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

User Manual

Cancer Incidence Public Dataset (32nd edition), data as of November 2024, release June 2025
Illinois Data – **IL8622.dat**
County Data – **CNTY8822.dat**
ZIP code data – **ZPCD8822.dat**
Cook County with Cook suburbs and Chicago – **COOK1322.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 any endorsement of the users' opinion or conclusions by the Department, and noneshould 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 V32, 1986-2022, data as of November 2024.

Technical Notes

The Illinois public data set is in fixed ASCII format and contains sanitized individual records of cancer incidence among Illinois residents diagnosed from 1986 through 2022. Separate files are available for the state, counties, ZIP codes, and Cook County with Cook suburbs and Chicago. 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, in the county and ZIP code files, the diagnosis year is a five-year aggregate (1988-92, 1993-97, 1998-02, 2003-07, 2008-12, 2013-17 and 2018-22) and five-year aggregate for Cook County with Cook suburbs and Chicago is (2013-17 and 2018-22). 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 Illinois incidence data for invasive cancers only, except for cancers of the bladder and carcinoma *in-situ* of the breast is provided in a separate category. Non-melanoma skin cancers, cases reported withunknown 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 Quattro Pro may not work since the number of cases in some of these files is too large).

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 (May Clinic through 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 cancer registries, the extent of their identification of our-of-state residents, and their standards of quality (see Cancer in North American: 2011-2015, 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, based on new information. For this reason, an updated version of previous years data is published each year (e.g. data for diagnosis year 2021 reported with this version of the public data set may differ from the last version). Users of the 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%		
1991	88%		
1992	91%	1988-92	89%
1993	92%		
1994	97%		

For data from 1995-2022 the revised method was used. A detailed description of the revised method isin the Winter 2001 edition of the NAACCR Narrative and can be located on-line.
https://narrative.naaccr.org/past-narrative-issues/

1995	100%		
1996	100%		
1997	100%	1993-97	98%
1998	100%		
1999	100%		
2000	100%		
2001	100%		
2002	100%	1998-02	100%
2003	100%		
2004	100%		
2005	100%		
2006	100%		
2007	100%	2003-07	100%
2008	100%		
2009	100%		
2010	100%		
2011	100%		
2012	100%	2008-12	100%
2013	100%		
2014	100%		
2015	100%		
2016	100%		
2017	100%	2013-17	100%
2018	100%		
2019	100%		
2020	100%		
2021	100%		
2022	100%	2018-22	100%

In addition to the estimates for completeness of case ascertainment (as of 11/24), NAACCR has developed a certification process that reviews registry data for completeness, accuracy, and timeliness of reporting. The criteria for silver and gold certification can be found on the NAACCR web site. https://www.naaccr.org/certification-criteria/

As of November 2024, ISCR data met the criteria for gold certification for diagnosis years 1995-2022. Certification status for 2022 data will be awarded in June 2025. [NOTE: diagnosis year 1995 now meets criteria for gold certification.]

