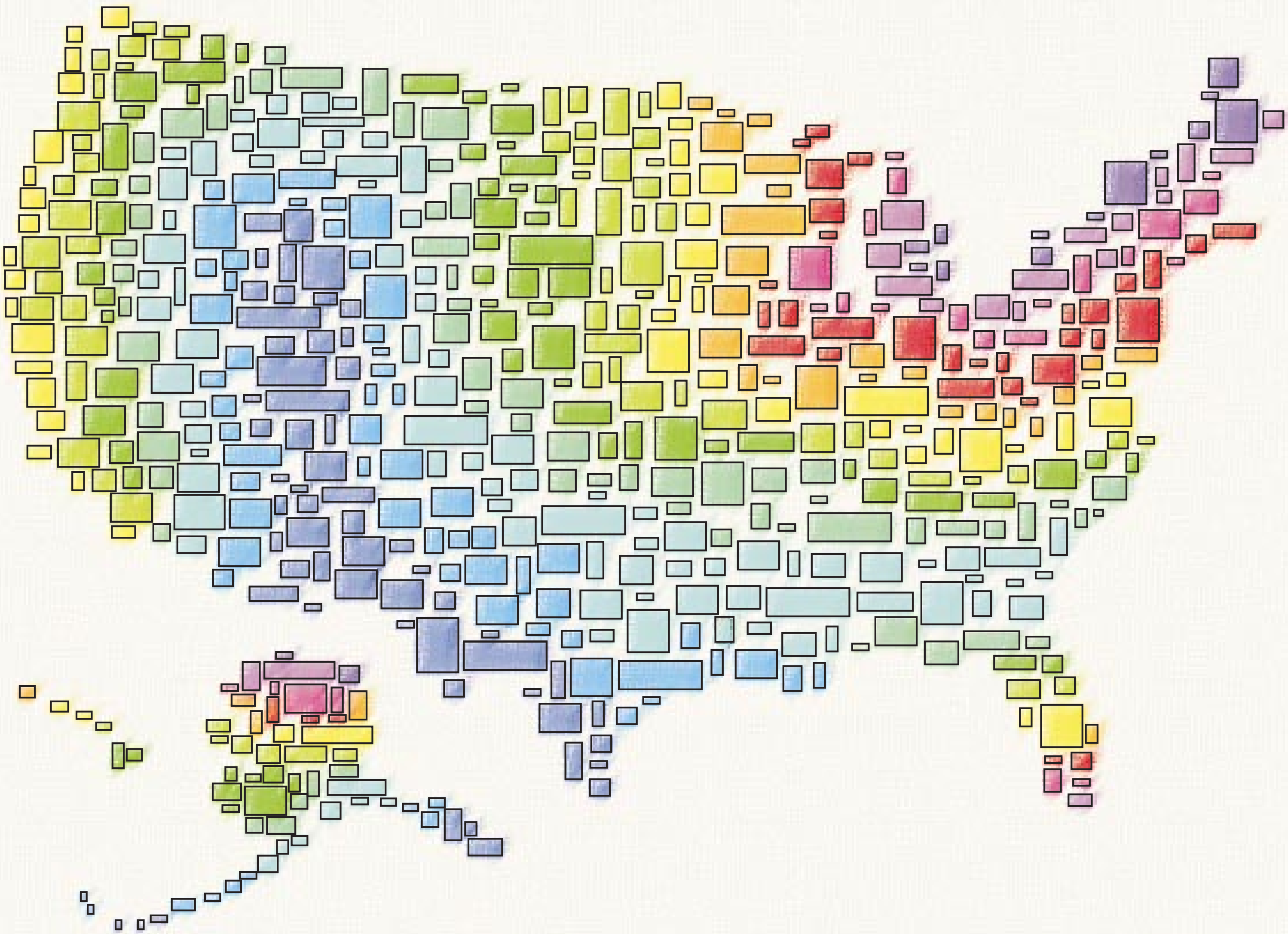
Atlas of United States Mortality

RO T E S SAS A
ATIONAL ENTER FOR EALT STATISTI S



US AM A A UMA S S
ENTERS FOR ISEASE ONTROL AND RE ENTION
ATIONAL ENTER FOR EALT STATISTI S

CENTERS FOR DISEASE CONTROL AND PREVENTION

Atlas of United States Mortality

Linda Williams Pickle
Michael Mungiole
Gretchen K. Jones
Andrew A. White

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
December 1996

DHHS Publication No. (PHS) 97-1015

National Center for Health Statistics

Suggested citation
Pickle LW, Mungiole M, Jones GK, White AA. Atlas of United States mortality. Hyattsville, Maryland: National Center for Health Statistics. 1996.

