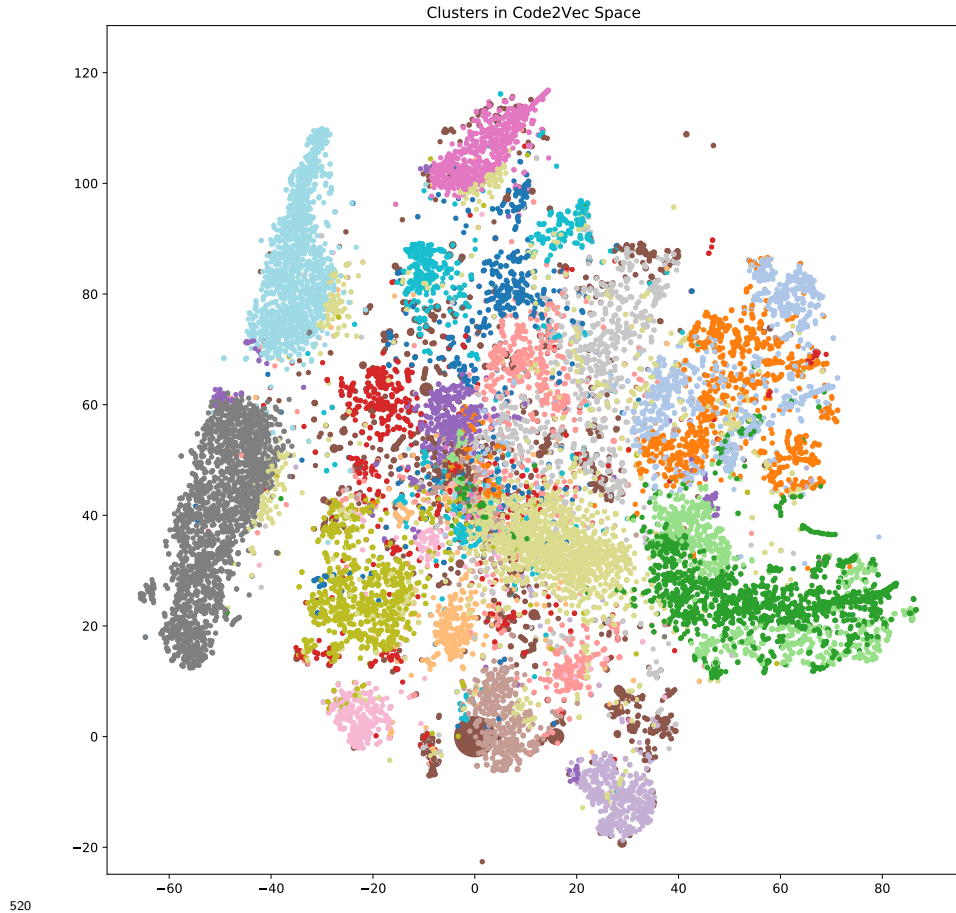


Appendix A. Diagnosis Code Clusters

Appendix A.1. *t*-SNE Visualization of Code2Vec Clusters



Each spot represents a diagnosis code in Code2Vec embedding space. The the codes belonging to the same cluster have the same color. The size of the spot represents the frequency of the code in the claim history records.

Appendix A.2. Top-15 Frequent Codes in Clusters

The top-15 frequent codes in each cluster are shown below. The color of the subtitle is consistent with the visualization. For each cluster, the codes are presented from the most frequent the the least frequent. The codes are followed by their description in the ICD-9-CM or ICD-10-CM codebase. If description is not given, that code was incorrectly recorded in the dataset that can not be found in the codebases.

