Illinois Department of Public Health
Division of Epidemiologic Studies
Illinois State Cancer Registry

User Manual – Chicago Community Areas
Cancer Incidence Public Dataset (30th edition), data as of November 2022, released April 2023

Chicago Community Areas
CA1120.dat

Data Use Agreement

By using these data, you signify your agreement to comply with the following statutorily based requirements.

The Illinois Health and Hazardous Substances Registry Act (410 ILCS 525/12) provides data collected by the Illinois State Cancer Registry (ISCR) be made available to the public. However, the identification or contact of individuals is prohibited.

In an effort to exclude identifying information on individual patients, the data (e.g. age, race, Hispanic ethnicity, year of diagnosis and type of cancer) have been aggregated into categories within individual records, the number of which depends on the size of the geographic area.

These data are provided as a public service for the purpose of statistical reporting and analysis only. There should be no attempt to learn the identity of any person included in these data. If the identity of any person is discovered inadvertently, no disclosure or other use of the identity will be made.

Uses of these data do not constitute an endorsement of the user's opinion or conclusions by the Department and none should be inferred.

Citation
Please reference the source of these data in any published document as follows; Illinois Department of Public Health, Illinois State Cancer Registry, public data set v30, 1986-2020, data as of November 2022.

Technical Notes

The Illinois public dataset is in ASCII format and contains sanitized individual records of cancer incidence among Illinois residents diagnosed from 1986 through 2019. Separate files are available for the state, counties, ZIP code areas, Cook County with Cook suburbs and Chicago, and Chicago community areas. Confidentiality is maintained by aggregating data within individual records into categories, the number of which depends on the size of the geographic area. Individual year of diagnosis is available for the Illinois data file. However, for the county and ZIP code files, the diagnosis year is a five-year aggregate (1986-90, 1991-95, 1996-00, 2001-05, 2006-10, 2011-2015 and 2016-20). For Cook County with Cook suburbs and Chicago, and Community Areas 2011-15 and 2016-20 are available. The following article describes the method used to measure uniqueness of the files.

Howe HL, Lake AJ, Shen T. Method to assess Identifiability in electronic data files. Am J epidemiol 2006; doi:10.1093/aje/kwk049.

The files include incidence data for invasive cancers only with the exception of cancer of the bladder. Carcinoma *insitu* of the breast is provided in a separate category. Non- melanoma skin cancers, cases reported with unknown or "other" sex, and cases with an unknown age are omitted.

The ASCII data files are NOT incorporated into a software program and therefore require the use of other statistical or database software packages for data analysis (spreadsheet programs such as Excel or QPRO may not work since the number of records in each file is in excess of two million).

Identification of cancer cases in ISCR is dependent upon reporting by hospitals, free- standing clinics, radiation treatment facilities, laboratories and physician offices as mandated by state law.

In addition, ISCR has agreements with other central cancer registries in Arkansas, California, Florida, Indiana, Iowa, Kentucky, Michigan, Minnesota (Mayo Clinic, through October 2005), Missouri, Mississippi (through 2004), North Carolina, Washington, Wisconsin, and Wyoming (through February 2008), to exchange cancer data. Completeness of out-of-state reporting depends upon the years of operation of these other central registries, the extent of their identification of out-of-state residents, and their standards of quality (see Cancer in North America: 2012-2016, Volume Two). https://www.naaccr.org/cancer-in-north-america-cina-volumes/

A death certificate clearance process involving follow-back of cancer deaths, in an effort to identify missed cases, has served as an additional means of case identification since August 1993.

Cancer cases are reported continuously to central registries in accordance with statutory reporting requirements. Cancer registries continue to revise and update data on the basis of new information. For this reason, an updated version of previous years data is published each year (e.g. data for diagnosis year 2010 reported with this version of the public data set may differ from the last version). Users of Illinois cancer data should be mindful of the "data as of" date.

