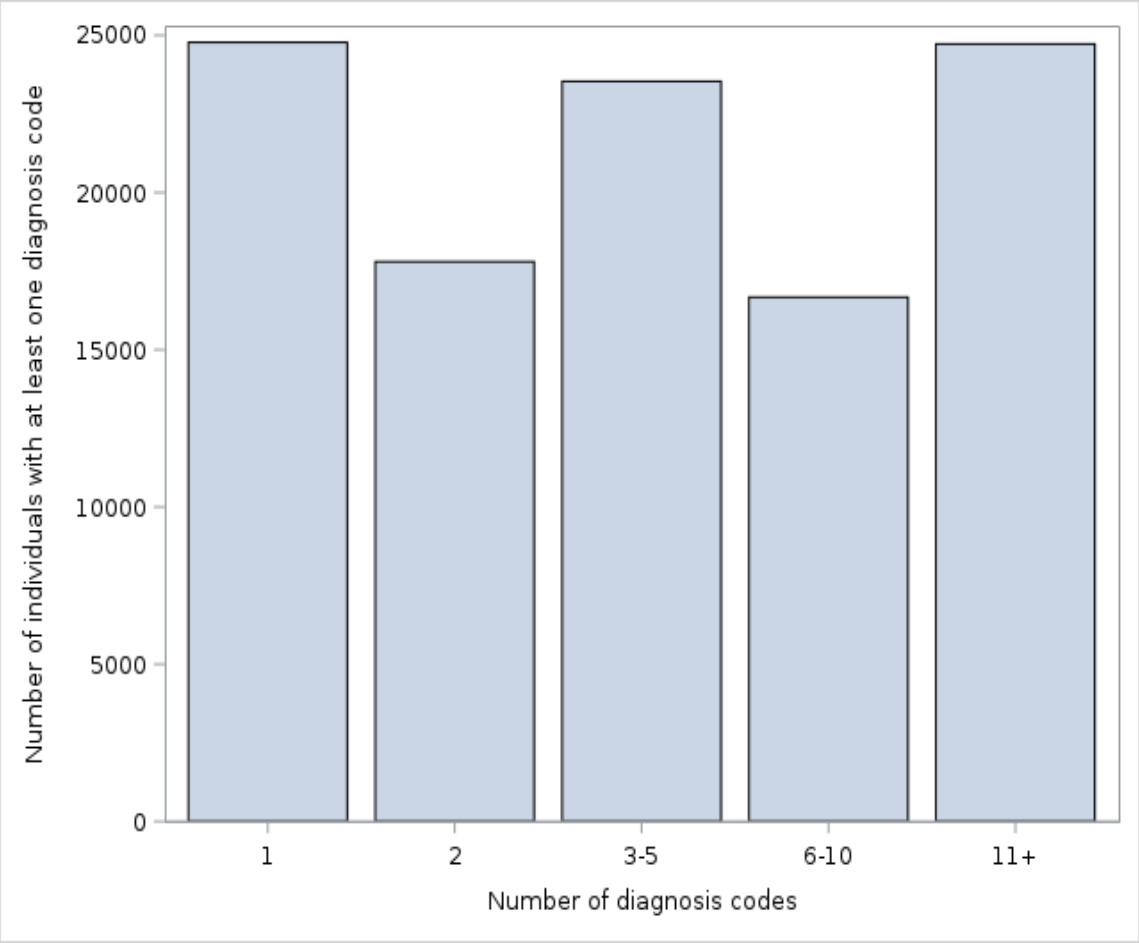


Figure H65. Number of diagnosis codes by individual, hay fever