cluster #0 •

ICD10-R531 Weakness

ICD10-R569 Unspecified convulsions

ICD10-I639 Cerebral infarction, unspecified
 535 ICD10-R2689 Other abnormalities of gait and mobility
 ICD10-J9600 Acute respiratory failure, unspecified whether with hypoxia or hypercapnia
 ICD10-R4182 Altered mental status, unspecified
 ICD10-R269 Unspecified abnormalities of gait and mobility
 540 ICD10-G40909 Epilepsy, unspecified, not intractable, without status epilepticus
 ICD10-I959 Hypotension, unspecified
 ICD10-C7931 Secondary malignant neoplasm of brain
 ICD10-R413 Other amnesia
 ICD10-R45851 Suicidal ideations
 545 ICD10-Z8673 Personal history of transient ischemic attack (TIA), and cerebral infarction without residual deficits
 ICD10-J9690 Respiratory failure, unspecified, unspecified whether with hypoxia or hypercapnia
 ICD10-J690 Pneumonitis due to inhalation of food and vomit
 550 **cluster #1 •**
 ICD09-496 Chronic airway obstruction, not elsewhere classified
 ICD09-42731 Atrial fibrillation
 ICD09-25002 Diabetes mellitus without mention of complication, type II or unspecified type, uncontrolled
 555 ICD09-78609 Other respiratory abnormalities
 ICD09-41401 Coronary atherosclerosis of native coronary artery
 ICD09-41400 Coronary atherosclerosis of unspecified type of vessel, native or graft
 ICD09-78701 Nausea with vomiting
 560 ICD09-1749 Malignant neoplasm of breast (female), unspecified
 ICD09-486 Pneumonia, organism unspecified
 ICD09-5920 Calculus of kidney
 ICD09-V1582 Personal history of tobacco use
 ICD09-7802 Syncope and collapse
 565 ICD09-56210 Diverticulosis of colon (without mention of hemorrhage)
 ICD09-7823 Edema
 ICD09-78906 Abdominal pain, epigastric
cluster #2 •
 ICD09-V700 Routine general medical examination at a health care facility
 570 ICD09-4011 Benign essential hypertension
 ICD09-2449 Unspecified acquired hypothyroidism
 ICD09-53081 Esophageal reflux
 ICD09-V7231 Routine gynecological examination
 ICD09-78079 Other malaise and fatigue
 575 ICD09-V7612 Other screening mammogram
 ICD09-78650 Chest pain, unspecified
 ICD09-311 Depressive disorder, not elsewhere classified
 ICD09-7295 Pain in limb
 ICD09-3051 Tobacco use disorder

580 ICD09-30000 Anxiety state, unspecified
 ICD09-78900 Abdominal pain, unspecified site
 ICD09-7231 Cervicalgia
 ICD09-4779 Allergic rhinitis, cause unspecified
 cluster #3 •

585 ICD10-R21 Rash and other nonspecific skin eruption
 ICD10-L700 Acne vulgaris
 ICD10-D485 Neoplasm of uncertain behavior of skin
 ICD10-L309 Dermatitis, unspecified
 ICD10-L821 Other seborrheic keratosis
 590 ICD10-D225 Melanocytic nevi of trunk
 ICD10-L814 Other melanin hyperpigmentation
 ICD10-L570 Actinic keratosis
 ICD10-L989 Disorder of the skin and subcutaneous tissue, unspecified
 ICD10-L720 Epidermal cyst
 595 ICD10-L299 Pruritus, unspecified
 ICD10-L400 Psoriasis vulgaris
 ICD10-L409 Psoriasis, unspecified
 ICD10-L089 Local infection of the skin and subcutaneous tissue, unspecified
 ICD10-L298 Other pruritus
 600 cluster #4 •

ICD09-I10
 ICD09-E11.9
 ICD09-616.10 Vaginitis and vulvovaginitis, unspecified
 ICD09-E78.5
 605 ICD09-V06.1 Need for prophylactic vaccination and inoculation against diphtheria-
 tetanus-pertussis, combined [DTP] [DTaP]
 ICD09-V20.2 Routine infant or child health check (dotless)
 ICD09-E03.9
 ICD09-Z00.00
 610 ICD09-Z79.899
 ICD09-Z12.31
 ICD09-V74.5 Screening examination for venereal disease
 ICD09-M54.5
 ICD09-Z23
 615 ICD09-218.9 Leiomyoma of uterus, unspecified
 ICD09-626.2 Excessive or frequent menstruation
 cluster #5 •

ICD09-V70.0 Routine general medical examination at a health care facility
 ICD09-401.9 Unspecified essential hypertension
 620 ICD09-272.4 Other and unspecified hyperlipidemia
 ICD09-250.00 Diabetes mellitus without mention of complication, type II or
 unspecified type, not stated as uncontrolled
 ICD09-780.79 Other malaise and fatigue
 ICD09-401.1 Benign essential hypertension
 625 ICD09-724.2 Lumbago