The following completeness of case ascertainment for each year of the Illinois data and for each five-year group of the county, zip code and Cook County data was estimated using the standards developedby the North American Association of Central Cancer Registries (NAACCR) and adopted by the NationalProgram of Cancer Registries. This method uses a sex and site-specific incidence-to-mortality ratio for whites, in order to compare the number of reported cases to the number of expected cases. This method was used for data from 1986-1994 year of diagnosis. A detailed description of the original method is in the November 1996 issue of the NAACCR Newsletter and can be located on-line. https://narrative.naaccr.org/past-narrative-issues/

DX Year	Completeness	DXYrGrp	Completeness
1986	88%		
1987	90%		
1988	87%		
1989	88%		
1990	89%	1986-90	87%
1991	88%		
1992	91%		
1993	92%		
1994	97%		

For data from 1995-2018 the revised method was used. A detailed description of the revised method is in the Winter 2001 edition of the NAACCR Narrative and can be located on-line. https://narrative.naaccr.org/past-narrative-issues/

1995	100%	1991-95	94%
1996	100%		
1997	100%		
1998	100%		
1999	100%		
2000	100%	1996-00	100%
2001	100%		
2002	100%		
2003	100%		
2004	100%		
2005	100%	2001-05	100%
2006	100%		
2007	100%		
2008	100%		
2009	100%		
2010	100%	2006-10	100%
2011	100%		
2012	100%		
2013	100%		
2014	100%		
2015	100%	2011-15	100%
2016	100%		
2017	100%		
2018	100%		
2019	100%		
2020	100%	2016-20	100%

In addition to the estimates for completeness of case ascertainment (as of 11/22), NAACCR has developed a certification process that reviews registry data for completeness, accuracy, and timeliness of reporting. The criteria for silver and goldcertification can be found on the NAACCR web site. https://www.naaccr.org/certification-criteria/

As of November 2022, ISCR data met the criteria for gold certification for diagnosis years 1995-2020. Certification status for 2020 data will be awarded in June 2023. [NOTE: diagnosis year 1995 now meets criteria for gold certification.]

	Criteria			Missing Data Fields			
Year	Pass Edits (%)	DCO*(%)	Unresolved Duplicates** (%)	Sex (%)	Age (%)	County (%)	Race*** (%)
1986	~	~	~	0.0	0.0	0.0	0.2
1987	~	~	~	0.0	0.0	0.0	0.2
1988	~	~	0.04	0.0	0.0	0.0	0.2
1989	~	~	0.04	0.0	0.0	0.0	0.2
1990	100	~	0.04	0.0	0.0	0.0	0.2
1991	100	~	0.04	0.0	0.0	0.0	0.5
1992	100	~	0.04	0.0	0.0	0.0	0.3
1993	100	2.2	0.04	0.0	0.0	0.0	0.2
1994	100	6.1	0.06	0.0	0.0	0.0	0.2
1995	100	2.6	0.03	0.0	0.0	0.0	0.3
1996	100	1.8	0.02	0.0	0.0	0.0	0.3
1997	100	1.8	0.09	0.0	0.0	0.0	0.5
1998	100	1.5	0.03	0.0	0.0	0.0	0.7
1999	100	1.8	0.02	0.0	0.0	0.0	0.7
2000	100	2.3	0.03	0.0	0.0	0.0	0.7
2001	100	2.4	0.00	0.0	0.0	0.0	0.6
2002	100	2.6	0.00	0.0	0.0	0.0	0.8
2003	100	1.5	0.02	0.0	0.0	0.0	0.8
2004	100	1.7	0.01	0.0	0.0	0.0	0.8
2005	100	1.9	0.00	0.0	0.0	0.0	1.0
2006	100	2.0	0.00	0.0	0.0	0.0	0.7
2007	100	1.2	0.00	0.0	0.0	0.0	0.7
2008	100	1.6	0.07	0.0	0.0	0.0	0.9
2009	100	1.6	0.03	0.0	0.0	0.0	1.0
2010	100	1.8	0.03	0.0	0.0	0.0	1.0
2011	100	1.8	0.00	0.0	0.0	0.0	1.2
2012	100	0.9	0.02	0.0	0.0	0.0	1.0
2013	100	1.0	0.02	0.0	0.0	0.0	0.9
2014	100	2.1	0.02	0.0	0.0	0.0	1.0
2015	100	2.1	0.02	0.0	0.0	0.0	1.0
2016	100	1.9	0.09	0.0	0.0	0.0	0.8
2017	100	1.9	0.06	0.0	0.0	0.0	0.8
2018	100	1.4	0.07	0.0	0.0	0.0	1.4
2019	100	0.9	0.02	0.0	0.0	0.0	1.5
2020	100	1.2	0.02	0.0	0.0	0.0	2.0