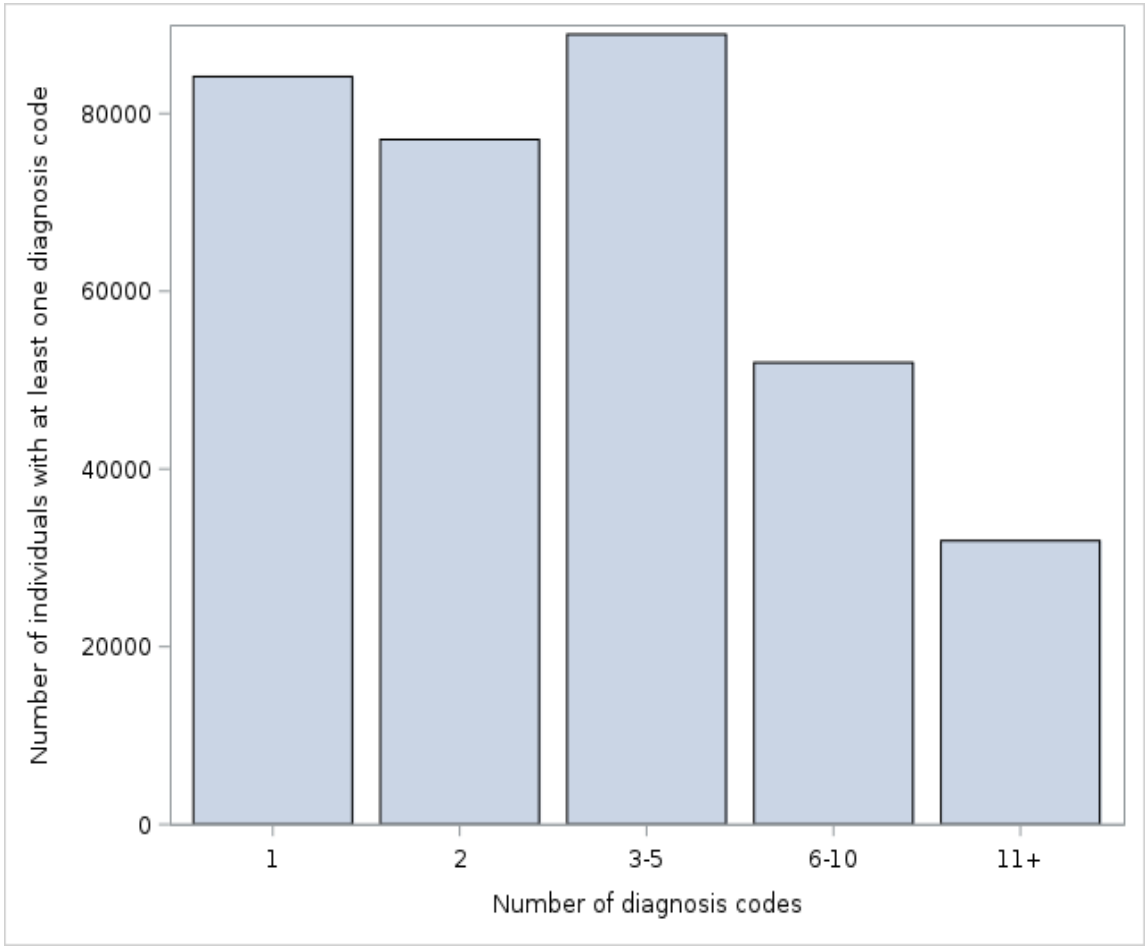


Figure H66. Number of diagnosis codes by individual, sinusitis

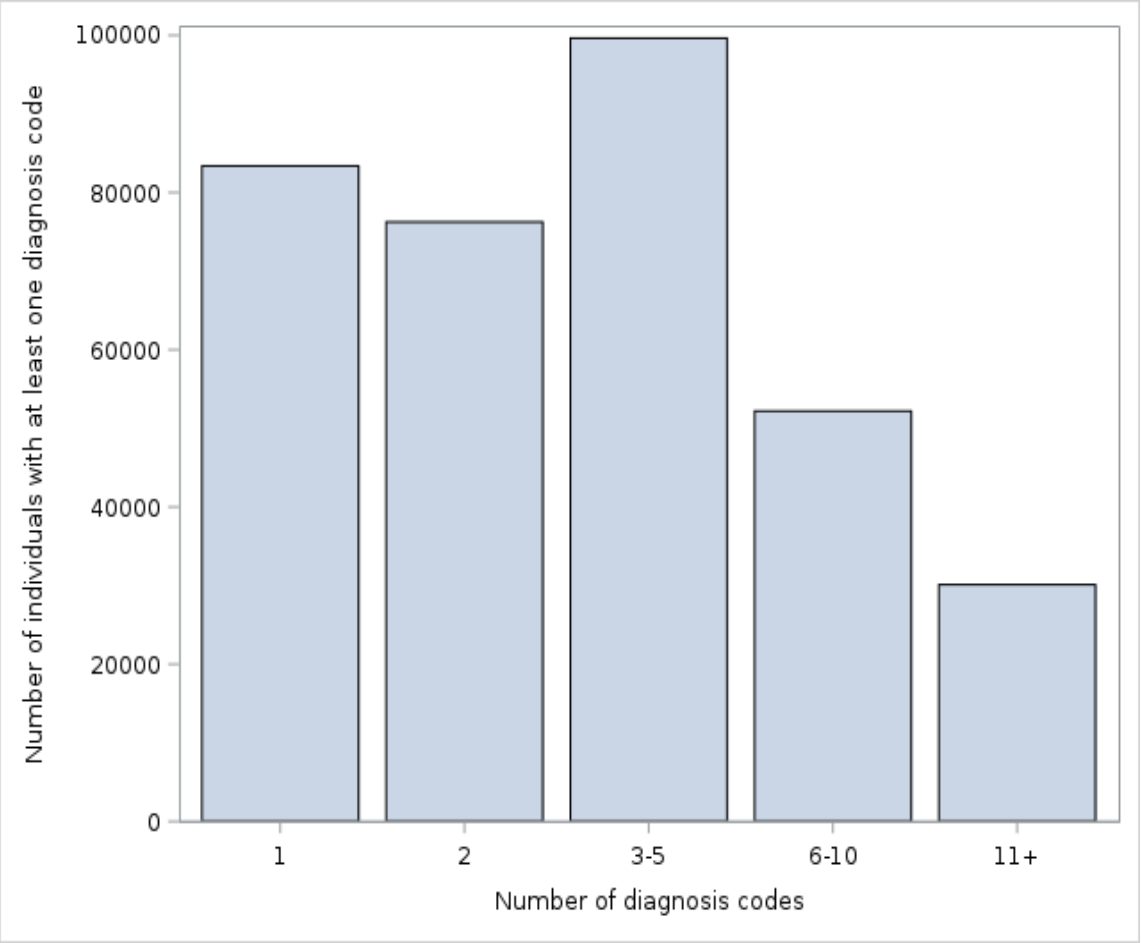


Figure H67. Number of diagnosis codes by individual, chronic rhinitis, nasopharyngitis, and pharyngitis