ICD09-V72.31 Routine gynecological examination
 ICD09-244.9 Unspecified acquired hypothyroidism
 ICD09-V76.12 Other screening mammogram
 ICD09-268.9 Unspecified vitamin D deficiency
 630 ICD09-285.9 Anemia, unspecified
 ICD09-786.50 Chest pain, unspecified
 ICD09-530.81 Esophageal reflux
 ICD09-789.00 Abdominal pain, unspecified site
 cluster #6 •
 635 ICD10-M549 Dorsalgia, unspecified
 ICD10-G894 Chronic pain syndrome
 ICD10-M5136 Other intervertebral disc degeneration, lumbar region
 ICD09-7242 Lumbago
 ICD10-M5126 Other intervertebral disc displacement, lumbar region
 640 ICD10-C7951 Secondary malignant neoplasm of bone
 ICD10-M47816 Spondylosis without myelopathy or radiculopathy, lumbar region
 ICD10-G629 Polyneuropathy, unspecified
 ICD10-M4806 Spinal stenosis, lumbar region
 ICD10-M47812 Spondylosis without myelopathy or radiculopathy, cervical region
 645 ICD10-M4802 Spinal stenosis, cervical region
 ICD10-M461 Sacroiliitis, not elsewhere classified
 ICD10-M961 Postlaminectomy syndrome, not elsewhere classified
 ICD10-M5137 Other intervertebral disc degeneration, lumbosacral region
 ICD09-7245 Backache, unspecified
 650 cluster #7 •
 ICD10-I2510 Atherosclerotic heart disease of native coronary artery without
 angina pectoris
 ICD10-N179 Acute kidney failure, unspecified
 ICD10-I509 Heart failure, unspecified
 655 ICD10-R600 Localized edema
 ICD10-I480 Paroxysmal atrial fibrillation
 ICD10-Z7901 Long term (current) use of anticoagulants
 ICD10-E1142 Type 2 diabetes mellitus with diabetic polyneuropathy
 ICD10-I739 Peripheral vascular disease, unspecified
 660 ICD10-E1140 Type 2 diabetes mellitus with diabetic neuropathy, unspecified
 ICD10-D62 Acute posthemorrhagic anemia
 ICD10-R262 Difficulty in walking, not elsewhere classified
 ICD10-M109 Gout, unspecified
 ICD10-M7989 Other specified soft tissue disorders
 665 ICD10-K7460 Unspecified cirrhosis of liver
 ICD10-R0989 Other specified symptoms and signs involving the circulatory and
 respiratory systems
 cluster #8 •
 ICD10-E7849 Other hyperlipidemia
 670 ICD10-M7918 Myalgia, other site
 ICD10-M7981 Nontraumatic hematoma of soft tissue

ICD10-Y95 Nosocomial condition
 ICD10-Z883 Allergy status to other anti-infective agents status
 ICD10-Z8614 Personal history of Methicillin resistant Staphylococcus aureus
 675 infection
 ICD10-C162 Malignant neoplasm of body of stomach
 ICD10-F40240 Claustrophobia
 ICD10-T402X5A Adverse effect of other opioids, initial encounter
 ICD10-B3789 Other sites of candidiasis
 680 ICD10-M7591 Shoulder lesion, unspecified, right shoulder
 ICD10-T8612 Kidney transplant failure
 ICD10-M65819 Other synovitis and tenosynovitis, unspecified shoulder
 ICD09-7944 Nonspecific abnormal results of function study of kidney
 ICD10-M67813 Other specified disorders of tendon, right shoulder
 685 cluster #9 •
 ICD10-Z36 Encounter for antenatal screening for chromosomal anomalies
 ICD10-Z3482 Encounter for supervision of other normal pregnancy, second
 trimester
 ICD10-Z3A39 39 weeks gestation of pregnancy
 690 ICD10-080 Encounter for full-term uncomplicated delivery
 ICD10-Z3A01 Less than 8 weeks gestation of pregnancy
 ICD10-009523 Supervision of elderly multigravida, third trimester
 ICD10-Z331 Pregnant state, incidental
 ICD10-034211 Maternal care for low transverse scar from previous cesarean de-
 695 livery
 ICD10-Z3A38 38 weeks gestation of pregnancy
 ICD10-Z3403 Encounter for supervision of normal first pregnancy, third trimester
 ICD10-Z3A37 37 weeks gestation of pregnancy
 ICD10-Z3490 Encounter for supervision of normal pregnancy, unspecified, un-
 700 specified trimester
 ICD10-Z3A36 36 weeks gestation of pregnancy
 ICD10-Z3401 Encounter for supervision of normal first pregnancy, first trimester
 ICD10-Z3402 Encounter for supervision of normal first pregnancy, second trimester
 cluster #10 •
 705 ICD10-0 ICD10-I10 Essential (primary) hypertension
 ICD10-Z0000 Encounter for general adult medical examination without abnor-
 mal findings
 ICD10-E119 Type 2 diabetes mellitus without complications
 ICD10-Z23 Encounter for immunization
 710 ICD10-E785 Hyperlipidemia, unspecified
 ICD10-Z01419 Encounter for gynecological examination (general) (routine) with-
 out abnormal findings
 ICD10-Z1231 Encounter for screening mammogram for malignant neoplasm of
 breast
 715 ICD10-E039 Hypothyroidism, unspecified
 ICD10-M545 Low back pain
 ICD10-Z1211 Encounter for screening for malignant neoplasm of colon

