



©Tony Savino/The Image Works

# B Technical Notes and Methods

## Study Population

The *Atlas of Stroke Hospitalizations Among Medicare Beneficiaries* presents stroke hospitalization data for fee-for-service Medicare beneficiaries ages 65 years and older who resided in the United States, Puerto Rico, or the U.S. Virgin Islands during 1995–2002.

On average, there were 27,759,446 Medicare beneficiaries in our study population per year during 1995–2002. Nearly 60% of the beneficiaries were women and approximately 40% were men. The racial/ethnic composition of the Medicare beneficiaries during this period was 7.8% black, 1.5% Hispanic, 87.7% white, and 3% other. Hispanics may be underrepresented in this Atlas in part because of the methods used for reporting race and ethnicity (see Definitions, Race/Ethnicity, on page 176 for more details). The age distribution of beneficiaries was 52.6% in the youngest age group (65–74 years), 34.8% in the middle age group (75–84 years), and 12.6% in the oldest age group ( $\geq 85$  years). For more information about the study population, see Section 1: National Patterns of Stroke Hospitalizations.

## Data Sources

### **Medicare Hospital Claims Data**

We obtained hospital claims data for Medicare beneficiaries ages 65 years and older for each year during 1995–2002 from the Centers for Medicare & Medicaid Services (CMS) Medicare Provider Analysis and Review (MEDPAR) file, Part A. We abstracted data on beneficiaries from all 50 states and

the District of Columbia, as well as on those from Puerto Rico and the U.S. Virgin Islands.

We included all hospitalizations among Medicare beneficiaries meeting our inclusion criteria for which the principal (i.e., first-listed) diagnosis was stroke, as indicated by *International Classification of Diseases, 9<sup>th</sup> Revision, Clinical Modification*<sup>1</sup> (ICD-9-CM) codes 430–434 and 436–438, but excluded those for which the principal diagnosis was transient ischemic attack (ICD-9-CM code 435).

The data in the MEDPAR Part A files do not distinguish between repeat hospitalizations and first-time hospitalizations. Therefore, the data presented in this *Atlas* refer to the number and rate of hospitalizations for stroke rather than to the number of people who were hospitalized for stroke, since some people were hospitalized more than once during the period covered in this *Atlas* (1995–2002).

### **Medicare Denominator Files**

We obtained the total number of Medicare beneficiaries ages 65 years and older who resided in the United States, Puerto Rico, or the U.S. Virgin Islands during 1995–2002 from the Medicare Denominator Files. The Medicare Denominator Files are annual files that contain data on all Medicare beneficiaries enrolled or entitled to benefits in a given year. Beneficiaries were excluded from the denominator if they were members of a health maintenance organization (HMO), died before July 1 of any of the years included in this *Atlas*, or were younger than 65 years on July 1 of any of the years included in this *Atlas*.

## **Area Resource File**

The Area Resource File<sup>2</sup> is a collection of data from more than 50 sources, including the following: the American Medical Association, American Hospital Association, U.S. Census Bureau, Centers for Medicare & Medicaid Services, Bureau of Labor Statistics, and National Center for Health Statistics. The Area Resource File is designed to be used by planners, policy makers, researchers, and others interested in the nation's health care delivery system and factors that may impact health status and health care in the United States.

We obtained the 2002 data on the types of health care facilities (i.e., all short-term general hospitals and those with neurological services, emergency departments, and rehabilitation care services) from the 2004 Area Resource File. We obtained the 2002 data on the number of stroke specialists (i.e., neurologists and neurosurgeons) from the 2005 Area Resource File. We used the contiguity matrix from the 2002 Area Resource File to spatially smooth the stroke hospitalization data.