Year	Completeness (NAACCR Method)[a] (% As of 11-24)	Pass EDITS (%)	DCO[b] (%)	Unresolved Duplicate[c] (%)	Missing Data Fields			
					Sex (%)	Age (%)	County (%)	Race (%)
1986	88	~	~	~	0.0	0.0	0.0	0.2
1987	90	~	~	~	0.0	0.0	0.0	0.1
1988	87	~	~	0.04	0.0	0.0	0.0	0.2
1989	88	~	~	0.04	0.0	0.0	0.0	0.2
1990	89	100	~	0.04	0.0	0.0	0.0	0.2
1991	88	100	~	0.04	0.0	0.0	0.0	0.5
1992	91	100	~	0.04	0.0	0.0	0.0	0.3
1993	92	100	2.2	0.04	0.0	0.0	0.0	0.2
1994	97	100	6.1	0.06	0.0	0.0	0.0	0.2
1995	99	100	2.6	0.03	0.0	0.0	0.0	0.3
1996	100	100	1.8	0.02	0.0	0.0	0.0	0.3
1997	100	100	1.8	0.09	0.0	0.0	0.0	0.4
1998	100	100	1.5	0.03	0.0	0.0	0.0	0.6
1999	100	100	1.8	0.02	0.0	0.0	0.0	0.6
2000	100	100	2.3	0.03	0.0	0.0	0.0	0.6
2001	100	100	2.4	0.00	0.0	0.0	0.0	0.6
2002	100	100	2.6	0.00	0.0	0.0	0.0	0.7
2003	100	100	1.5	0.02	0.0	0.0	0.0	0.8
2004	100	100	1.7	0.01	0.0	0.0	0.0	0.7
2005	100	100	1.9	0.00	0.0	0.0	0.0	0.9
2006	100	100	2.0	0.00	0.0	0.0	0.0	0.7
2007	100	100	1.2	0.00	0.0	0.0	0.0	0.7
2008	100	100	1.6	0.07	0.0	0.0	0.0	0.8
2009	100	100	1.6	0.03	0.0	0.0	0.0	0.9
2010	100	100	1.8	0.03	0.0	0.0	0.0	0.9
2011	100	100	1.8	0.00	0.0	0.0	0.0	1.1
2012	100	100	0.8	0.02	0.0	0.0	0.0	1.0
2013	100	100	1.0	0.02	0.0	0.0	0.0	0.8
2014	100	100	2.1	0.02	0.0	0.0	0.0	0.9
2015	100	100	2.1	0.09	0.0	0.0	0.0	0.9
2016	100	100	1.9	0.09	0.0	0.0	0.0	0.8
2017	100	100	1.9	0.06	0.0	0.0	0.0	0.7
2018	100	100	1.3	0.07	0.0	0.0	0.0	1.2
2019	100	100	0.9	0.02	0.0	0.0	0.0	1.3
2020	100	100	1.2	0.02	0.0	0.0	0.0	1.7
2021	100	100	1.2	0.00	0.0	0.0	0.0	1.8
2022	100	100	0.9	0.00	0.0	0.0	0.0	2.2

~ not applicable

a. For data prior to 1995, the NAACCR's completeness estimating algorithm (version 1) was used. For data on or after 1995, the NAACCR's completeness estimating algorithm (version 2) was used.

b. DCO follow back not started until end of 1993 reporting year.

c. NAACCR's duplicate protocol was run for each year at the time of data submission for registry certification.

Cancer coding changes during 2001-2022

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 astrocytoma 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's 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/WHO2008 – 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 W Jefferson St, 3rd Fl
Springfield, IL 62761
Phone 217-785-1873
Email – dph.ISCRinquiries@illinois.gov

File Layout for State file IL8622.dat (number of records 2,329,491)

Record Format (all fields are numeric).

Data Field	Positions	Length
sex code	1-1	1
diagnosis year	2-5	4
report source	6-6	1
stage of disease at diagnosis	7-7	1
histology (4-digit code)	8-11	4
behavior code	12-12	1
SEER groups	13-17	5
pediatric codes level 1	18-19	2
pediatric codes level 2	20-22	3
5-yr age at diagnosis group	23-24	2
race group code	25-25	1
Hispanic code	26-26	1
method of diagnosis code	27-27	1

Codes for Data

fieldsSex code

1 male
2 female

Diagnosis Year

1986	2001	2016
1987	2002	2017
1988	2003	2018
1989	2004	2019
1990	2005	2020
1991	2006	2021
1992	2007	2022
1993	2008	
1994	2009	
1995	2010	
1996	2011	
1997	2012	
1998	2013	
1999	2014	
2000	2015	

Report Source

1 hospital or clinic
2 radiation treatment centers/medical oncology centers
3 laboratory only
4 physician's office
5 nursing home or hospice
6 autopsy only
7 death certificate only
8 other hospital out-patient units/surgery

Stage of disease at diagnosis

0 in situ
1 localized
2 regional
3 distant
9 unknown, unstaged, or unspecified

Histology (4-digit code)

NOTE: Although cases diagnosed in 1986 and 2000 were reported with The International Classification of Diseases for Oncology Version 2 (ICD-O-2) codes 2, and cases diagnosed between 2001-2022, were reported with version 3 (ICD-O-3) codes3, all cases were converted to version 3 codes and grouped according to site group definitions established by the SEER program of the National Cancer Institute (NCI) and also used by NAACCR. These standardized classification schemes allow direct comparisons of Illinois data with international, national and state publications. Kaposi's sarcoma and mesothelioma are classified as separate site groups. This change has a little or no impact on cancer incidence rates for a few specific cancers, compared to using the previous site grouping method.

ICD-O-3 for specific codes see:
Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, Whelan S. International Classification of Diseases for Oncology, 3rd ed Geneva: World Health Organization; 2000.

