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

User Manual

Cancer Incidence Public Dataset (27th edition), data as of November 2019, released April 2020
Illinois Data – IL8617.dat
County Data – CNTY8817.dat
ZIP code data – ZPCD8817.dat
Cook County with Cook suburbs and Chicago – COOK0817.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 none should be inferred.

Citation

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 2017. 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 and 2013-17) and five-year aggregate for Cook County with Cook suburbs and Chicago is (2008-12 and 2013-17). The following article describes the method used to measure uniqueness of the files.


The files include Illinois incidence data for invasive cancers only with the exception of cancers of the bladder. Carcinoma in-situ 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 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 on the basis of new information. For this reason, an updated version of previous years data are 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 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 developed by the North American Association of Central Cancer Registries (NAACCR) and adopted by the National Program 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/

<table>
<thead>
<tr>
<th>DX Year</th>
<th>Completeness</th>
<th>DXYrGrp</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>87%</td>
<td>1988-92</td>
<td>88%</td>
</tr>
<tr>
<td>1989</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>91%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>92%</td>
<td>1993-97</td>
<td>94%</td>
</tr>
<tr>
<td>1994</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For data from 1995-2016 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/

<table>
<thead>
<tr>
<th>DX Year</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>100%</td>
</tr>
<tr>
<td>1996</td>
<td>100%</td>
</tr>
<tr>
<td>1997</td>
<td>100%</td>
</tr>
<tr>
<td>1998</td>
<td>100%</td>
</tr>
<tr>
<td>1999</td>
<td>100%</td>
</tr>
<tr>
<td>2000</td>
<td>100%</td>
</tr>
<tr>
<td>2001</td>
<td>100%</td>
</tr>
<tr>
<td>2002</td>
<td>100%</td>
</tr>
<tr>
<td>2003</td>
<td>100%</td>
</tr>
<tr>
<td>2004</td>
<td>100%</td>
</tr>
<tr>
<td>2005</td>
<td>100%</td>
</tr>
<tr>
<td>2006</td>
<td>100%</td>
</tr>
<tr>
<td>2007</td>
<td>100%</td>
</tr>
<tr>
<td>2008</td>
<td>100%</td>
</tr>
<tr>
<td>2009</td>
<td>100%</td>
</tr>
<tr>
<td>2010</td>
<td>100%</td>
</tr>
<tr>
<td>2011</td>
<td>100%</td>
</tr>
<tr>
<td>2012</td>
<td>100%</td>
</tr>
<tr>
<td>2013</td>
<td>100%</td>
</tr>
<tr>
<td>2014</td>
<td>100%</td>
</tr>
<tr>
<td>2015</td>
<td>100%</td>
</tr>
<tr>
<td>2016</td>
<td>100%</td>
</tr>
<tr>
<td>2017</td>
<td>100%</td>
</tr>
</tbody>
</table>

In addition to the estimates for completeness of case ascertainment (as of 11/19), 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 website. https://www.naaccr.org/certification-criteria/
As of November 2019, ISCR data met the criteria for gold certification for diagnosis years 1995-2017. Certification status for 2017 data will be awarded in June 2020. [NOTE: diagnosis year 1995 now meets criteria for gold certification.]