Library of Congress Cataloging–in–Publication Data
Atlas of United States mortality / Linda Williams Pickle ... [et al.].
p. cm.
Includes bibliographical references.
ISBN 0-8406-0521-8
1. Mortality–United States–Maps. 2. United States–Statistics, Vital–Maps. I. Pickle, Linda Williams. II. National Center for Health Statistics (U.S.)
[DNLM: 1. Mortality–United States–maps. WA 900 AA1 A881 1996]
G1201.E24 A8 1996 <G&M>
304.6'4'0973022–DC21 96-45521
CIP
MAPS

Acknowledgments

No project of this magnitude can be accomplished without the assistance of many others. The authors wish to thank the following people who contributed greatly to making this atlas a reality:

Manning Feinleib, who as former Director of the National Center for Health Statistics (NCHS), had the vision and made the decision to produce an NCHS atlas;

Charles M. Croner, who laid the foundation for the project during its early years and served as a valuable link to the professional geographer community;

Monroe Sirken, who gave the project a high priority on a crowded research agenda and encouraged us to examine the cognitive aspects of statistical mapping;

Harry M. Rosenberg and Jeffrey D. Maurer, who provided valuable feedback throughout the course of this project and critiqued prototype map designs;

Douglas Herrmann, who helped lead the Cognitive Aspects of Graphs and Maps working group and provided valuable insight into the cognitive interpretation of mapped information;

Michelle L. Canham, Association of Schools of Public Health Intern, who conducted the literature reviews of the causes of death included in this atlas and summarized this background material for the initial drafts of the “Results” section;

Lester R. Curtin, for valuable discussions concerning statistical models and demography and many helpful suggestions during the peer review process;

Russell D. Wolfinger, SAS Institute, Inc., who provided valuable technical assistance in the modeling effort;

Joseph Sedransk, Case Western Reserve University, for insightful discussions concerning parametric estimation methods;

Jimmie D. Givens, who ensured that the authors had sufficient computer resources for this project;

Diane M. Makuc and Deborah D. Ingram, who provided the Health Service Area definitions and chose names for Health Service Areas for this atlas project;

Barbara Foley Wilson, for leading focus group discussions and experiments of statistical map reading;

Karen R. Whitaker, for helping with recruitment and preparation of test materials for in-house cognitive experiments;

NCHS staff, who served as technical reviewers and advisors: Ronette R. Briefel; Charles M. Croner; Lester R. Curtin; Lois A. Fingerhut; Katherine Flegal; Mary Anne Freedman; Richard F. Gillum; Deborah D. Ingram; Lillian M. Ingster; Felicia LeClere; Diane M. Makuc; Jeffrey D. Maurer; Margaret A. McDowell; Geraldine M. McQuillan; Harry M. Rosenberg; Diane K. Wagener; Ronald W. Wilson; and Karen A. Johnson, National Institutes of Health;

Numerous other NCHS staff members, who participated in focus group discussions and experiments of the cognitive aspects of statistical map interpretation;

John T. Behrens, Arizona State University; Cynthia A. Brewer, The Pennsylvania State University; Melody Carswell, University of Kentucky; Daniel B. Carr, George Mason University; Reid Hastie, University of Colorado; Stephan Lewandowsky, University of Western Australia; Alan M. MacEachren, The Pennsylvania State University; and Katherine Hansen Simonson, Sandia National Laboratories, whose collaborative research with NCHS, specifically referenced in the “Graphical design” and “Statistical methods” sections, enabled us to add innovative design and analytic elements to this atlas;

Mark Wherley and William P. Vancura of the Deasy GeoGraphics Laboratory, Department of Geography, The Pennsylvania State University, who processed and enhanced the draft map files to produce the high quality images included in this publication; and finally,

Patricia A. Vaughan, visual information specialist, for her many creative ideas and dedication during the design and production phases for this project and Patricia Keaton-Williams, editor, for reviewing the text and graphics.