## **Definitions**

### **Selected Comorbidities**

Our analyses of selected comorbidities were based on the following ICD-9-CM codes listed in diagnosis fields 2–10 (given a principal diagnosis of stroke as defined by ICD-9-CM codes 430–434 and 436–438):

#### **Atrial Fibrillation (427.3)**

- Atrial fibrillation (427.31)
- Atrial flutter (427.32)

#### **Diabetes (250)**

- Diabetes mellitus without mention of complication (250.0)
- Diabetes with ketoacidosis (250.1)
- Diabetes with hypertension (250.2)
- Diabetes with coma (250.3)
- Diabetes with renal manifestations (250.4)
- Diabetes with ophthalmic manifestations (250.5)
- Diabetes with neurological manifestations (250.6)
- Diabetes with peripheral circulatory disorders (250.7)
- Diabetes with other specified manifestations (250.8)
- Diabetes with unspecified complication (250.9)

#### **Hypertension (401–405)**

- Essential hypertension (401)
- Hypertensive heart disease (402)
- Hypertensive renal disease (403)
- Hypertensive heart and renal disease (404)
- Secondary hypertension (405)

### **Hospital Charges**

Hospital charges, defined as the total charges for stroke hospitalizations among Medicare beneficiaries, were obtained from the MEDPAR Part A files. This category includes the following inpatient hospital charges included in Part A of Medicare insurance: a semi-private room, meals, regular nursing services, operating and recovery rooms, intensive care, inpatient prescription drugs, laboratory tests, X-rays, psychiatric care, inpatient rehabilitation, and long-term care hospitalization when medically necessary, as well as other medically necessary services and supplies provided in the hospital.<sup>3</sup>

### ***Hospital Discharge Status***

Discharge destinations are defined by Medicare. They include the following categories:

**Home**—residence or place other than a health care facility where patient receives self-care, health service care, or IV drug therapy.

**Skilled Nursing Facility**—facility that meets specific regulatory certification requirements and that primarily provides inpatient skilled nursing care and related services to patients who require medical, nursing, or rehabilitative services but does not provide the level of care or treatment available in a hospital.

**Other Care Facility**—intermediate care, short-term care, or other type of facility. This category does not include skilled nursing facilities.

**Any Care Facility**—skilled nursing, intermediate care, short-term care, or other type of facility. This category includes skilled nursing facilities and other care facilities (listed above).

**Died Before Discharge**—patient died during the hospital stay.

**Died Within 30 Days**—patient died within 30 days of admission to the hospital for a stroke.

### ***Race/Ethnicity***

In this publication, we only present data on patients in three major racial/ethnic categories from the MEDPAR file (blacks, Hispanics, and whites).<sup>4</sup> We identified the race/ethnicity of each beneficiary on the basis of the race code on the claim record for a patient's hospital stay. Since race and Hispanic ethnicity were not reported separately in the Medicare databases, the categories of black, Hispanic, and

white are mutually exclusive.<sup>4</sup> Therefore, a person who is Hispanic and white was reported as either Hispanic or white.

This reporting practice can result in misclassification of race and ethnicity. According to 1996 data, the probability that the racial/ethnic designation on Medicare claims data is correct is 96.6% for whites and 95.5% for blacks, but only 19.4% for Hispanics.<sup>4</sup> At the same time, the probability that a person identified as Hispanic in the Medicare data set is actually Hispanic is 98%. Together, these data suggest that Hispanics are underreported in the Medicare data sets and that this underreporting could introduce bias into the results presented here.

We recognize that racial categories reflect socially distinct groups of people but not biological differences.<sup>5,6</sup> Moreover, we recognize that each racial/ethnic group includes people of considerable diversity with regard to culture, socioeconomic status, heritage, and area of residence.