<table>
<thead>
<tr>
<th>Year</th>
<th>Pass Edits (%)</th>
<th>DCO*(%)</th>
<th>Unresolved Duplicates** (%)</th>
<th>Sex (%)</th>
<th>Age (%)</th>
<th>County (%)</th>
<th>Race*** (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1987</td>
<td>~</td>
<td>~</td>
<td>~</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1988</td>
<td>~</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1989</td>
<td>~</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1990</td>
<td>100</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1991</td>
<td>100</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>1992</td>
<td>100</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>1993</td>
<td>100</td>
<td>~</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>1994</td>
<td>100</td>
<td>2.2</td>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>1995</td>
<td>100</td>
<td>2.6</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>1996</td>
<td>100</td>
<td>1.8</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>1997</td>
<td>100</td>
<td>1.8</td>
<td>0.09</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.6</td>
</tr>
<tr>
<td>1998</td>
<td>100</td>
<td>1.5</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>1999</td>
<td>100</td>
<td>1.8</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2000</td>
<td>100</td>
<td>2.3</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2001</td>
<td>100</td>
<td>2.4</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2002</td>
<td>100</td>
<td>2.6</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2003</td>
<td>100</td>
<td>1.5</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2004</td>
<td>100</td>
<td>1.7</td>
<td>0.01</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.9</td>
</tr>
<tr>
<td>2005</td>
<td>100</td>
<td>1.9</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2006</td>
<td>100</td>
<td>2.0</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2007</td>
<td>100</td>
<td>1.2</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>2008</td>
<td>100</td>
<td>1.7</td>
<td>0.07</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2009</td>
<td>100</td>
<td>1.6</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2010</td>
<td>100</td>
<td>1.8</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2011</td>
<td>100</td>
<td>1.8</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.4</td>
</tr>
<tr>
<td>2012</td>
<td>100</td>
<td>0.9</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
</tr>
<tr>
<td>2013</td>
<td>100</td>
<td>1.0</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2014</td>
<td>100</td>
<td>2.1</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2015</td>
<td>100</td>
<td>2.1</td>
<td>0.02</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
</tr>
<tr>
<td>2016</td>
<td>100</td>
<td>2.0</td>
<td>0.09</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2017</td>
<td>100</td>
<td>2.1</td>
<td>0.06</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
</tbody>
</table>

~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 2001-2017

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’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
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 lori.koch@illinois.gov
### File Layout for State file IL8617.dat (number of records 1,947,354)

#### Record Format (all fields are numeric).

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

#### Codes for Data fields

**Sex code**

1  male  
2  female

**Diagnosis Year**

<table>
<thead>
<tr>
<th>Year</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>2001</td>
</tr>
<tr>
<td>1987</td>
<td>2002</td>
</tr>
<tr>
<td>1988</td>
<td>2003</td>
</tr>
<tr>
<td>1989</td>
<td>2004</td>
</tr>
<tr>
<td>1990</td>
<td>2005</td>
</tr>
<tr>
<td>1991</td>
<td>2006</td>
</tr>
<tr>
<td>1992</td>
<td>2007</td>
</tr>
<tr>
<td>1993</td>
<td>2008</td>
</tr>
<tr>
<td>1994</td>
<td>2009</td>
</tr>
<tr>
<td>1995</td>
<td>2010</td>
</tr>
<tr>
<td>1996</td>
<td>2011</td>
</tr>
<tr>
<td>1997</td>
<td>2012</td>
</tr>
<tr>
<td>1998</td>
<td>2013</td>
</tr>
<tr>
<td>1999</td>
<td>2014</td>
</tr>
<tr>
<td>2000</td>
<td>2015</td>
</tr>
<tr>
<td></td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td>2017</td>
</tr>
</tbody>
</table>
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 2001-2017, were reported with version 3 (ICD-O-3) codes 3, 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:

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
Pediatric codes level 1 (ICCC ICD-O-3)


1 I Leukemias, myeloproliferative diseases, and myelodysplastic 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
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.


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 or not microscopically confirmed
File Layout for county file CNTY8817.dat (number of records 1,857,941)

Record Format (all fields are numeric)

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Positions</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>sex code</td>
<td>1-1</td>
<td>1</td>
</tr>
<tr>
<td>diagnosis year group</td>
<td>2-2</td>
<td>1</td>
</tr>
<tr>
<td>county at diagnosis code</td>
<td>3-5</td>
<td>3</td>
</tr>
<tr>
<td>report source</td>
<td>6-6</td>
<td>1</td>
</tr>
<tr>
<td>stage of disease at diagnosis</td>
<td>7-7</td>
<td>1</td>
</tr>
<tr>
<td>cancer site group codes</td>
<td>8-9</td>
<td>2</td>
</tr>
<tr>
<td>age at diagnosis group code</td>
<td>10-10</td>
<td>1</td>
</tr>
<tr>
<td>race group code</td>
<td>11-11</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic code</td>
<td>12-12</td>
<td>1</td>
</tr>
</tbody>
</table>