ICD10-E559 Vitamin D deficiency, unspecified
 ICD10-E782 Mixed hyperlipidemia
 720 ICD10-K219 Gastro-esophageal reflux disease without esophagitis
 ICD10-R05 Cough
 cluster #11 •
 ICD10-H2513 Age-related nuclear cataract, bilateral
 ICD10-H04123 Dry eye syndrome of bilateral lacrimal glands
 725 ICD10-H524 Presbyopia
 ICD10-H5213 Myopia, bilateral
 ICD10-H2512 Age-related nuclear cataract, left eye
 ICD10-H538 Other visual disturbances
 ICD10-H2511 Age-related nuclear cataract, right eye
 730 ICD10-H52223 Regular astigmatism, bilateral
 ICD10-H40013 Open angle with borderline findings, low risk, bilateral
 ICD10-Z961 Presence of intraocular lens
 ICD10-H1045 Other chronic allergic conjunctivitis
 ICD10-H269 Unspecified cataract
 735 ICD10-H527 Unspecified disorder of refraction
 ICD10-H5203 Hypermetropia, bilateral
 ICD10-H1013 Acute atopic conjunctivitis, bilateral
 cluster #12 •
 ICD10-4019
 740 ICD10-V700
 ICD10-F314 Bipolar disorder, current episode depressed, severe, without psychotic features
 ICD10-F1123 Opioid dependence with withdrawal
 ICD10-F901 Attention-deficit hyperactivity disorder, predominantly hyperactive type
 745 ICD10-2724
 ICD10-7242
 ICD10-25000
 ICD10-F4001 Agoraphobia with panic disorder
 750 ICD10-F429 Obsessive-compulsive disorder, unspecified
 ICD10-78650
 ICD10-4011
 ICD10-F259 Schizoaffective disorder, unspecified
 ICD10-F913 Oppositional defiant disorder
 755 ICD10-4779
 cluster #13 •
 ICD10-M79671 Pain in right foot
 ICD10-M79672 Pain in left foot
 ICD10-M25572 Pain in left ankle and joints of left foot
 760 ICD10-M25571 Pain in right ankle and joints of right foot
 ICD10-M79674 Pain in right toe(s)
 ICD10-M79675 Pain in left toe(s)
 ICD10-L84 Corns and callosities

ICD10-S93401A Sprain of unspecified ligament of right ankle, initial encounter
 765 ICD10-M2011 Hallux valgus (acquired), right foot
 ICD10-M79673 Pain in unspecified foot
 ICD10-S93402A Sprain of unspecified ligament of left ankle, initial encounter
 ICD10-M2012 Hallux valgus (acquired), left foot
 ICD10-M2041 Other hammer toe(s) (acquired), right foot
 770 ICD10-M19072 Primary osteoarthritis, left ankle and foot
 ICD10-M779 Enthesopathy, unspecified
 cluster #14 •
 ICD10-E11.9 Type 2 diabetes mellitus without complications
 ICD10-E78.5 Hyperlipidemia, unspecified
 775 ICD10-E03.9 Hypothyroidism, unspecified
 ICD10-Z79.899 Other long term (current) drug therapy
 ICD10-Z12.31 Encounter for screening mammogram for malignant neoplasm of breast
 ICD10-Z00.00 Encounter for general adult medical examination without abnormal findings
 780 ICD10-M54.5 Low back pain
 ICD10-K21.9 Gastro-esophageal reflux disease without esophagitis
 ICD10-R10.9 Unspecified abdominal pain
 ICD10-R07.9 Chest pain, unspecified
 785 ICD10-F41.9 Anxiety disorder, unspecified
 ICD10-J02.9 Acute pharyngitis, unspecified
 ICD10-N39.0 Urinary tract infection, site not specified
 ICD10-R53.83 Other fatigue
 ICD10-G47.33 Obstructive sleep apnea (adult) (pediatric)
 790 cluster #15 •
 ICD10-Z87891 Personal history of nicotine dependence
 ICD10-A419 Sepsis, unspecified organism
 ICD10-R918 Other nonspecific abnormal finding of lung field
 ICD10-E860 Dehydration
 795 ICD10-E871 Hypo-osmolality and hyponatremia
 ICD10-J90 Pleural effusion, not elsewhere classified
 ICD10-D72829 Elevated white blood cell count, unspecified
 ICD10-C20 Malignant neoplasm of rectum
 ICD10-G8918 Other acute postprocedural pain
 800 ICD10-N400 Benign prostatic hyperplasia without lower urinary tract symptoms
 ICD10-D696 Thrombocytopenia, unspecified
 ICD10-R1110 Vomiting, unspecified
 ICD10-R188 Other ascites
 ICD10-J9811 Atelectasis
 805 ICD10-C189 Malignant neoplasm of colon, unspecified
 cluster #16 •
 ICD10-M25532 Pain in left wrist
 ICD10-S0990XA Unspecified injury of head, initial encounter
 ICD10-W19XXXA Unspecified fall, initial encounter

