



---

## **Indirect Sources of Information on Hispanic Ethnicity Improve Classification of Illinois Cancer Cases and Deaths**

---

Hispanic status is often misclassified on cancer incidence report forms and death certificates. Hispanic cases or decedents are erroneously classified as non-Hispanic because of insufficient information to confirm ethnicity. Spanish surnames have been used as a means to correct for such error.

Take a moment to review the following list of surnames and circle those you think have at least a 90 percent chance of belonging to individuals of Hispanic origin.

ABREGO	FIGUEROA	JASSO	OLMOS	SANCHEZ
ADAME	FLORES	LARA	PABON	SAUCEDA
BETANCOURT	GARCIA	LIMON	PENA	SOTO
BURGOS	GONZALES	LOPEZ	PEREZ	TAMEZ
CADENA	GONZALEZ	MALAVE	QUIROZ	TELLO
CEJA	HARO	MARTINEZ	RAMIREZ	TORRES
DAVILA	HEREDIA	MATA	RASCON	ULLOA
DELGADILLO	HERNANDEZ	NAJERA	RIVERA	URENA
ECHEVARRIA	IBARRA	NIEVES	RODRIGUEZ	URIAS
ELIZONDO	IRIZARRY	OCASIO	SAMANIEGO	ZUNIGA

In March 1996, a Spanish surname product became available through special study of a 1990 census data sample (1,868,781 records) used to estimate undercount. A classification scheme was created using both the relationship of the surname to the self-identification Hispanic origin response on the census form and to the frequency of occurrence of the surname in the sample. The surnames then were grouped into the following 28 Spanish surname categories.

Heavily Hispanic	Generally Hispanic	Moderately Hispanic	Occasionally Hispanic	Rarely Hispanic
0101	0201	0301	0405	5001
0102	0202	0302	0410	5005
0105	0205	0305	0415	5010
0110	0210	0310	0425	5025
0115	0215	0315		5100
0125	0225	0325		5500

The first two characters denote Hispanic classification (01 heavily Hispanic, 02 generally Hispanic, 03 moderately Hispanic, 04 occasionally Hispanic, and 50 rarely Hispanic). The criterion for placing a surname within each category depended on the proportion of householders with the surname responding affirmatively to the Hispanic origin item on the 1990 census form. The criterion for each of the five Spanish surname categories was as follows.

Spanish Surname Classification	Proportion Self-Identified As Hispanic
01 Heavily Hispanic	Over 75 Percent
02 Generally Hispanic	50 Percent < x ≤ 75 Percent
03 Moderately Hispanic	25 Percent < x ≤ 50 Percent
04 Occasionally Hispanic	5 Percent < x ≤ 25 Percent
50 Rarely Hispanic	5 Percent ≥ x

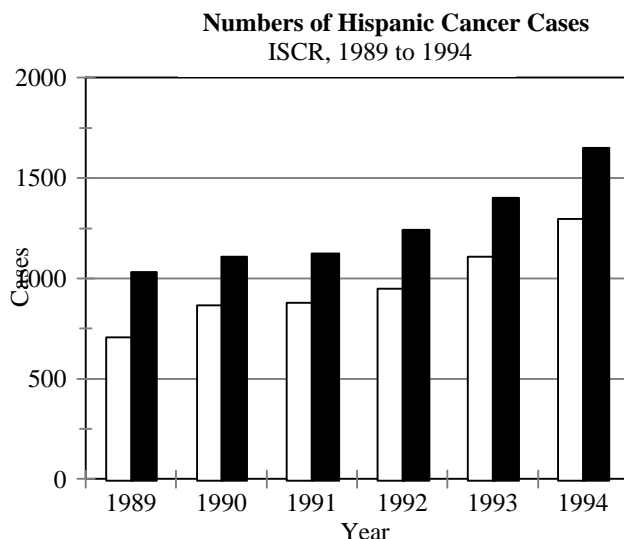
The third and fourth characters represent the frequency with which a name occurred in the study sample. (For example, category 0125 includes surnames that are heavily Hispanic with at least 25 observations.) The approach is described in more detail elsewhere.<sup>1</sup>

If you selected all of the surnames in the table above to be associated with at least a 90 percent predictive value for Hispanic status, you were correct. They all belong in the 0125 category of the census bureau Spanish surname schematic classification system. Shown below is the same list with the proportion self-identified as Hispanic during the 1990 census following the surname. The 12 highlighted surnames can identify 20 percent of the United States Hispanic population.