~not applicable

*DCO follow-back not started until end of 1993 reporting year

NAACCR's duplicate protocol was run for each year at the time of data submission for registry certification*race 98 "other unspecified" for diagnosis year 1991+ is included with unknown to be compatible with national reporting

Cancer coding changes during 2001-2020

Cancer Incidence - several definitional changes occurred in some histology and behavior codes in ICD-O-3 that affected the inclusion and exclusion of reportable cancers diagnosed beginning in 2001.The changes predominately affected leukemias, lymphomas, and cancer of the ovary. One category of change between ICD-O-2 and ICD-O-3 is the manner in which leukemias and lymphomas are classified and coded. Although conversion of histology codes from ICD-O-2 to ICD-O-3 for cases diagnosed prior to 2001 will help to minimize these differences, some minor differences may still exist, particularly with respect to some relatively rare lymphocytic cancers that can be coded to either leukemia or lymphoma.

Starting with ICD-O-3, several myelodysplastic diseases and syndromes are considered malignant, and therefore are now reportable for cases diagnosed in 2001 and later and are included in these data. Leukemias that represent a disease progression from one of the myelodysplastic diseases or syndromes diagnosed in 2001 and forward are no longer reportable.

For pediatric cancers, differences in incidence rates may be due to changes between the second and third edition of the International Classification of Childhood Cancers (ICCC). Two changes in the ICCC-3 classification are main contributors to this change. 1) Burkitt lymphoma and unspecified lymphoma, which were separated from non-Hodgkin lymphoma previously, are combined with non-Hodgkin lymphoma; 2) Some lymphomas, which were grouped in the miscellaneous lymphoreticular neoplasms previously, are now included in the non-Hodgkin lymphoma category.

Pilocytic astrocytoma is considered to have uncertain behavior in the published version of ICD-O-3, but is reportable as a malignant cancer in North America. Including the childhood astrocytomas in the category of malignant brain tumors may introduce differences between childhood brain cancer rates in North America compared to other areas of the world that may not include these tumors as malignant.

In addition, mesothelioma and Kaposi sarcoma cases are reported as separate categories. This change has little or no impact on most rates for specific cancers.

Coding changes for 2010 cases:

SEER Site recode ICD-O-3/WHO 2008 is used for cases diagnosed 2010 and based on ICD-O-3, updated for Hematopoietic codes based on WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008). http://seer.cancer.gov/tools/heme

If you have questions about the dataset, please contact:

Lori Koch
Illinois Department of Public Health
Division of Epidemiologic Studies
535 West Jefferson Street, 3rd Fl
Springfield, IL 62761 phone (217) 785-1873
Email lori.koch@illinois.gov

File layout for Chicago Community Areas CA1120.DAT (records – 120,880)
Record Format (all fields are numeric)

Data Field	Positions	Length
Chicago Community Areas	1-2	2
Diagnosis year group	3-3	1
Cancer site group	4-5	2
Age at diagnosis group code	6-6	1
Sex code	7-7	1
Race group code	8-8	1
Stage of disease at diagnosis	9-9	1
Hispanic code	10-10	1
Geocode precision	11-11	1

Codes for data fields
Chicago Community Areas
(See end of the file for census tract groups for community areas.
Modifications made per request of CDPH, based on 2010 census codes.)

1 Rogers Park	26 West Garfield Park	51 South Deering
2 West Ridge	27 East Garfield Park	52 East Side
3 Uptown	28 Near West Side	53 West Pullman
4 Lincoln Square	29 North Lawndale	54 Riverdale
5 North Center	30 South Lawndale	55 Hegewisch
6 Lake View	31 Lower West Side	56 Garfield Ridge
7 Lincoln Park	32 Loop	57 Archer Heights
8 Near North Side	33 Near South Side	58 Brighton Park
9 Edison Park	34 Armour Square	59 McKinley Park
10 Norwood Park	35 Douglas	60 Bridgeport
11 Jefferson Park	36 Oakland	61 New City
12 Forest Glen	37 Fuller Park	62 West Elsdon
13 North Park	38 Grand Boulevard	63 Gage Park
14 Albany Park	39 Kenwood	64 Clearing
15 Portage Park	40 Washington Park	65 West Lawn
16 Irving Park	41 Hyde Park	66 Chicago Lawn
17 Dunning	42 Woodlawn	67 West Englewood
18 Montclare	43 South Shore	68 Englewood
19 Belmont Cragin	44 Chatham	69 Greater Grand Crossing
20 Hermosa	45 Avalon Park	70 Ashburn
21 Avondale	46 South Chicago	71 Auburn Gresham
22 Logan Square	47 Burnside	72 Beverly
23 Humboldt Park	48 Calumet Heights	73 Washington Heights
24 West Town	49 Roseland	74 Mount Greenwood
25 Austin	50 Pullman	75 Morgan Park
		76 O'Hare
		77 Edgewater