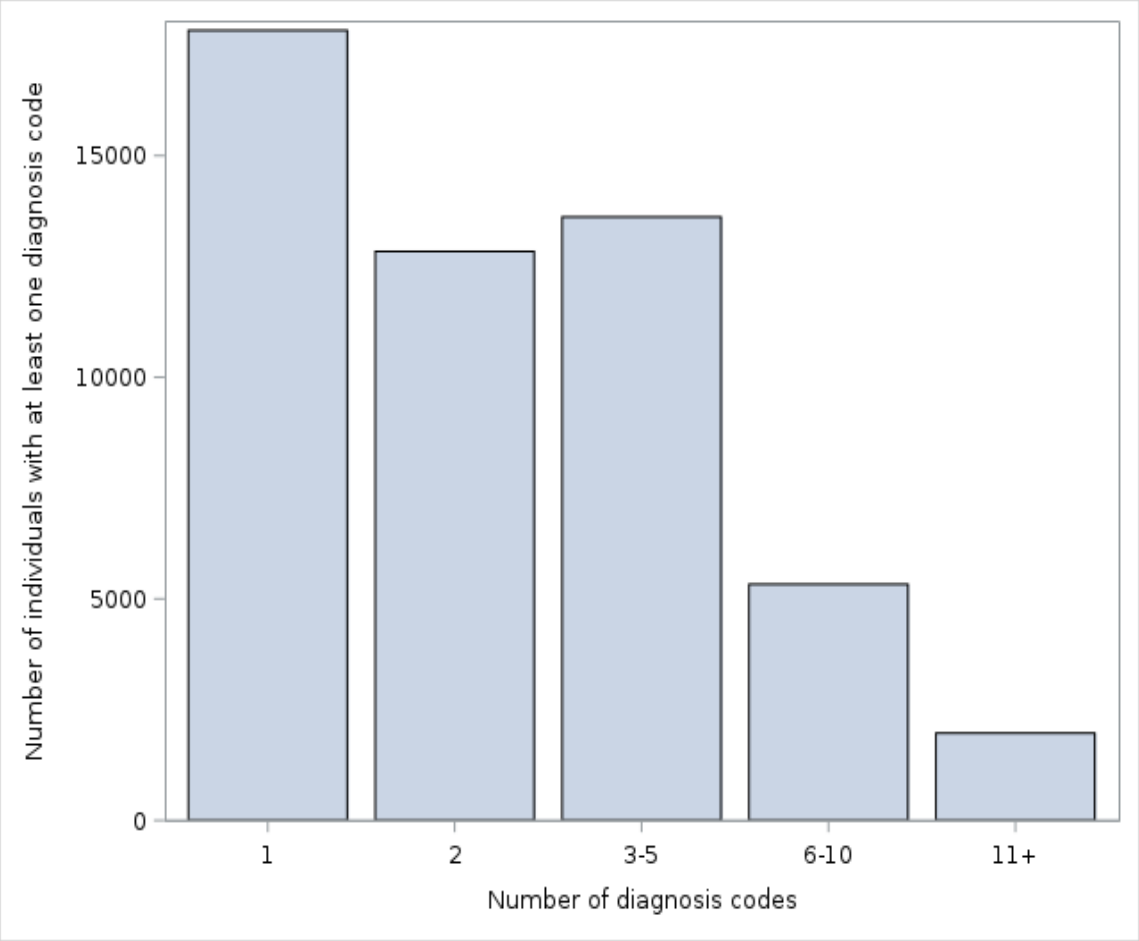


Figure H68. Number of diagnosis codes by individual, dermatitis