### ***Stroke Subtypes***

Analyses of stroke subtypes were based on the following ICD-9-CM codes listed as the principal diagnosis on the hospital claims:

Hemorrhagic (430–432)

Ischemic (433–434)

Ill-Defined (436–437)

Late effects of cerebrovascular disease (438)

## Statistical Analysis

### ***Rationale for Spatial Smoothing***

County estimates of stroke hospitalization rates and percentages of stroke hospitalizations by discharge status and selected comorbidities can be unstable in counties with small populations. This problem is particularly relevant in analyses of geographic disparities among racial and ethnic groups because many counties have small or nonexistent minority populations. We used two approaches to reduce the statistical instability of county-level estimates: (1) the aggregation of all data for 1995–2002 and (2) the application of a statistical procedure known as spatial smoothing.

Several techniques exist for spatially smoothing data; each technique has its own set of strengths and limitations.<sup>7</sup> We chose to spatially smooth the data using a spatial moving average. Numerators (e.g., the number of stroke hospitalizations) and denominators (e.g., number of Medicare beneficiaries) for each county were combined with the numerators and denominators of immediate neighboring counties (i.e., contiguous counties), and then divided by the number of years in the study period to produce an average annual estimate. Thus, a single county's smoothed estimate represents an average of the rates of that county and all its contiguous neighbors for the study period.

As with all small area estimates of health outcomes, the county-level estimates of stroke hospitalization data presented in this Atlas are understood to be approximations of the actual geographic disparities that exist.<sup>8</sup>

### ***Calculation of Spatially Smoothed and Age-Adjusted Stroke Hospitalization Rates***

We calculated spatially smoothed and age-adjusted county-level stroke hospitalization rates for the total population and by racial/ethnic group for all strokes combined and by stroke subtype. We used the contiguity matrix for all U.S. counties from the 2002 Area Resource File database to facilitate spatial smoothing of stroke hospitalization data.

To calculate spatially smoothed and age-adjusted rates, we first summed the number of stroke hospitalizations (the numerator) in each county during 1995–2002 by age group (65–74, 75–84, and >85 years) to create summary numerators for the 8-year study period. Likewise, we summed the number of Medicare beneficiaries (the denominator) in each county during 1995–2002 by age group to create summary denominators for the 8-year study period. Next, for each age group, we combined the summary numerator of each index county with the summary numerators of its neighboring counties, and then divided this number by 8. We also combined the summary denominator of each index county with the summary denominators of its neighboring counties, and then divided this number by 8.

Next, we divided each newly calculated numerator by each newly calculated denominator to produce average annual spatially smoothed stroke hospitalization rates for each age group. Finally, we used the 2000 U.S. standard population weights for ages 65 years and older (see page 180) to calculate age-adjusted rates.

The result was average annual spatially smoothed and age-adjusted estimates of stroke hospitalization rates at the county level. We repeated these calculations to produce rates by stroke subtype for each of the three racial/ethnic groups.

Two constraints were applied to the calculation of county stroke hospitalization rates. For each racial and ethnic group, a stroke hospitalization rate was not calculated for any county for which the total number of race-specific stroke hospitalizations in that county and its neighboring counties was fewer than 20 during the study period (1995–2002).<sup>9</sup> To avoid calculating rates for counties that had no members of a particular racial/ethnic group but whose neighbors did have significant population sizes, race/ethnicity-specific rates were not calculated for a county with fewer than 10 Medicare beneficiaries of a particular racial/ethnic group during 1995–2002.

### **Standard Population Weights**

Because we calculated age-adjusted stroke hospitalization rates only for people ages 65 years and older and not for all age groups, we had to recalculate the standard population weights for the 2000 U.S. standard population for this age group. New weights for people ages 65 years and older were calculated using a two-step procedure.

First, we summed the published standard population weights for ages 65–74, 75–84, and ≥85.<sup>10</sup> Second, for each age group, we divided the original weight (based on the total age distribution) by the sum of the weights for ages 65 and older. The resulting quotients are the standard population weights used in this *Atlas*.