Behavior code (5th digit of morphology code)

2 in-situ
3 malignant

SEER Groups

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

20010 lip
20020 tongue
20030 salivary gland
20040 floor of mouth
20050 gum & other mouth
20060 nasopharynx
20070 tonsil

20080 oropharynx
20090 hypopharynx
20100 other oral cavity & pharynx
21010 esophagus
21020 stomach
21030 small intestine
21041 cecum
21042 appendix
21043 ascending colon
21044 hepatic flexure
21045 transverse colon
21046 splenic flexure
21047 descending colon
21048 sigmoid colon
21049 large intestine NOS
21051 rectosigmoid junction
21052 rectum
21060 anus, anal canal, anorectum
21071 liver
21072 intrahepatic bile duct
21080 gallbladder
21090 other biliary
21100 pancreas
21110 retroperitoneum
21120 peritoneum, omentum, mesentery
21130 other digestive organs
22010 nose, nasal cavity, middle ear
22020 larynx
22030 lung & bronchus
22050 pleura
22060 trachea, mediastinum, other respiratory organs
23000 bones & joints
24000 soft tissue including heart
25010 melanoma of skin
25020 other non-epithelial skin
26000 breast
27010 cervix uteri
27020 corpus uteri
27030 uterus NOS
27040 ovary
27050 vagina
27060 vulva
27070 other female genital organs
28010 prostate
28020 testis
28030 penis
28040 other male genital organs
29010 urinary bladder
29020 kidney & renal pelvis
29030 ureter
29040 other urinary organs

30000 eye & orbit
31010 brain
31040 cranial nerves other nervous system
32010 thyroid
32020 other endocrine including thymus
33011 hodgkin lymphoma - nodal
33012 hodgkin lymphoma - extranodal
33041 non-hodgkin lymphoma - nodal
33042 non-hodgkin lymphoma - extranodal
34000 myeloma
35011 acute lymphocytic leukemia
35012 chronic lymphocytic leukemia
35013 other lymphocytic leukemia
35021 acute myeloid leukemia
35022 chronic myeloid leukemia
35023 other myeloid/monocytic leukemia
35031 acute monocytic leukemia
35041 other acute leukemia
35043 aleukemic, subleukemic & NOS
36010 mesothelioma
36020 kaposi sarcoma
37000 miscellaneous

Pediatric codes level 1 (ICCC ICD-0-3)
NOTE:The following article describes the current method used to group pediatric cancer: Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International classification of Childhood Cancer, Third Edition. Cancer. 2005;103:1457-67

1 I Leukemias, myeloproliferative diseases, and mylodysplastic diseases
2 II Lymphomas and reticuloendothelial neoplasms
3 III Central nervous system and miscellaneous intracranial & intraspinal neoplasms
4 IV Neuroblastoma and other peripheral nervous cell tumors
5 V Retinoblastoma
6 VI Renal tumors
7 VII Hepatic tumors
8 VIII Malignant bone tumors
9 IX Soft tissue and other extraosseous sarcomas
10 X Germ cell tumors, trophoblastic tumors, and neoplasms of gonads
11 XI Other malignant epithelial neoplasms and malignant melanomas
12 XII Other and unspecified malignant neoplasms
98 pediatric case (aged 0-19) not classified by ICCC3
99 not pediatric case (age > 19)

Pediatric codes level 2 (ICCC ICD-O-3)
11 I(a) Lymphoid leukemias
12 I(b) Acute myeloid leukemias
13 I(c) Chronic myeloproliferative diseases
14 I(d) Myelodysplastic syndrome and other myeloproliferative diseases
15 I(e) Unspecified and other specified leukemias
21 II(a) Hodgkin lymphomas
22 II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)