Diagnosis year group code

1   2011-2015
2   2016-2020

Cancer site group (based on SEER group codes*)

1   colorectal (21041-21052)
2   lung and bronchus (22030)
3   female breast-invasive (26000, behavior code 3)
4   cervix (27010)
5   prostate (28010)
6   all other cancers (all other codes)
7   female breast-insitu (26000, behavior code 2)

* SEER group codes
The recode, primary site code, and morphology codes used to categorize cancer incidence into
the following SEER groups can be found on the SEER website
http://seer.cancer.gov/siterecode/

Age at diagnosis group code

1   <5
2   5-14
3   15-34
4   35-44
5   45-54
6   55-64
7   65+

Sex code
1   male
2   female

Race group code

1   white
2   black
3   other*
9   unknown (includes unknown race and race code 98 'other unspecified' to be
        compatible with national reporting of race)

* Other race includes Asian-American, Pacific Islanders, American Indians, Alaska Natives, and all
other races. In order to improve the quality of cancer surveillance data on American Indians the
race has been enhanced with a linkage to the Indian Health Services patient registration database.
This linkage identifies cancer cases among American Indians who were misclassified as non-Indians
in the registry database.

Stage of disease at diagnosis

0   in situ
1   localized
2   regional
3   distant metastases/systemic disease
9   unknown, unstaged, or unspecified

Hispanic code

1   Hispanic *
0   Non-Hispanic (includes unknown)

*Hispanic ethnicity was enhanced according to the NAACCR Hispanic identification algorithm (NHIA). NHIA is a generally reliable method to enhance the ethnic identification of the Latino population in the United States.

Howe HL. Evaluation of NHIA Submission for 1997-2001. Springfield, Ill.: North American Association of Central Cancer Registries, October 2004.

Schenke N and Gentleman JF On judging the significance of differences by examining the overlap between confidence intervals. The American Statistician 2001;55:182-18

Geocode Precision (used to group community areas level of precision)