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

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
<table>
<thead>
<tr>
<th>County</th>
<th>Race Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crawford</td>
<td>all race only</td>
</tr>
<tr>
<td>Cumberland</td>
<td>all race only</td>
</tr>
<tr>
<td>DeKalb</td>
<td>all race only</td>
</tr>
<tr>
<td>DeWitt</td>
<td>all race only</td>
</tr>
<tr>
<td>Douglas</td>
<td>all race only</td>
</tr>
<tr>
<td>DuPage</td>
<td>all race, white, black, other and Hispanic/non-Hispanic</td>
</tr>
<tr>
<td>Edgar</td>
<td>all race only</td>
</tr>
<tr>
<td>Edwards</td>
<td>all race only</td>
</tr>
<tr>
<td>Effingham</td>
<td>all race only</td>
</tr>
<tr>
<td>Fayette</td>
<td>all race only</td>
</tr>
<tr>
<td>Ford</td>
<td>all race only</td>
</tr>
<tr>
<td>Franklin</td>
<td>all race only</td>
</tr>
<tr>
<td>Fulton</td>
<td>all race only</td>
</tr>
<tr>
<td>Gallatin</td>
<td>all race only</td>
</tr>
<tr>
<td>Greene</td>
<td>all race only</td>
</tr>
<tr>
<td>Grundy</td>
<td>all race only</td>
</tr>
<tr>
<td>Hamilton</td>
<td>all race only</td>
</tr>
<tr>
<td>Hancock</td>
<td>all race only</td>
</tr>
<tr>
<td>Hardin</td>
<td>all race only</td>
</tr>
<tr>
<td>Henderson</td>
<td>all race only</td>
</tr>
<tr>
<td>Henry</td>
<td>all race only</td>
</tr>
<tr>
<td>Iroquois</td>
<td>all race only</td>
</tr>
<tr>
<td>Jackson</td>
<td>all race only</td>
</tr>
<tr>
<td>Jasper</td>
<td>all race only</td>
</tr>
<tr>
<td>Jefferson</td>
<td>all race only</td>
</tr>
<tr>
<td>Jersey</td>
<td>all race only</td>
</tr>
<tr>
<td>JoDaviess</td>
<td>all race only</td>
</tr>
<tr>
<td>Johnson</td>
<td>all race only</td>
</tr>
<tr>
<td>Kane</td>
<td>all race, white, black and Hispanic/non-Hispanic</td>
</tr>
<tr>
<td>Kankakee</td>
<td>all race, white, black</td>
</tr>
<tr>
<td>Kendall</td>
<td>all race only</td>
</tr>
<tr>
<td>Knox</td>
<td>all race only</td>
</tr>
<tr>
<td>Lake</td>
<td>all race, white, black, other and Hispanic/non-Hispanic</td>
</tr>
<tr>
<td>LaSalle</td>
<td>all race only</td>
</tr>
<tr>
<td>Lawrence</td>
<td>all race only</td>
</tr>
<tr>
<td>Lee</td>
<td>all race only</td>
</tr>
<tr>
<td>Livingston</td>
<td>all race only</td>
</tr>
<tr>
<td>Logan</td>
<td>all race only</td>
</tr>
<tr>
<td>McDonough</td>
<td>all race only</td>
</tr>
<tr>
<td>McHenry</td>
<td>all race only</td>
</tr>
<tr>
<td>McLean</td>
<td>all race only</td>
</tr>
<tr>
<td>Macon</td>
<td>all race, white, black</td>
</tr>
<tr>
<td>Macoupin</td>
<td>all race only</td>
</tr>
<tr>
<td>Madison</td>
<td>all race, white, black</td>
</tr>
<tr>
<td>Marion</td>
<td>all race only</td>
</tr>
<tr>
<td>Marshall</td>
<td>all race only</td>
</tr>
<tr>
<td>Mason</td>
<td>all race only</td>
</tr>
<tr>
<td>Massac</td>
<td>all race only</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>in-situ</td>
</tr>
<tr>
<td>1</td>
<td>localized</td>
</tr>
<tr>
<td>2</td>
<td>regional</td>
</tr>
<tr>
<td>3</td>
<td>distant metastases/systemic disease</td>
</tr>
<tr>
<td>9</td>
<td>Unknown, unstaged or unspecified</td>
</tr>
</tbody>
</table>