Contents

List of tables

List of figures

Introduction

Maps have played a fundamental role in public health since the mid-1800's. Soon after a call for studying the geographic patterns of disease (1), Dr. John Snow linked the London cholera epidemic to a contaminated water supply (2). For over a hundred years afterward, however, the usefulness of mapping health outcomes in the United States was limited to either detailed views of a single area or national maps at a State or regional level. Then in 1975, when computer systems had become sufficiently powerful to automate the mapping process, the National Cancer Institute published maps of U.S. cancer death rates at the small-area level (3). Previously unnoticed clusters of high-rate counties on these maps led to numerous field studies, which uncovered, for example, the links between shipyard asbestos exposure and lung cancer (4) and snuff dipping and oral cancer (5). This first atlas demonstrated that mapping small-area death rates could be a valuable public health tool by generating etiologic hypotheses and identifying high-rate communities where intervention efforts might be warranted. Its publication was followed by others from the National Cancer Institute (6–9) and instigated similar efforts around the world (10). Following the success of these atlases in advancing the understanding of cancer etiology, this monograph presents maps of the leading causes of death in the United States for the period 1988–92.

The research underlying this project has led to improved statistical methods for modeling death rates and innovative presentation formats for maps and graphics based on cognitive research (11). In this atlas, information previously available only in tabular form or summarized on a single map is presented on multiple maps and graphs. Broad geographic patterns by age group are highlighted by application of a new smoothing algorithm, and the geographic unit for mapping is defined on the basis of patterns of health care. These new features allow the public health researcher to examine the data at several geographic levels – to read an approximate rate for an area, to discern clusters of similar-rate areas, to visualize broad geographic patterns, and to compare regional rates. With these additional tools, important geographic patterns of cause-specific mortality in the United States can more easily be identified.

Although many causes of death included in this atlas have been mapped before, previous efforts have focused on a limited range of causes (3, 6–9, 12) or have presented data only at the State level (13). This is the first publication of maps of all leading causes of death in the United States on a small-area scale. Comparisons of map patterns across causes of death, sex, or race can provide clues to disease etiology. For this reason unlike many earlier atlases, separate maps by sex and race are included in the same volume, using consistent methods of presentation.

Data sources

The death rates mapped in this volume were calculated from information recorded on all U.S. death certificates to residents of the 50 States and the District of Columbia for 1988–92 and from population data for 1990.

Mortality. Numbers of deaths by age, race, sex, place of residence, and cause of death are based on original death certificates reported to the National Center for Health Statistics (NCHS) from the States. Death certificates with age not stated were excluded (0.025 percent of total). Race was classified following standard procedures for all United States vital statistics (14). Hispanic origin is classified separately from race. Hispanics are included in the data mapped in this atlas according to their race (white or black) as reported on the death certificate; but Hispanics with no racial designation are included in the "White" category. Deaths of persons of races classified as other than white or black were not mapped. Further details on the methods of data collection and processing of death certificates may be found in the Technical Appendix of *Vital Statistics of the United States, 1990, Volume II, Mortality, Part A* (14).

Population. Population counts from the 1990 Census (15), classified by age, race, sex, and county, were multiplied by 5 to create a denominator corresponding to the 5 years of mortality data. Rates computed using such a "person years at risk" denominator are often termed "average annual" rates. In the few instances where the calculated number of person years was less than the reported number of deaths, as when deaths occurred in a sparsely populated county before census enumeration, the person years at risk were inflated to equal the total number of deaths due to any cause.

Causes of death

Mapped causes. The underlying causes of death were initially coded according to the *Ninth Revision, International Classification of Diseases* (ICD–9)(16) then aggregated according to the "List of 72 Selected Causes of Death," which is used in NCHS publications of tabular mortality statistics. This atlas includes maps of rates for the top ranking 11 causes of death from this 72-cause list (14), based on numbers of deaths in 1988–92, as well as the four leading cancer sites, motor vehicle injuries, and suicide and homicide by firearms—for a total of 18 causes of death. Specific definitions of the mapped causes of death are provided in table 1. Total mortality rates due to any cause of death are also mapped. Unintentional injuries, homicides, and suicides are referred to as "external" causes of death. The other causes are referred to as "natural" causes. Death rates for a number of these causes are being used to monitor the health status of Americans at the State and national levels (17).

Quality of data. The issue of accuracy of cause-of-death statistics is fundamental to the interpretation of patterns shown on these maps (18–20). Because this accuracy has been challenged with regard to previously published mortality atlases, what follows is a discussion of potential sources of error in death certificate processing and reporting, and the means by which NCHS sought to compensate for these errors.