1   address level (98.9% of cases – see map)
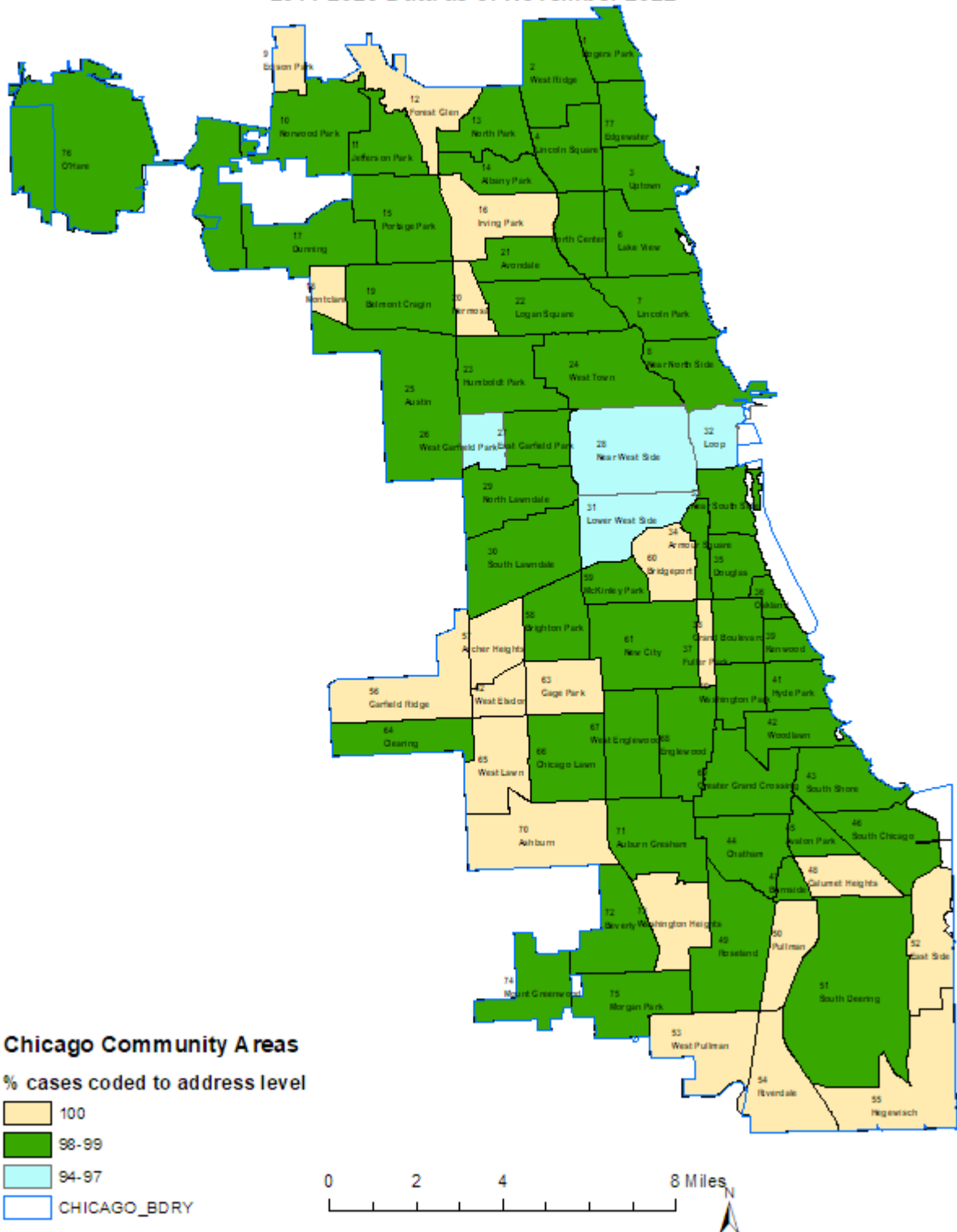2   ZIP+4
3   ZIP+2
4   ZIP