23 II(c) Burkitt lymphoma
24 II(d) Miscellaneous lymphoreticular neoplasms
25 II(e) Unspecified lymphomas
31 III(a) Ependymomas and choroid plexus tumor
32 III(b) Astrocytomas
33 III(c) Intracranial and intraspinal embryonal tumors
34 III(d) Other gliomas
35 III(e) Other specified intracranial and intraspinal neoplasms
36 III(f) Unspecified intracranial and intraspinal neoplasms
41 IV(a) Neuroblastoma and ganglioneuroblastoma
42 IV(b) Other peripheral nervous cell tumors
50 V Retinoblastoma
61 VI(a) Nephroblastoma and other nonepithelial renal tumors
62 VI(b) Renal carcinomas
63 VI(c) Unspecified malignant renal tumors
71 VII(a) Hepatoblastoma
72 VII(b) Hepatic carcinomas
73 VII(c) Unspecified malignant hepatic tumors
81 VIII(a) Osteosarcomas
82 VIII(b) Chondrosarcomas
83 VIII(c) Ewing tumor and related sarcomas of bone
84 VIII(d) Other specified malignant bone tumors
85 VIII(e) Unspecified malignant bone tumors
91 IX(a) Rhabdomyosarcomas
92 IX(b) Fibrosarcomas, peripheral nerve sheath tumors & other fibrous neoplasms
93 IX(c) Kaposi sarcoma
94 IX(d) Other specified soft tissue sarcomas
95 IX(e) Unspecified soft tissue sarcomas
101 X(a) Intracranial and intraspinal germ cell tumors
102 X(b) Malignant extracranial and extragonadal germ cell tumors
103 X(c) Malignant gonadal germ cell tumors
104 X(d) Gonadal carcinomas
105 X(e) Other and unspecified malignant gonadal tumors
111 XI(a) Adrenocortical carcinomas
112 XI(b) Thyroid carcinomas
113 XI(c) Nasopharyngeal carcinomas
114 XI(d) Malignant melanomas
115 XI(e) Skin carcinomas
116 XI(f) Other and unspecified carcinomas
121 XII(a) Other specified malignant tumors
122 XII(b) Other unspecified malignant tumors
998 pediatric case (aged 0-19) not classified in ICCC3
999 not pediatric case (age > 19)

5-yr age at diagnosis groups

0	<1
1	1-4
2	5-9
3	10-14
4	15-19
5	20-24

6 25-29
7 30-34
8 35-39
9 40-44
10 45-49
11 50-54
12 55-59
13 60-64
14 65-69
15 70-74
16 75-79
17 80-84
18 85+

Race group code

1 white
2 black
3 other*
9 unknown (includes race code 98 "other unspecified" for diagnosis years 1991+ 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.

Hispanic code

0 non-Hispanic (including unknown)
1 Hispanic*
9 Hispanic data not available for diagnosis years 1986-1989

*Hispanic ethnicity was enhanced according to the NAACCR Hispanic Identification algorithm (NHIA). NHIA is a generally reliable method to enhance the ethnic identification of 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;55:186-186.

Method of Diagnosis Code

1 positive histology
2 positive exfoliative cytology, no positive histology
3 positive histology & positive immunophenotyping and/or positive genetic studies
4 positive microscopic confirmation, method not specified
5 positive laboratory test or marker study
6 direct visualization without microscopic confirmation
7 radiography & other imaging techniques without microscopic confirmation
8 clinical diagnosis only (other than 5,6 or 7)
9 unknown whether microscopically confirmed

File Layout for county file CNTY8822.dat (number of records 2,240,077)

Record Format (all fields are numeric)

Data Field	Positions	Length
sex code	1-1	1
diagnosis year group	2-2	1
county at diagnosis code	3-5	3
report source	6-6	1
stage of disease at diagnosis	7-7	1
cancer site group codes	8-9	2
age at diagnosis group code	10-10	1
race group code	11-11	1
Hispanic code	12-12	1

Codes for data fields

Sex code

1 male
2 female

Diagnosis Year group (5yr groups)

1 1988-1992
2 1993-1997
3 1998-2002
4 2003-2007
5 2008-2012
6 2013-2017
7 2018-2022

County at diagnosis code with race and Hispanic availability

1	Adams	all race only
3	Alexander	all race only
5	Bond	all race only
7	Boone	all race only
9	Brown	all race only
11	Bureau	all race only
13	Calhoun	all race only
15	Carroll	all race only
17	Cass	all race only
19	Champaign	all race, white, black
21	Christian	all race only
23	Clark	all race only
25	Clay	all race only
27	Clinton	all race only
29	Coles	all race only