ABREGO	100	FIGUEROA	98	JASSO	100	OLMOS	92	<b>SANCHEZ</b>	96
ADAME	100	FLORES	94	LARA	96	PABON	93	SAUCEDA	100
BETANCOURT	92	<b>GARCIA</b>	95	LIMON	96	PENA	96	SOTO	96
BURGOS	93	<b>GONZALES</b>	92	<b>LOPEZ</b>	95	<b>PEREZ</b>	96	TAMEZ	100
CADENA	93	<b>GONZALEZ</b>	98	MALAVE	97	QUIROZ	97	TELLO	97
CEJA	96	HARO	92	<b>MARTINEZ</b>	96	<b>RAMIREZ</b>	97	<b>TORRES</b>	95
DAVILA	98	HEREDIA	98	MATA	95	RASCON	100	ULLOA	100
DELGADILLO	98	<b>HERNANDEZ</b>	97	NAJERA	94	<b>RIVERA</b>	96	URENA	100
ECHEVARRIA	90	IBARRA	97	NIEVES	97	<b>RODRIGUEZ</b>	97	URIAS	100
ELIZONDO	98	IRIZARRY	97	OCASIO	100	SAMANIEGO	100	ZUNIGA	97

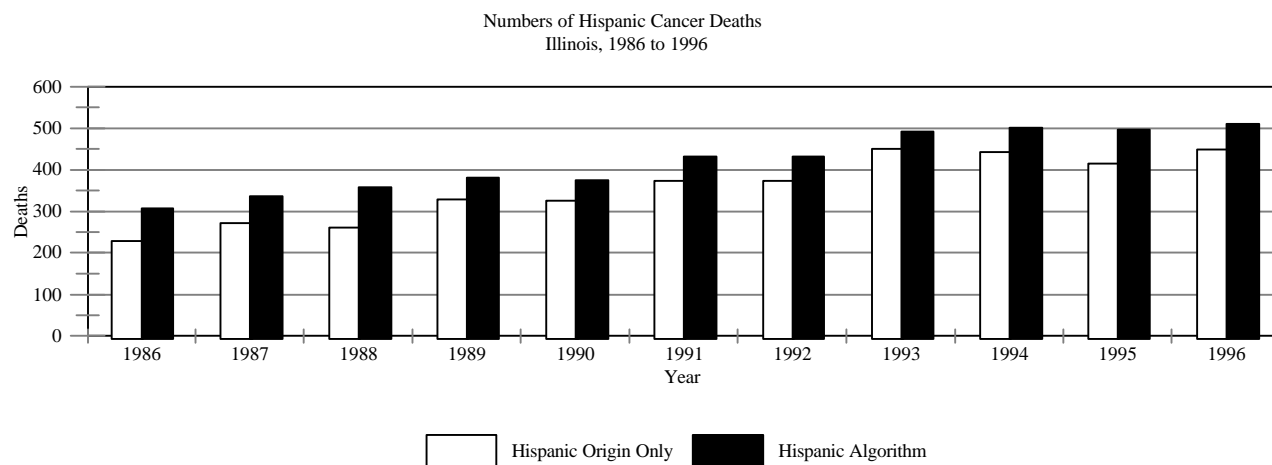
Using the census Spanish surname study scheme, two comprehensive studies were conducted to evaluate Hispanic identification on the ISCR database and on Illinois master death files. Last and maiden names on ISCR from 1989, when the Hispanic origin data element was implemented, to 1994 and on Illinois cancer death certificate files for 1986 to 1996 were classified according to the 1990 census Spanish surname study schematic codes. In addition, Hispanic non-U.S. birthplaces (Central America, Cuba, Mexico, Puerto Rico, South America, and Spain) also were studied as an indirect identifier of Hispanic status. The goal of these studies was to develop an algorithm based on direct and indirect evidence that reduced misclassification of Hispanic cancer cases and decedents in Illinois. The indirect identification component of the algorithm included cases/decedents with surnames in the heavily Hispanic categories (0101, 0102, 0105, 0110, 0115, and 0125) and/or cases/decedents born in Hispanic non-U.S. birthplaces with special adjustments that maximized correct identification indirectly. Detailed discussion of the methodology is presented elsewhere.<sup>2,3</sup>

The figure shows the numbers of cases/decedents identified as Hispanic using the Hispanic origin data element only compared with numbers using the Hispanic algorithm.