The quality of cause-of-death determination in the United States is affected by the accuracy and completeness of information—from medical diagnosis to final coding and processing of underlying cause of death. Beginning with mortality data for 1968, the underlying cause of death has been determined by an NCHS computerized system that consistently applies World Health Organization coding and selection rules to each death certificate using all conditions reported by the medical certifier (21). This system was last amended to incorporate the classification for Human immunodeficiency virus (HIV) infection, beginning with data year 1987. Automation of this task and cross-verification of medical condition coding has reduced errors in assigning underlying cause from death certificate information to less than 1 percent (20). However, the completeness and accuracy of the information supplied on the certificate and the decedent's medical diagnosis remain as potential sources of error (22).

There are indications that the quality of medical information on the death certificate has improved over time. In particular, there has been a steady reduction of deaths assigned to the residual and nonspecific category of Symptoms, signs, and ill-defined conditions (ICD–9 categories 780–799) from 1.5 percent before 1988 to 1.07 percent in 1992 (14). In addition, the number of medical conditions reported on death certificates has increased suggesting more detailed diagnostic information and greater care in completing the medical certification of death (23). Validation studies suggest that, for most broad categories, the reported underlying cause of death agrees well with hospital records of the decedents (24–26). However, for deaths that occur away from a medical setting—as for an unobserved sudden death—the medical certifier may not have

Table 1 - Causes of death for the NCHS mortality atlas: Definitions and map titles

Cause of death	*Abbreviated map titles*	*ICD–9 category numbers included*	*U.S. age-adjusted death rate, 1988–92*[1]			
			White male	*Black male*	*White female*	*Black female*
Diseases of heart	Heart disease	390–398, 402, 404–429	205.0	282.9	104.8	171.7
Malignant neoplasms						
All	All cancer	140–208	160.0	245.0	111.2	136.2
Trachea, bronchus, & lung	Lung cancer	162	56.1	84.5	25.7	26.0
Colon & rectum	Colorectal cancer	153–154, 159.0	16.6	21.1	11.2	15.4
Prostate	Prostate cancer	185	15.0	34.6	No data	No data
Female breast	Breast cancer	174	No data	No data	22.7	27.4
Cerebrovascular diseases	Stroke	430–438	28.1	56.6	23.8	43.4
Unintentional injuries						
All	Unintentional injuries	E800–E949	46.1	64.2	17.7	20.9
Motor vehicle injuries	Motor vehicle injuries	E810–E825	25.6	28.1	10.8	9.1
Chronic obstructive pulmonary diseases	COPD	490–496	27.5	26.4	15.5	11.0
Pneumonia & influenza	Pneumonia & influenza	480–487	17.2	27.9	10.3	13.5
Diabetes mellitus	Diabetes	250	11.1	23.7	9.4	24.9
Suicide						
All	Suicide	E950–E959	19.9	12.4	4.8	2.3
By firearms	Firearm suicide	E955.0–E955.4	12.9	7.8	2.0	1.0
Chronic liver disease & cirrhosis	Liver disease	571	11.7	19.5	4.9	8.4
Human immunodeficiency virus infection[2]	HIV	*042–*044	14.7	46.9	1.1	10.3
Homicide & legal intervention						
All	Homicide	E960–E978	8.7	66.2	2.8	13.2
By firearms	Firearm homicide	E965.0–E965.4, E970	5.7	49.4	1.4	6.6
All causes	All causes	All	648.5	1071.6	373.3	589.7

[1] Rate per 100,000 standard million population (See table 2.).
[2] Beginning with data for 1987, the National Center for Health Statistics introduced categories *042–*044 for classifying and coding human immunodeficiency virus infection. The asterisks before the categories are not part of the *Ninth Revision, International Classification of Diseases*.

sufficient information about the decedent's medical history to correctly report the underlying and contributing causes of death. For example, long-term diabetics are at high risk of heart disease and stroke as a consequence of their disease, but studies have shown substantial underreporting of diabetes on their death certificates (27). Other errors may occur where the cause of death is classified to a related, but incorrect, disorder or to a nonspecific disease category. The latter type of error can be addressed by grouping the related causes that are often confused or by not subsetting the broadly specified disease for analysis.

The potential for errors in assigning underlying cause of death was considered in defining the causes to map. Cause groups were created for this project that were broad enough to avoid these problems, yet specific enough to be meaningful for etiologic research. For example, cancers of the colon and rectum were combined because of the potential for misclassification between these diagnoses (25). All chronic obstructive pulmonary diseases (COPD) (including chronic bronchitis, emphysema, and asthma, ICD–9 categories 490–496) were combined for mapping because approximately 75 percent of all COPD deaths were coded as "Other," with the majority of these coded as "Not otherwise specified."