### Cancer site group (based on SEER group codes)

<table>
<thead>
<tr>
<th>Code</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>oral cavity &amp; pharynx (20010-20100)</td>
</tr>
<tr>
<td>2</td>
<td>esophagus (21010)</td>
</tr>
<tr>
<td>3</td>
<td>stomach (21020)</td>
</tr>
<tr>
<td>4</td>
<td>colorectal (21041-21052)</td>
</tr>
<tr>
<td>5</td>
<td>liver (21071)</td>
</tr>
<tr>
<td>6</td>
<td>pancreas (21100)</td>
</tr>
<tr>
<td>7</td>
<td>lung &amp; bronchus (22030)</td>
</tr>
<tr>
<td>8</td>
<td>bone (23000)</td>
</tr>
<tr>
<td>9</td>
<td>melanomas (25010)</td>
</tr>
<tr>
<td>10</td>
<td>breast-invasive only (26000, behavior code 3)</td>
</tr>
<tr>
<td>11</td>
<td>cervix (27010)</td>
</tr>
<tr>
<td>12</td>
<td>uterus (27020-27030)</td>
</tr>
<tr>
<td>13</td>
<td>ovary (27040)</td>
</tr>
<tr>
<td>14</td>
<td>prostate (28010)</td>
</tr>
<tr>
<td>15</td>
<td>testis (28020)</td>
</tr>
<tr>
<td>16</td>
<td>bladder (29010)</td>
</tr>
<tr>
<td>17</td>
<td>kidney (29020)</td>
</tr>
<tr>
<td>18</td>
<td>nervous system (31010-31040)</td>
</tr>
<tr>
<td>19</td>
<td>Hodgkin's lymphomas (33011-33012)</td>
</tr>
<tr>
<td>20</td>
<td>non-Hodgkin's lymphomas (33041-33042)</td>
</tr>
<tr>
<td>21</td>
<td>myelomas (34000)</td>
</tr>
<tr>
<td>22</td>
<td>leukemias (35011-35043)</td>
</tr>
<tr>
<td>23</td>
<td>all other sites</td>
</tr>
<tr>
<td></td>
<td>(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)</td>
</tr>
<tr>
<td>24</td>
<td>breast-in-situ only (26000, behavior code 2)</td>
</tr>
</tbody>
</table>

### Age at diagnosis group code

<table>
<thead>
<tr>
<th>Code</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>2</td>
<td>5-14</td>
</tr>
<tr>
<td>3</td>
<td>15-34</td>
</tr>
<tr>
<td>4</td>
<td>35-44</td>
</tr>
<tr>
<td>5</td>
<td>45-54</td>
</tr>
<tr>
<td>6</td>
<td>55-64</td>
</tr>
<tr>
<td>7</td>
<td>65-74</td>
</tr>
<tr>
<td>8</td>
<td>75+</td>
</tr>
</tbody>
</table>
**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. 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.