Application of the Hispanic algorithm identified on average an additional 292 cancer cases per year and an additional 64 cancer deaths per year as Hispanic that were not documented with Hispanic origin data elements. These cases/decedents have over a 90 percent chance of being Hispanic based on their surname and/or birthplace status. Hispanic identification was enhanced by about 25 percent to 30 percent on the ISCR database and about 15 percent to 20 percent on Illinois cancer death certificates with the indirect Hispanic algorithm component. When comparing Hispanic cancer incidence and mortality rates calculated using Hispanic origin only and the Hispanic algorithm, those produced with those using the Hispanic algorithm demonstrated the least difference from rates reported by the Surveillance, Epidemiology, and Ends Results (SEER) program of the National Cancer Institute.<sup>4</sup>

The use of the Hispanic algorithm should produce more valid and reliable Hispanic cancer incidence and mortality data for Illinois than Hispanic origin data elements alone. Illinois Hispanic cancer incidence and mortality data using estimates determined with the algorithm are published elsewhere.<sup>5</sup>



## References

1. Word DL and Perkins RC. *Building a Spanish Surname List for the 1990's--A New Approach to An Old Problem*. Population Division Working Paper No. 13. February 1996.
2. Dolecek TA, Howe HL. Hispanic Identification in the Illinois State Cancer Registry. Epidemiologic Report Series 98:2. Springfield, Ill.: Illinois Department of Public Health, June 1998.
3. Dolecek TA, Howe HL. Hispanic Identification on Illinois Cancer Death Certificates. Epidemiologic Report Series

98:3. Springfield, Ill.: Illinois Department of Public Health, August 1998.

4. Miller BA, Kolonel LN, Bernstein L, Young, Jr. JL, Swanson GM, West D, Key CR, Liff JM, Glover CS, Alexander GA, et al. (eds). *Racial/Ethnic Patterns of Cancer in the United States 1988-1992*, National Cancer Institute. NIH Pub. No. 96-4104. Bethesda, MD, 1996.
5. Dolecek TA, Howe HL, Snodgrass JL. Illinois Cancer Statistics Review: Incidence 1986 to 1995, Mortality 1986 to 1996. Epidemiologic Report Series 98:4. Springfield, Ill.: Illinois Department of Public Health, August 1998.

## ILLINOIS STATE CANCER REGISTRY

### Update of State of Illinois Cancer Incidence Projections

The fall 1997 newsletter presented cancer incidence projections for the state of Illinois through 1998 and the detailed methodology by which these data were established. The 1998 update of these projections is provided below and includes projections for 1999. The estimates are listed by sex for all cancers and for selected cancer types.

**Projected Cancer Incidence  
All Sites and Selected Sites, Illinois 1996-1999  
(Revised July 1998)**

	1996	1997	1998	1999
<b>All Sites</b>	54,407	54,822	55,236	55,650
<b>All Sites (Males)</b>	27,378	27,607	27,836	28,066
<b>All Sites (Females)</b>	27,030	27,215	27,400	27,585
<b>Colorectal</b>	6,724	6,785	6,847	6,908
<b>Colorectal (Males)</b>	3,257	3,290	3,324	3,357
<b>Colorectal (Females)</b>	3,466	3,495	3,523	3,552
<b>Lung</b>	8,631	8,683	8,735	8,787
<b>Lung (Males)</b>	5,132	5,171	5,210	5,249
<b>Lung (Females)</b>	3,500	3,512	3,525	3,538
<b>Invasive Breast (Females)</b>	8,134	8,198	8,262	8,326
<b>Invasive Cervix</b>	722	729	735	742
<b>Prostate</b>	7,120	7,171	7,222	7,272

### Discrepancy Reports Re-instituted

Recently ISCR staff began using the EDITS program to identify discrepancies in data submissions. All abstracts, electronic and manual submissions, are run through the EDITS program before they are entered onto the ISCR database. Individual error reports are then mailed to facilities for reconciliation. Individual quarterly discrepancy reports also are prepared to show each facility's discrepancy rate for the preceding three months.

The most common discrepancies identified using the EDITS software are as follows:

- ! Conflicts between county code and ZIP code;
- ! Submission of Extent of Disease and TNM staging for sites that do not have a valid TNM or EOD stage;
- ! Reporting the stage of disease as distant without information recorded in the site of distant mets;
- ! Date treatment initiated inconsistent with reported treatment or diagnosis date;
- ! The use of punctuation in the address; and
- ! Using the alpha letter 'o' (O) in the staging fields instead of zero (0).

The EDITS program is available to reporting facilities. The ISCR version contains modifications to the EDITS program that reflect changes specific to Illinois. If you are interested in obtaining this modified version of the EDITS software, please call Mike Goza at 217-557-4089. We strongly encourage you to order and use this software!