810 ICD10-M25521 Pain in right elbow
ICD10-T148 Other injury of unspecified body region, initial encounter
ICD10-M79644 Pain in right finger(s)
ICD10-M79645 Pain in left finger(s)
ICD10-X58XXXA Exposure to other specified factors, initial encounter
815 ICD10-Z4789 Encounter for other orthopedic aftercare
ICD10-W010XXA Fall on same level from slipping, tripping and stumbling with-
out subsequent striking against object, initial encounter
ICD10-V892XXA Person injured in unspecified motor-vehicle accident, traffic,
initial encounter
820 ICD10-S83512A Sprain of anterior cruciate ligament of left knee, initial en-
counter
ICD10-S83511A Sprain of anterior cruciate ligament of right knee, initial en-
counter
ICD10-Z4802 Encounter for removal of sutures
825 ICD10-T1490 Injury, unspecified, initial encounter
cluster #17 ●
ICD10-Z4931 Encounter for adequacy testing for hemodialysis
ICD10-R4782 Fluency disorder in conditions classified elsewhere
ICD10-Z4901 Encounter for fitting and adjustment of extracorporeal dialysis
830 catheter
ICD10-L97513 Non-pressure chronic ulcer of other part of right foot with necro-
sis of muscle
ICD10-S83282D Other tear of lateral meniscus, current injury, left knee, subse-
quent encounter
835 ICD10-J338 Other polyp of sinus
ICD10-I0981 Rheumatic heart failure
ICD10-T451X5S Adverse effect of antineoplastic and immunosuppressive drugs,
sequela
ICD10-N3943 Post-void dribbling
840 ICD10-E0869 Diabetes mellitus due to underlying condition with other specified
complication
ICD10-O208 Other hemorrhage in early pregnancy
ICD10-Q383 Other congenital malformations of tongue
ICD10-N320 Bladder-neck obstruction
845 ICD10-H5052 Exophoria
ICD10-K6819 Other retroperitoneal abscess
cluster #18 ●
ICD10-Z00129 Encounter for routine child health examination without abnor-
mal findings
850 ICD10-Z00121 Encounter for routine child health examination with abnormal
findings
ICD10-J020 Streptococcal pharyngitis
ICD10-Z6852 Body mass index (BMI) pediatric, 5th percentile to less than 85th
percentile for age
855 ICD10-R0981 Nasal congestion

ICD10-F802 Mixed receptive-expressive language disorder
 ICD10-J3489 Other specified disorders of nose and nasal sinuses
 ICD10-Z3800 Single liveborn infant, delivered vaginally
 ICD10-J00 Acute nasopharyngitis [common cold]
 860 ICD10-H6123 Impacted cerumen, bilateral
 ICD10-B9789 Other viral agents as the cause of diseases classified elsewhere
 ICD10-J0390 Acute tonsillitis, unspecified
 ICD10-H903 Sensorineural hearing loss, bilateral
 ICD10-H6691 Otitis media, unspecified, right ear
 865 ICD10-H109 Unspecified conjunctivitis
 cluster #19 •
 ICD10-401.9
 ICD10-V70.0
 ICD10-272.4
 870 ICD10-250.00
 ICD10-724.2
 ICD10-V04.81
 ICD10-V72.31
 ICD10-401.1
 875 ICD10-786.50
 ICD10-780.79
 ICD10-V76.12
 ICD10-789.00
 ICD10-530.81
 880 ICD10-477.9
 ICD10-729.5