Amended data. The numbers of deaths that occurred in Alabama, Alaska, Hawaii, and New Jersey for the years 1988–92 are in error, because NCHS did not receive changes to the causes of death made at the State level (14, 28, 29). These differences are concentrated among selected causes of death, primarily the external causes. For example, the largest discrepancy found was for suicides in Alaska — State records indicated 360 suicides during 1988–92, compared to 237 suicides reported to NCHS, a 34-percent deficit. A comparison of annual death rates for 1979–92 due to unintentional injuries, motor vehicle injuries, suicides, and homicides in Alaska shows that rates for 1988–92 are in line with previous years except for suicides, which may be understated in this atlas. The reader is cautioned against overinterpretation of small-area rates for external causes in the previously mentioned States.

Incidence versus mortality. Although public health researchers would prefer to examine patterns of incidence rather than mortality, no nationwide registries exist for noncommunicable diseases. Though some problems surely remain in the mortality data presented here, the experience of National Cancer Institute epidemiologists in successfully following leads generated by cancer mortality atlases demonstrates the utility of mapping these data (30).

Geographic unit

Deaths were initially assigned to a county (or equivalent administrative unit, such as independent city or parish) according to the residence of the deceased, regardless of the place of death. These 3,141 geographic units were then aggregated into Health Service Areas (HSA's) by a cluster analysis of where residents aged 65 years and over obtained routine short-term hospital care in 1988 (31) (appendix I). An HSA may be thought of as an area that is relatively self-contained with respect to hospital care. The original 802 HSA's defined by Makuc et al. (31) were supplemented to include Alaska and Hawaii. Also, several of the original HSA's were combined to achieve a minimum HSA size of 250 square miles for better visibility on the maps. Only New York City remains as a small but populous HSA; its HSA is enlarged for visibility east of its actual location and is labeled "NYC" on the maps. Several other major cities, such as Washington, D.C., were grouped with surrounding counties by the original cluster analysis. The final boundary file includes 805 HSA's.

HSA's are a compromise—in size and number—between the 3,141 counties and 50 States. Data for many of the causes included in this atlas are too sparse to provide stable 5-year rates at the county level, but mapping at the State level would mask many interesting geographic patterns in the data. (In fact, mapping at the HSA or county level may mask interesting local variations in the data. However, in addition to sparse data and confidentiality concerns, most States do not geocode death certificate addresses below the county level.)

Previously published cancer atlases (3, 6, 8, 9) mapped according to county or State Economic Area, aggregations of counties according to demographic and economic conditions in 1960. HSA's, defined on the basis of 1988 health care utilization, are more likely relevant for mapping current death rates and provide a reasonable spatial filter for detecting variations in death rates across the United States.

A map of the HSA boundaries is provided in appendix I, along with a listing of HSA names keyed to the boundary map by number. Each HSA name includes at least one county name and, in some cases, the name of a major city or town. These are provided for convenience in identifying HSA's on the maps. Each HSA that included counties from two

States (77 of 805 HSA's) was assigned to the State where the majority of its population lived. Further details are provided in appendix I.