## ADVERSE PREGNANCY OUTCOMES REPORTING SYSTEM

### Update on Electronic Reporting

Since December 1996 reporting facilities have had an option to report cases electronically. The software is free and can be run on most personal computers. When requested, an employee from the Division will visit the reporting facility to install the software and train the hospital staff. Additional technical support is available from the APORS central office by telephone. Marjorie Higgins coordinates electronic reporting and assists hospital reporters when needed and Mike Goza performs the site visits. To date, the software has been installed at 39 of the 163 reporting hospitals. About 16 percent of all APORS reports are

received electronically.

Electronic reporting saves time, reduces effort and improves data quality. Hospitals may send the reports on diskette or by E-mail to APORS. Reports sent by E-mail reduces the transit time to minutes. Electronic reporting reduces by approximately one-third the amount of time to process each report since manual data entry is not necessary. This also eliminates



potential data entry errors. The information is then verified, coded and added to the data base.

The APORS program is working to improve and expand electronic reporting so that more Illinois hospitals can adopt this method. APORS' goal is to have all hospitals reporting electronically within three years.

Further information about electronic reporting is available from the APORS program. Please contact Marjorie Higgins at 217-557-4090 for assistance.

## CENSUS OF FATAL OCCUPATIONAL INJURIES

### Occupation-related Fatalities among Women

Women's job related injuries and fatalities are substantially below those of men, largely because of their lower proportions in high-risk industries and occupations. Data from 1992-1996 from the U.S. Bureau of Labor Statistics (BLS) reveal that women in the United States comprise less than half the total workforce; women incurred less than one-tenth of the job related fatal injuries. Illinois data for 1992 through 1996 showed that 6 percent (n=76) of all occupational fatalities (n=1,264) were women.



Homicides and transportation events accounted for nearly two-thirds of the work injury deaths among women in the U.S. As shown in the accompanying table, deaths in these categories among Illinois women accounted for 79 percent of all occupational fatalities. Seventy percent of the homicides in women were from shootings. The fatalities occurred primarily during robbery attempts mostly in the industrial grouping, retail trade, such as grocery stores, gasoline service stations, eating and drinking establishments. Sixty percent of the women, who were murdered, worked in service occupations as cashiers, clerks, waitresses, sales supervisors or proprietors. Most of these women were between the ages of 25 and 40 years old.

A total of 43 percent of the transportation fatalities among women were caused by automobile crashes. Occupations for these fatalities included social service, health care workers, sales workers or governmental agency workers. An additional 37 percent of the transportation fatalities among women were from bus or truck crashes. The occupations were listed as drivers or laborers. The age range of women who suffered fatal transportation accidents were between 40 and 80 years of age.

In sum, women do not suffer fatal accidents at the same rate as men when comparing events, as shown in the table below. Thirty-nine percent of the women who suffered fatal occupational injuries died as the results of homicides, while only 17 percent of the men suffered from homicides. More time and data, however, are needed to ascertain if serious injuries incurred by women will increase relative to their increased employment in high-risk occupations

### Job-related Fatalities by Event or Exposure<sup>a</sup> and Gender

Event or Exposure	Female		Male		Number
	Number	Percent	Number	Percent	
Homicides	30	39.5	204	17.2	234
Transportation	30	39.5	401	33.8	431
Fire and Explosion	-	-	41	3.5	41
Falls	5	6.6	185	15.6	190
Contact with Objects	7	9.2	209	17.6	216
Harmful Exposure	4	5.3	148	12.5	152
Total	76	100.0	1,188	100.0	1,264

SOURCE: Illinois Department of Public Health, Division of Epidemiologic Studies, on behalf of the U.S. Bureau of Labor Statistics, in cooperation with state and federal agencies, Census of Fatal Occupational Injuries, 1992-1996.

<sup>a</sup>Based on 1992 -1996 BLS Occupational Injury and Illness Classification Structures

NOTE: Percent does not add to 100 due to rounding.

The REGISTRY NEWSLETTER is published by the Division of Epidemiologic Studies, Illinois Department of Public Health, 605 W. Jefferson St., Springfield, IL 62762, 217-785-1873, TTY (hearing impaired use only) 800-547-0466.

Jim Edgar, Governor  
 John R. Lumpkin, MD, Director of Public Health  
 Holly L. Howe, PhD, Chief, Division of Epidemiologic Studies  
 Therese A. Dolecek, PhD, Senior Epidemiologist  
 Jan Snodgrass, Manager, Illinois State Cancer Registry  
 Trish Egler, Manager, Adverse Pregnancy Outcomes Reporting System  
 Bess Edmondson, Staff Member, Census of Fatal Occupational Injuries



Printed by Authority of the State of Illinois  
 P.O. #549135 2M 12/98