### **2000 U.S. Standard Population Weights**

Age Group (yrs)	Population	Weight
<1	3,795,000	0.013818
1–4	15,192,000	0.055317
5–14	39,977,000	0.145565
15–24	38,077,000	0.138646
25–34	37,233,000	0.135573
35–44	44,659,000	0.162613
45–54	37,030,000	0.134834
55–64	30,531,000	0.111170
65–74	18,136,000	0.066037
75–84	12,315,000	0.044841
≥85	4,259,000	0.015508
Total	274,634,000	1.000000

Source: National Center for Health Statistics, CDC, 2001<sup>10</sup>

### **2000 U.S. Standard Population Weights for Age Groups ≥65 Years**

Age Group (yrs)	Population	Weight
65–74	18,136,000	0.522501
75–84	12,315,000	0.354797
≥85	4,259,000	0.122702
Total ≥65	34,710,000	1.000000

## Calculation of Spatially Smoothed Stroke Hospitalization Percentages

### ***Percentages of Stroke Hospitalizations with Selected Discharge Outcomes***

We calculated spatially smoothed percentages of stroke hospitalizations with selected discharge outcomes at the county level by race/ethnicity for each of the following discharge outcomes: home, skilled nursing facility, other care facility, any care facility, died before discharge, and died within 30 days of stroke hospitalization.

To calculate spatially smoothed percentages, we first summed the number of stroke hospitalizations for each discharge outcome (the numerator) in each county during 1995–2002 to create summary numerators for the 8-year study period. Likewise, we summed the number of stroke hospitalizations (the denominator) in each county during 1995–2002 to create summary denominators for the 8-year study period.

Next, for each discharge outcome, we combined the summary numerator of each index county with the summary numerators of its neighboring counties, and then divided this number by 8. We also combined the summary denominator of each index county with the summary denominators of its neighboring counties, and then divided this number by 8. Finally, we divided each newly calculated numerator by each newly calculated denominator to produce average annual spatially smoothed estimates of stroke hospitalization percentages at the county level for each discharge outcome by race/ethnicity.

Three constraints were applied to the calculation of county stroke percentages for each discharge outcome. For each racial and ethnic group, a stroke hospitalization percentage was not calculated for any county for which the total number of race-specific stroke hospitalizations in that county and its neighboring counties was fewer than 20 during the study period (1995–2002).<sup>9</sup> To avoid calculating percentages for counties that had no members of a particular racial/ethnic group but whose neighbors did have significant population sizes, race/ethnicity-specific percentages were not calculated for a county with fewer than 10 Medicare beneficiaries of a particular racial/ethnic group during 1995–2002. Furthermore, county percentages with a relative standard error  $\geq 30\%$  were considered unstable<sup>11</sup> and were categorized as “insufficient data.”

### ***Percentages of Stroke Hospitalizations with Selected Comorbidities***

We calculated spatially smoothed percentages of stroke hospitalizations with selected comorbidities at the county level by race/ethnicity for each of the following comorbidities: hypertension, diabetes, and atrial fibrillation.

To calculate spatially smoothed percentages, we first summed the number of stroke hospitalizations for each comorbidity (the numerator) in each county during 1995–2002 to create summary numerators for the 8-year study period. Likewise, we summed the number of stroke hospitalizations (the denominator) in each county during 1995–2002 to create summary denominators for the 8-year study period. Next, for each comorbidity, we combined the summary

numerator of each index county with the summary numerators of its neighboring counties, and then divided this number by 8. We also combined the summary denominator of each index county with the summary denominators of its neighboring counties, and then divided this number by 8. Finally, we divided each newly calculated numerator by each newly calculated denominator to produce average annual spatially smoothed estimates of stroke hospitalization percentages at the county level for each comorbidity by race/ethnicity.