31	Cook	all race, white, black, other and Hispanic/non-Hispanic
33	Crawford	all race only
35	Cumberland	all race only
37	DeKalb	all race only
39	DeWitt	all race only
41	Douglas	all race only
43	DuPage	all race, white, black, other and Hispanic/non-Hispanic
45	Edgar	all race only
47	Edwards	all race only
49	Effingham	all race only
51	Fayette	all race only
53	Ford	all race only
55	Franklin	all race only
57	Fulton	all race only
59	Gallatin	all race only
61	Greene	all race only
63	Grundy	all race only
65	Hamilton	all race only
67	Hancock	all race only
69	Hardin	all race only
71	Henderson	all race only
73	Henry	all race only
75	Iroquois	all race only
77	Jackson	all race only
79	Jasper	all race only
81	Jefferson	all race only
83	Jersey	all race only
85	JoDaviess	all race only
87	Johnson	all race only
89	Kane	all race, white, black and Hispanic/non-Hispanic
91	Kankakee	all race, white, black
93	Kendall	all race only
95	Knox	all race only
97	Lake	all race, white, black, other and Hispanic/non-Hispanic
99	LaSalle	all race only
101	Lawrence	all race only
103	Lee	all race only
105	Livingston	all race only
107	Logan	all race only
109	McDonough	all race only
111	McHenry	all race only
113	McLean	all race only
115	Macon	all race, white, black
117	Macoupin	all race only
119	Madison	all race, white, black
121	Marion	all race only
123	Marshall	all race only
125	Mason	all race only

127	Massac	all race only
129	Menard	all race only
131	Mercer	all race only
133	Monroe	all race only
135	Montgomery	all race only
137	Morgan	all race only
139	Moultrie	all race only
141	Ogle	all race only
143	Peoria	all race, white, black
145	Perry	all race only
147	Piatt	all race only
149	Pike	all race only
151	Pope	all race only
153	Pulaski	all race only
155	Putnam	all race only
157	Randolph	all race only
159	Richland	all race only
161	Rock Island	all race, white, black
163	St Clair	all race, white, black
165	Saline	all race only
167	Sangamon	all race, white, black
169	Schuyler	all race only
171	Scott	all race only
173	Shelby	all race only
175	Stark	all race only
177	Stephenson	all race only
179	Tazewell	all race only
181	Union	all race only
183	Vermilion	all race, white, black
185	Wabash	all race only
187	Warren	all race only
189	Washington	all race only
191	Wayne	all race only
193	White	all race only
195	Whiteside	all race only
197	Will	all race, white, black and Hispanic/non-Hispanic
199	Williamson	all race only
201	Winnebago	all race, white, black
203	Woodford	all race only

Report Source

1 hospital or clinic
2 radiation treatment centers/medical oncology centers
3 laboratory only
4 physician's office
5 nursing home or hospice
6 autopsy only
7 death certificate only
8 other hospital out-patient units/surgery

Stage of disease at diagnosis

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

Cancer site group (based on SEER group codes)

1 oral cavity & pharynx (20010-20100)
2 esophagus (21010)
3 stomach (21020)
4 colorectal (21041-21052)
5 liver (21071)
6 pancreas (21100)
7 lung & bronchus (22030)
8 bone (23000)
9 melanomas (25010)
10 breast-invasive only (26000, behavior code 3)
11 cervix (27010)
12 uterus (27020-27030)
13 ovary (27040)
14 prostate (28010)
15 testis (28020)
16 bladder (29010)
17 kidney (29020)
18 nervous system (31010-31040)
19 Hodgkin's lymphomas (33011-33012)
20 non-Hodgkin's lymphomas (33041-33042)
21 myelomas (34000)
22 leukemias (35011-35043)
23 all other sites
(21030,21060,21072,21080,21090,21110,21120,21130,22010,22020,22050,22060,
24000,25020,27050,27060,27070,28030,8040,29030,29040,30000,32010,32020,
36010,36020,37000)
24 breast-in-situ only (26000, behavior code 2)

Age at diagnosis group code

1 < 5
2 5-14
3 15-34
4 35-44
5 45-54
6 55-64
7 65-74
8 75+

Race group code

1 white
2 black
3 other*****
9 unknown (includes unknown race, race code 98 'other unspecified' for diagnosis yrs. 1991+, or race suppressed)

* other race includes Asian-American, Pacific Islanders, American Indians, Alaskan Natives, and all other races. 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.

NOTE: Data for white residents, black residents and residents of all other races, are not available for every county (see list of counties above). Fifteen counties have sufficiently large black populations and three counties had sufficiently large other populations to allow meaningful statistics for the race group. For the remaining counties, race data are suppressed, therefore, you must use the state file to obtain a rate by race for Illinois.

Hispanic codes

1 Hispanic*
0 non-Hispanic (including unknown for appropriate counties)
9 Hispanic data suppressed or not available for diagnosis years 1988-1992

NOTE: Five counties have sufficiently large Hispanic populations (see list of counties above) to allow for meaningful statistics for the ethnicity group. For the remaining counties, Hispanic data are suppressed; therefore, you must use the state file to obtain a rate by Hispanic ethnicity for Illinois

*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;55:182-186.