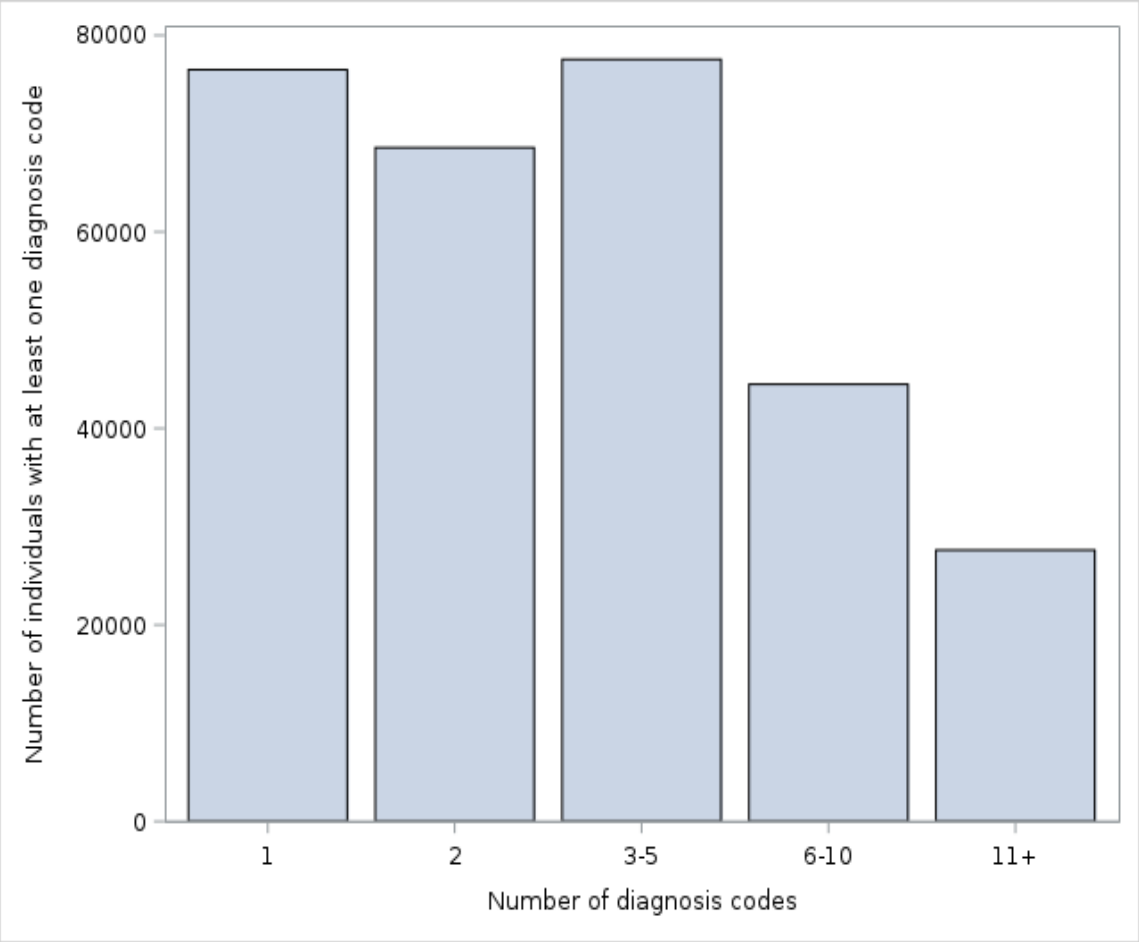
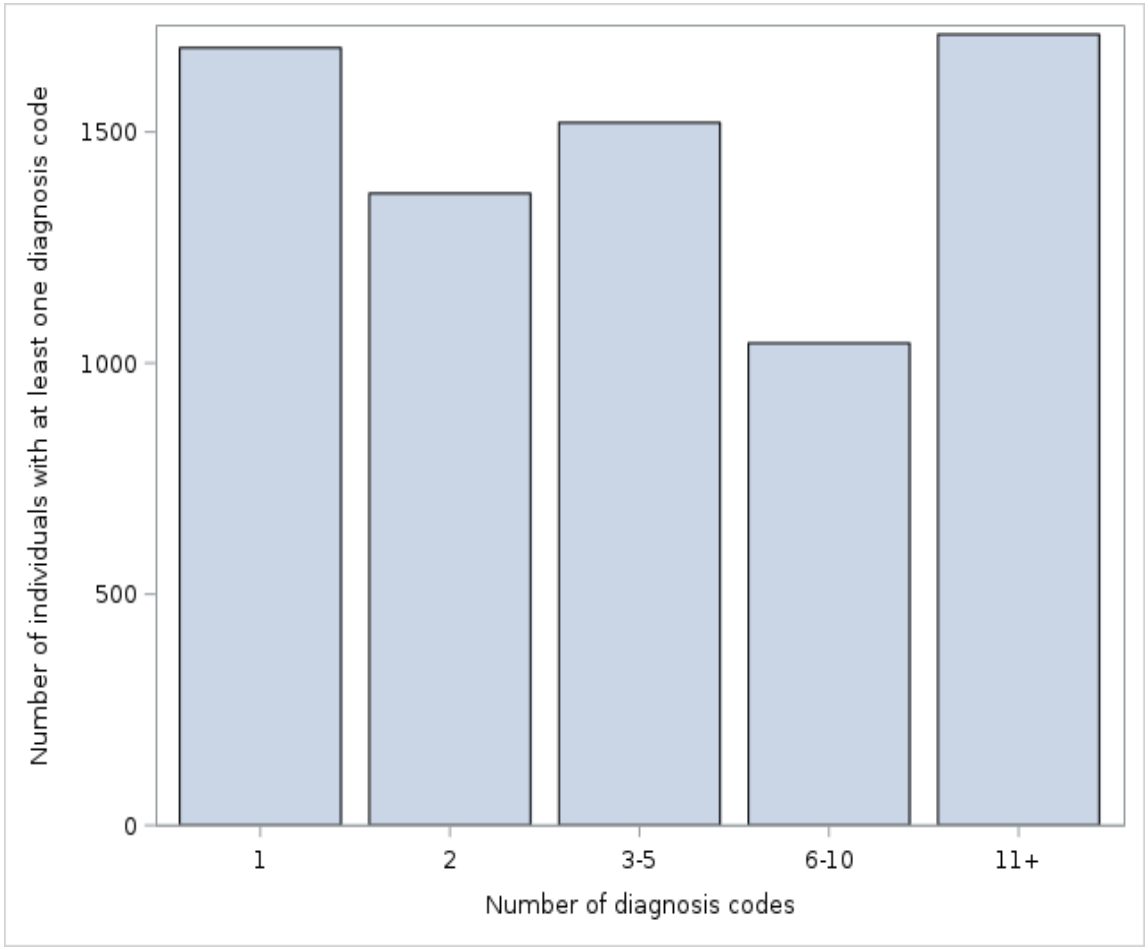