Three constraints were applied to the calculation of county stroke percentages with selected comorbidities. For each racial and ethnic group, a stroke hospitalization percentage was not calculated for any county for which the total number of race-specific stroke hospitalizations in that county and its neighboring counties was fewer than 20 during the study period (1995–2002).<sup>9</sup> To avoid calculating percentages for counties that had no members of a particular racial/ethnic group but whose neighbors did have significant population sizes, race/ethnicity-specific percentages were not calculated for a county with fewer than 10 Medicare beneficiaries of a particular racial/ethnic group during 1995–2002. Furthermore, county percentages with a relative standard error  $\geq 30\%$  were considered unstable<sup>11</sup> and were categorized as “insufficient data.”

### ***Contiguity Matrix***

We used the contiguity matrix for all U.S. counties from the Area Resource File<sup>2</sup> database to spatially smooth stroke hospitalization data. Counties are

considered contiguous by water rights to other counties when they both border the same body of water.

Islands and peninsulas are considered contiguous to neighboring counties on the basis of shared water rights or accessibility.

## **Geographic Information**

### ***Spatial Geometry***

The geographic data used for the county-level maps in this publication came from the Environmental Systems Research Institute’s (ESRI) ArcUSA database, which includes the spatial geometry and characteristics of all U.S. counties. ESRI produced this database by updating 1973 digital line-graph data produced by the U.S. Geological Survey to reflect changes in county boundaries through 1988. The maps were drawn on a scale of 1:2,000,000, which is sufficiently detailed to identify major county features. Hospitalization data were linked to county-level jurisdiction on the basis of the modified Federal Information Processing Standard (FIPS) codes (see pages 183–184).

### ***Map Projections***

We used several different map projections to produce the county-level maps in this publication. For the contiguous United States, we used an Albers conic equal area projection; for Alaska, Puerto Rico, and the U.S. Virgin Islands, we used a Miller’s Cylindrical projection; and for Hawaii, we used geographic coordinates (latitude and longitude). Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands are not in proper geographic scale relative to

the continental United States on the national maps. By using a combination of different projections and scales, we were able to present a relatively familiar picture of these regions.

The coordinate information for the contiguous 48 states was derived from an Albers conic equal area projection with the following parameters:

Spheroid: Clarke 1866      Central Meridian: -96.0  
 1<sup>st</sup> Standard Parallel: 29.5      2<sup>nd</sup> Standard Parallel: 45.0  
 False Easting: 0.0      False Northing: 0.0  
 Reference Latitude: 37.5

The coordinate information for Alaska, Puerto Rico, and the U.S. Virgin Islands was projected using the Miller’s cylindrical projection with the following parameters:

Spheroid: Sphere  
 Central Meridian: -121.8497772217

### County FIPS Code Modifications

We used Federal Information Processing Standard (FIPS)<sup>12</sup> codes to link county-level data from multiple data sets. However, to ensure that data from these multiple data sets were linked accurately, we had to modify the FIPS codes by “incorporating” cities and some entire counties into adjacent counties. The total number of counties included in this *Atlas* is 3,187.

### Alaska

Original County	Original County FIPS Code	Incorporated into Adjacent County	Modified FIPS Code
Aleutians East	2013	Aleutian Islands	2010
Aleutians West	2016	Aleutian Islands	2010
Skagway-Hoonah-Angoon	2232	Skagway-Yakutat-Angoon	2231
Yakutat	2282	Skagway-Yakutat-Angoon	2231

### Florida

Original County	Original County FIPS Code	Incorporated into Adjacent County	Modified FIPS Code
Dade	12025	Miami-Dade	12086

### Hawaii

Original County	Original County FIPS Code	Incorporated into Adjacent County	Modified FIPS Code
Kalawao	15005	Maui	15009

### Montana

Original County	Original County FIPS Code	Incorporated into Adjacent County	Modified FIPS Code
Yellowstone National Park (Part), Montana	30113	Park	30067