NOTE: As noted above, some data for race and Hispanic ethnicity in the county file have been suppressed. Therefore, to compare county race or ethnicity data with IL you will need to obtain the state data from the IL8622 file.

File Layout for ZIP code file ZPCD8822.dat (number of records 2,239,859)

)Record format (all fields numeric)

Data Field	Positions	Length
sex code	1-1	1
diagnosis year group (5yr groups)	2-2	1
zip code at diagnosis	3-7	5
stage of disease at diagnosis	8-8	1
cancer site groups	9-10	2
age at diagnosis group code	11-11	1
latitude to the centroid of the zip code	12-24	13.7
longitude to the centroid of the zip code	25-37	13.7

Codes for data fields

Sex code

1 Male
2 Female

Diagnosis Year Group (5yr groups)

1 1988-1992
2 1993-1997
3 1998-2002
4 2003-2007
5 2008-2012
6 2013-2017
7 2018-2022

Zip code at diagnosis

Valid Illinois Zip codes

Stage of disease at diagnosis

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

Cancer site group (based on cancer groups at county level)

1 oral cavity & pharynx (1)
2 colorectal (4)
3 lung & bronchus (7)
4 breast-invasive female (10 for females only)
5 cervix (11)
6 prostate (14)
7 urinary system (16,17)
8 central nervous system (18)
9 leukemias and lymphomas(19,20,22)
10 all other cancers (2,3,5,6,8,9,12,13,15,21,and 23 (10 for males only)
11 breast-in-situ female (24 for females only)

NOTE: Due to very small numbers of cases, male breast in-situ cases have been omitted from the file (N=207 for all years combined 1987-2021)

Age at diagnosis group code

1	0-14
2	15-44
3	45-64
4	65+

NOTE: Zip code data are 100% complete with 99.9% accuracy. Please note the latitude and longitude provided are coded to the centroid of the zip code, not to a street-level address and are expressed in decimal degrees. Data were geocoded with MapMarker V31 using coordinate system North American Datum of 1983 (NAD83).

The following website has a utility available that will convert decimal degrees to degrees, minutes and seconds: https://www.fcc.gov/media/radio/dms-decimal

Latitude to the centroid of the zip code
Longitude to the centroid of the zip code

File Layout for COOK1322.dat (number of records 273,121)

Record format (all fields are numeric)

Data Field	Positions	Length
sex code	1-1	1
diagnosis year group	2-2	1
Cook county part	3-3	1
report source	4-4	1
stage of disease at diagnosis	5-5	1
cancer site group	6-7	2
age at diagnosis group code	8-8	1
race group code	9-9	1
Hispanic code	10-10	1

Codes for Data fields

Sex code

1 male
2 female

Diagnosis year group

1 2013-2017
2 2018-2022

Cook County Part

1 Chicago
2 Cook Suburbs

Report Source

1 hospital or clinic
2 radiation treatment centers/medical oncology centers
3 laboratory only
4 physician's office
5 nursing home or hospice
6 autopsy only
7 death certificate only
8 other hospital, out-patient units/surgery

Stage of disease at diagnosis

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

Cancer Site Group (based on SEER group codes)

1 oral cavity & pharynx (20010-20100)
2 esophagus (21010)
3 stomach (21020)
4 colorectal (21041-21052)
5 liver (21071)
6 pancreas (21100)
7 lung & bronchus (22030)
8 bone (23000)
9 melanomas (25010)
10 breast-invasive only (26000, behavior code 3)
11 cervix (27010)
12 uterus (27020-27030)
13 ovary (27040)
14 prostate (28010)
15 testis (28020)
16 bladder (29010)
17 kidney (29020)
18 nervous system (31010-31040)
19 Hodgkin's lymphomas (33011-33012)
20 non-Hodgkin's lymphomas (33041-33042)
21 myelomas (34000)
22 leukemias (35011-35043)
23 all other sites (21030,21060,21072,21080,21090,21110, 21120,21130,22010,22020,22050,22060, 24000,25020,27050,27060,27070,28030, 28040,29030,29040,30000,32010,32020, 36010,36020,37000)
24 breast-in-situ only (26000, behavior code 2)

Age at diagnosis group code

1 <5
2 5-14
3 15-34
4 35-44
5 45-54
6 55-64
7 65-74
8 75 +

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. 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.

Hispanic codes

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 internals. The American Statistician 2001;55:182-186