Community Area	4 digit tract	added new 4 digit tract from 2010 census	partial tracts/blocks included - NOTES
			taken out per request Eric Jones, CDPH (5/2012) outside city limits
			2010 census some new tracts split btwn CAs or partially added
1 Rogers Park	101 thru 109	8306	
2 West Ridge	201 thru 209		
3 Uptown	310 thru 321	8307	
4 Lincoln Square	401 thru 410	8308	
5 North Center	501 thru 515	8437	
6 Lake View	601 thru 634	8319 8320 8321	
7 Lincoln Park	701 thru 720	8325 8326	
8 Near North Side	801 thru 819	8422 8383	
9 Edison Park	901 thru 903		
10 Norwood Park	1001 thru 1007	(8104 blocks 3029 3032 3035 3039 3041 3048 3051 3018 3025 3026 3071 3034 3027 3028)	most of area 8104 8106 out of city limits
11 Jefferson Park	1101 thru 1105		
12 Forest Glen	1201 thru 1204		
13 North Park	1301 thru 1305	8318	
14 Albany Park	1401 thru 1408		
15 Portage Park	1501 thru 1512	8317	
16 Irving Park	1601 thru 1613		
17 Dunning	1701 thru 1711		
18 Montclare	1801 thru 1803	8316	
19 Belmont Cragin	1901 thru 1914	8315	
20 Hermosa	2001 thru 2006	8312	some 8312 in CA 22 BUT no population
21 Avondale	2101 thru 2109	8311 (8309 blocks 1000 thru 1005)	
22 Logan Square	2201 thru 2229	8322 8323 8324 8309 8310	some 8309 in CA21 some 8310 in CA24
23 Humboldt Park	2301 thru 2318	8421 8366 8367	
24 West Town	2401 thru 2436	8423 (8310 blocks 1000 thru 1017)	
25 Austin	2501 thru 2524	8313 8314	
26 West Garfield Park	2601 thru 2610		
27 East Garfield Park	2701 thru 2719	8368 thru 8371 8373 thru 8374	(part of 8371 block 1006 in CA28 but no population)
28 Near West Side	2801 thru 2843	8330 8378 8380 8381 8331 8333 8382 8329 8419 8429 (8371 block 1014)	(all of 8371 block 1014 in CA28)
29 North Lawndale	2901 thru 2927	8430 8387 8386 8434 8433 8431 8415 8416 8414	
30 South Lawndale	3001 thru 3020	8417 8407 8408 8435 8305	
31 Lower West Side	3101 thru 3115	8412 8413 8432	
32 Loop	3201 thru 3206	8391 8390	
33 Near South Side	3301 thru 3305	8410	(8411 block 3039 no population CA34)
34 Armour Square	3401 thru 3406	8411	(CA 60 tract 8400 block 2015 in CA34 but no population)
35 Douglas	3501 thru 3515	8392 8420 8395 8396	
36 Oakland	3601 thru 3605	8365 8364	
37 Fuller Park	3701 thru 3704	8355 8356	
38 Grand Boulevard	3801 thru 3820	8359 8358 8357 8360 8436	
39 Kenwood	3901 thru 3907		
40 Washington Park	4001 thru 4008	8361 8345	
41 Hyde Park	4101 thru 4114	8362 8363	
42 Woodlawn	4201 thru 4212	8344 8439	some 8439 in CA 43
43 South Shore	4301 thru 4314	8342 (8439 blocks 3000-3002, 4000-4002, 5000-5004)	
44 Chatham	4401 thru 4409	8424	
45 Avalon Park	4501 thru 4503	8343	
46 South Chicago	4601 thru 4610	8339	
47 Burnside	4701		
48 Calumet Heights	4801 thru 4805		
49 Roseland	4901 thru 4914	8340	
50 Pullman	5001 thru 5003		
51 South Deering	5101 thru 5105	8388	
52 East Side	5201 thru 5206		
53 West Pullman	5301 thru 5306		
54 Riverdale	5401		(8215 block 1094 outside city limit)
55 Hegewisch	5501 thru 5502		
56 Garfield Ridge	5601 thru 5613	8352	(part of 9801 blocks 1000-1010 no population)
57 Archer Heights	5701 thru 5705		
58 Brighton Park	5801 thru 5811	8428	
59 McKinley Park	5901 thru 5907	8403 8404	(tract 5907 block 2009 in CA60 but no population)
60 Bridgeport	6001 thru 6016	8402 8401 8398 8400 8397 8399	(some of 8400 in CA34 no population)
61 New City	6101 thru 6122	8426 8438	

62 West Elsdon	6201 thru 6204		
63 Gage Park	6301 thru 6309	8351	
64 Clearing	6401 thru 6408		**(part of 9801 block 1005 no population)**
65 West Lawn	6501 thru 6505		
66 Chicago Lawn	6601 thru 6611	8350	
67 West Englewood	6701 thru 6720	8349	
68 Englewood	6801 thru 6814	8347 8348 8346	
69 Greater Grand Crossing	6901 thru 6915	8418 8425	
70 Ashburn	7001 thru 7005		8209 out of city limits was never in CA
71 Auburn Gresham	7101 thru 7115		
72 Beverly	7201 thru 7207		
73 Washington Heights	7301 thru 7307		
74 Mount Greenwood	7401 thru 7404		
75 Morgan Park	7501 thru 7506		8233 out of city limits except block 1001 - no population
76 O'Hare (cont)	7608 thru 7609	**9800 (blocks 1039-1045 & 1059 outside city limits** **7707 blk 1065 has pop**	(7609 blks 1081-1084, 1096 outside city limits no population) 7705-7709 blocks outside city except 7707 block 1065 8400 8408 in Dupage County 8116 8117 outside city limits
77 Edgewater	301 thru 309		

Illinois Department of Public Health, Division of Epidemiologic Studies
Illinois State Cancer Registry, Community Areas of Chicago
Percent of case reports with an address level geocode
2011-2020 Data as of November 2022

Chicago Community Areas

% cases coded to address level

- 100
- 98-99
- 94-97
- CHICAGO_BDRY

0   2   4   8 Miles