## Virginia

Independent City	Independent City FIPS Code	Incorporated into Adjacent County	Modified FIPS Code
Bedford	51515	Bedford	51019
Bristol	51520	Washington	51119
Buena Vista	51530	Rockbridge	51163
Charlottesville	51540	Albemarle	51003
Clifton Forge	51560	Alleghany	51005
Colonial Heights	51570	Chesterfield	51041
Covington	51580	Alleghany	51005
Danville	51590	Pittsylvania	51143
Emporia	51595	Greensville	51081
Fairfax	51600	Fairfax	51059
Falls Church	51610	Fairfax	51059
Franklin	51620	Southampton	51175
Fredericksburg	51630	Spotsylvania	51177
Galax	51640	Grayson	51077
Harrisonburg	51660	Rockingham	51165
Hopewell	51670	Prince George	51149
Lexington	51678	Rockbridge	51163
Lynchburg	51680	Campbell	51031
Manassas	51683	Prince William	51153
Manassas Park	51685	Prince William	51153
Martinsville	51690	Henry	51089
Norton	51720	Wise	51195
Petersburg	51730	Dinwiddie	51053
Poquoson	51735	York	51199
Radford	51750	Montgomery	51121
Richmond	51760	Henrico	51087
Roanoke	51770	Roanoke	51161
Salem	51775	Roanoke	51161
South Boston	51780	Halifax	51083
Staunton	51790	Augusta	51015
Suffolk	51800	Suffolk	51123
Waynesboro	51820	Augusta	51015
Williamsburg	51830	James City	51095
Winchester	51840	Frederick	51069

## References

1. U.S. Department of Health and Human Services. *International Classification of Diseases, 9th Revision, Clinical Modification*. Washington, DC: U.S. Public Health Service, Health Care Financing Administration; 1980. DHHS Publication No. (PHS) 80-1260.
2. Bureau of Health Professions. Area Resource File. Rockville, MD: U.S. Department of Health and Human Services, Health Resources and Services Administration. Available at <http://www.arfsys.com>.
3. Centers for Medicare & Medicaid Services, Office of Research, Development and Human Services. *Health Care Financing Review, Statistical Supplement*. Baltimore, MD: U.S. Department of Health and Human Services; 2004. Available at <http://www.cms.hhs.gov/review/default.asp>.
4. Arday, SL, Arday DR, Monroe S, Zhang J. HCFA's racial and ethnic data: current accuracy and recent improvements. *Health Care Financing Review* 2000;21(4):107–16.
5. Cooper R. A note on the biological concept of race and its application in epidemiologic research. *American Heart Journal* 1984;108(3 Pt 2):715–22.
6. Kaplan J, Bennett T. Use of race and ethnicity in biomedical publication. *JAMA* 2003;289(20):2709–16.
7. Waller L, Gotway C. *Applied Statistics for Public Health Data*. Hoboken, NJ: John Wiley and Sons, Inc.; 2004.
8. Ricketts TC. Geographic information systems and public health. *Annual Review of Public Health* 2003;24:1–6.
9. Kochanek KD, Murphy SL. Deaths: final data for 2002. *National Vital Statistics Reports* 2004;53(5):111.
10. Klein RJ, Schoenborn CA. *Age-Adjustment Using the 2000 Projected U.S. Population*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; 2001.
11. Klein RJ, Proctor SE, Boudreault MA, Turczyn KM. *Healthy People 2010* criteria for data suppression. *Healthy People 2010 Statistical Notes*, No. 24. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; 2002.
12. National Institute of Standards and Technology. *Federal Information Processing Standards Publication 55-3: Codes for Named Populated Places, Primary County Divisions, and Other Locational Entities of the United States, Puerto Rico, and the Outlying Areas*. Gaithersburg, MD: U.S. Department of Commerce; 1994. Available at <http://www.itl.nist.gov/fipspubs/fip55-3.htm>.

---

Note: Web site addresses of nonfederal organizations are provided solely as a service to our readers. Provision of an address does not constitute an endorsement of this organization by CDC or the federal government, and none should be inferred. CDC is not responsible for the content of other organizations' Web pages.