Figure H69. Number of diagnosis codes by individual, psoriasis



Appendix I. Power calculations for health outcomes with mixed findings

Power calculations help researchers determine whether a study includes enough people, both overall and within comparison groups, to reliably detect a relationship between an exposure and a health outcome. These calculations take into account how common the health outcome is and how confident we want to be in the results.

For health outcomes where our findings were mixed or showed no clear association, we conducted power calculations for models that used distance to facilities as the exposure measure. Using an online calculator, we entered the percentage of people with each health outcome in ZIP codes within 5 miles of facilities and in ZIP codes farther away (shown in Appendix D), along with standard statistical inputs (such as power and significance levels). The calculator then estimated how many participants we would need, both in total and in each ZIP code group, to reliably detect an effect. These estimates can be compared to our actual sample sizes to assess whether our study was adequately powered to find an association for that particular outcome. We are unable to calculate these estimates for health conditions where the percent did not differ in the two groups. In some instances, the prevalence values presented in Appendix D appear identical because percentages were rounded to one decimal place, although small differences may be present in the underlying unrounded estimates; unrounded estimates were used to conduct the power calculations.

In most cases, the results showed that our study did not include enough data to reliably detect an effect, even if one exists. This means that the lack of statistically significant findings should not be interpreted as evidence that there is no relationship between residential industrial pollution exposure and these health outcomes among Louisiana residents. Rather, it suggests that our study may have been too small to detect an existing association. More research with larger sample sizes or different study designs is needed to better understand these potential relationships.

Table I1. Estimated sample sizes required to reach 80% power for infant health outcomes in distance-based exposure models.