**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 1989-1993  

**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.


*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 IL8614 file.
File Layout for ZIP code file ZPCD8817.dat (number of records 1,857,777)

Record format (all fields numeric)

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

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

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=164 for all years combined 1988-2017)
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 COOK0817.dat** (number of records 266,590)

Record format (all fields are numeric)

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Positions</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>sex code</td>
<td>1-1</td>
<td>1</td>
</tr>
<tr>
<td>diagnosis year group</td>
<td>2-2</td>
<td>1</td>
</tr>
<tr>
<td>Cook county part</td>
<td>3-3</td>
<td>1</td>
</tr>
<tr>
<td>report source</td>
<td>4-4</td>
<td>1</td>
</tr>
<tr>
<td>stage of disease at diagnosis</td>
<td>5-5</td>
<td>1</td>
</tr>
<tr>
<td>cancer site group</td>
<td>6-7</td>
<td>2</td>
</tr>
<tr>
<td>age at diagnosis group code</td>
<td>8-8</td>
<td>1</td>
</tr>
<tr>
<td>race group code</td>
<td>9-9</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic code</td>
<td>10-10</td>
<td>1</td>
</tr>
</tbody>
</table>

**Codes for Data fields**

**Sex code**
1  male  
2  female

**Diagnosis year group**
1  2008-2012  
2  2013-2017

**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)

<table>
<thead>
<tr>
<th>Cancer Site Group</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>oral cavity &amp; pharynx</td>
<td>20010-20100</td>
</tr>
<tr>
<td>esophagus</td>
<td>21010</td>
</tr>
<tr>
<td>stomach</td>
<td>21020</td>
</tr>
<tr>
<td>colorectal</td>
<td>21041-21052</td>
</tr>
<tr>
<td>liver</td>
<td>21071</td>
</tr>
<tr>
<td>pancreas</td>
<td>21100</td>
</tr>
<tr>
<td>lung &amp; bronchus</td>
<td>22030</td>
</tr>
<tr>
<td>bone</td>
<td>23000</td>
</tr>
<tr>
<td>melanomas</td>
<td>25010</td>
</tr>
<tr>
<td>breast-invasive only</td>
<td>26000</td>
</tr>
<tr>
<td>cervix</td>
<td>27010</td>
</tr>
<tr>
<td>uterus</td>
<td>27020-27030</td>
</tr>
<tr>
<td>ovary</td>
<td>27040</td>
</tr>
<tr>
<td>prostate</td>
<td>28010</td>
</tr>
<tr>
<td>testis</td>
<td>28020</td>
</tr>
<tr>
<td>bladder</td>
<td>29010</td>
</tr>
<tr>
<td>kidney</td>
<td>29020</td>
</tr>
<tr>
<td>nervous system</td>
<td>31010-31040</td>
</tr>
<tr>
<td>Hodgkin's lymphomas</td>
<td>33011-33012</td>
</tr>
<tr>
<td>non-Hodgkin's lymphomas</td>
<td>33041-33042</td>
</tr>
<tr>
<td>myelomas</td>
<td>34000</td>
</tr>
<tr>
<td>leukemias</td>
<td>35011-35043</td>
</tr>
<tr>
<td>all other sites</td>
<td>21030,21060,21072,21080,21090,21110,21120,21130,22010,22020,22050,22060,24000,25020,27050,27060,27070,28030,28040,29030,29040,30000,32010,7000,3020,36010,36020,37000</td>
</tr>
<tr>
<td>breast-in-situ only</td>
<td>26000</td>
</tr>
</tbody>
</table>

### Age at diagnosis group code

<table>
<thead>
<tr>
<th>Age at Diagnosis Group Code</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;5</td>
</tr>
<tr>
<td>2</td>
<td>5-14</td>
</tr>
<tr>
<td>3</td>
<td>15-34</td>
</tr>
<tr>
<td>4</td>
<td>35-44</td>
</tr>
<tr>
<td>5</td>
<td>45-54</td>
</tr>
<tr>
<td>6</td>
<td>55-64</td>
</tr>
<tr>
<td>7</td>
<td>65-74</td>
</tr>
<tr>
<td>8</td>
<td>75+</td>
</tr>
</tbody>
</table>

### Race group code

<table>
<thead>
<tr>
<th>Race Group Code</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>white</td>
</tr>
<tr>
<td>2</td>
<td>black</td>
</tr>
<tr>
<td>3</td>
<td>other*</td>
</tr>
<tr>
<td>9</td>
<td>unknown</td>
</tr>
</tbody>
</table>

* 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 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.


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