Health outcome	Total sample required for sufficient power	Sample required in each comparison group for sufficient power
Low birthweight	35,276	17,638
Preterm birth	132,134	66,067
Birth defects	4,440,258	2,220,129
Death before age 1, any cause	7,360,606	3,680,303

Table I2. Estimated sample sizes required to reach sufficient (80%) power for children's health outcomes in distance-based exposure models.		
Health outcome	**Total sample required for sufficient power**	**Sample required in each comparison group for sufficient power**
Other acquired anemias	413,426	206,713
All cancers	862,424	431,212
Asthma	7,975,160	3,987,580
Seasonal allergies/hay fever	5,082	2,541
Acute/chronic sinusitis	954	477
Chronic rhinitis, nasopharyngitis, and pharyngitis	123,570	61,785
ADHD, Conduct Disorders, and Hyperkinetic Syndrome	17,642	8,821
Intellectual disabilities and related conditions	4,215,506	2,107,753
Other developmental delays	94,562	47,281
Anxiety Disorders	35,158	17,579
Bipolar Disorder	856,416	428,208
Depressive Disorders	1,879,560	939,780
Diabetes type 1	1,097,302	548,651
Diabetes type 2	5,171,710	2,585,855
Thyroid gland disorders	331,908	165,954
Migraine and other chronic headache	54,556	27,278
Epilepsy	196,972	98,486

Table I3. Estimated sample sizes required to reach sufficient (80%) power for women's health outcomes in distance-based exposure models.		
Health outcome	**Total sample required for sufficient power**	**Sample required in each comparison group for sufficient power**
Female reproductive system cancers	3,749,376	1,874,688
Female infertility	3,293,066	1,646,533
Early menopause	470,860	235,430
Death before age 65, any cause	3,059,902	1,529,951

Table I4. Estimated sample sizes required to reach 80% power for pregnancy-related health outcomes in distance-based exposure models.		
Health outcome	**Total sample required for sufficient power**	**Sample required in each comparison group for sufficient power**
Other maternal hypertension	211,188	105,594
Gestational diabetes	216,730	108,365
Death before age 65, any cause	574,632	287,316

Table I5. Estimated sample sizes required to reach 80% power for men's health outcomes in distance-based exposure models.		
Health outcome	**Total sample required for sufficient power**	**Sample required in each comparison group for sufficient power**
Prostate cancer	8,297,762	4,148,881

Table I6. Estimated sample sizes required to reach 80% power for selected adult health outcomes in distance-based exposure models.		
Health outcome	**Total sample required for sufficient power**	**Sample required in each comparison group for sufficient power**
Lupus	748,390	374,195
Multiple sclerosis	3,919,532	1,959,766
Systemic sclerosis	470,860	235,430
Myositis	470,860	235,430
Colorectal cancer	5,171,710	2,585,855
Lung cancer	3,293,066	1,646,533
Thyroid cancer	2,039,380	1,019,690
All cancers	1,104,736	552,368
Heart attack	1,060,712	530,356
Heart failure and nonischemic heart disease	551,986	275,993
Ischemic heart disease	17,378	8,689
Peripheral vascular disease	54,872	27,436
Hypertension	141,104	70,552
Stroke	897,546	448,773
Hypothyroidism	20,684	10,342
Diabetes, type 2	630,816	315,408
Thyroid gland disorders	16,918	8,459
Other glucose metabolism disorder	539,860	269,930
Gastroesophageal reflux disease	13,458	6,729
Irritable bowel disorders	35,998	17,999
Anxiety disorders	5,182	2,591
Depressive disorders	20,164	10,082

Arthritis	17,166	8,583
Chronic kidney disease	309,544	154,772
Migraine and chronic headache	138,734	69,367
Epilepsy	197,722	98,861
Obesity	6,205,460	3,102,730
Asthma	94,502	47,251
COPD	15,328	7,664
Seasonal allergies/hay fever	76,068	38,034
Acute/chronic sinusitis	3,450	1,725
Chronic rhinitis, nasopharyngitis, and pharyngitis	887,270	443,635
Dermatitis and eczema	52,958	26,479
Death before age 65, any cause	2,498,608	1,249,304