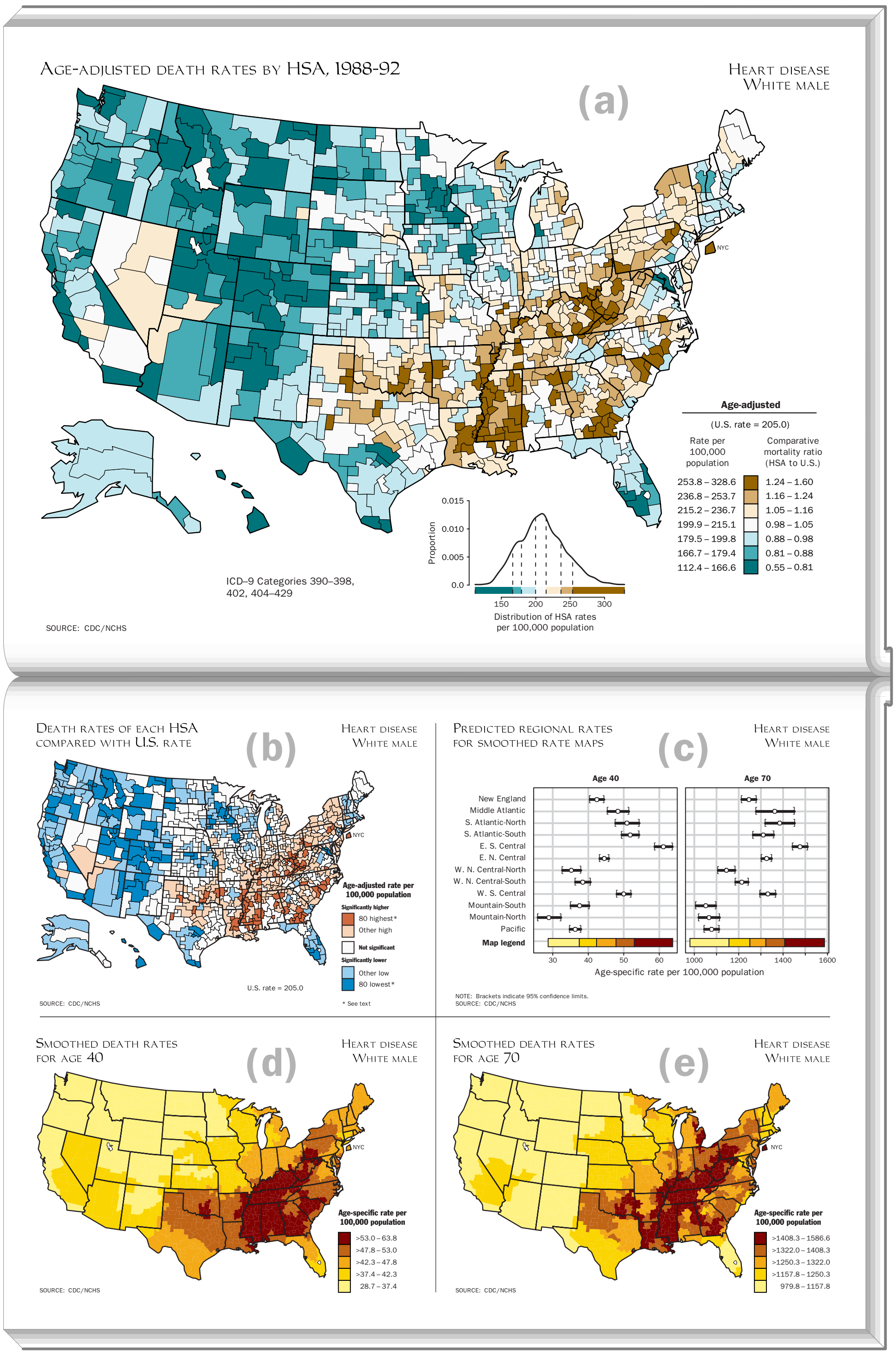
For simplicity of presentation, boundary lines on the base map have been smoothed to within 5 miles of their original location (32). In addition, islands that were combined with a continental HSA to meet minimum size requirements are not shown; deaths among these island residents are included in the rates of the larger continental HSA. All maps were drawn using an Albers equal area projection (33).

This report examines the geographic effects of region as well as HSA. In this atlas, "region" is used in the generic sense and is not to be confused with Census Regions (Northeast, Midwest, South, and West). Fourteen regions for whites and 12 regions for blacks were created by subdividing the nine Census Divisions (appendix I). For whites, the original South Atlantic, West North Central, and Mountain divisions were subdivided so that no region contained more than six States. Because of sparse populations, only the South Atlantic Division was subdivided for blacks. For whites and blacks, Alaska and Hawaii were considered as separate regions, apart from the remainder of the Pacific Division. Note that because an HSA can include counties from two States, the regional boundaries (appendix I) do not strictly follow State lines.

Reader's guide

To aid the reader, the layout of graphical elements on each two-page set is fixed in terms of placement on the page, titles, and colors. It takes only a few minutes to become familiar with this standard page layout (figure 1) and to read the "Graphical design" and "Statistical methods" sections, which explain the components below. The reader who does so will make full use of the integrated graphical presentation.

Figure 1. Graphical components of the two-page atlas layout



Graphical design

One of the unique features of this atlas is the use of cognitive research to guide its design. Early cognitive interviews and focus groups with typical atlas users at NCHS demonstrated a clear effect of a map's graphical design and page layout on the user's understanding of the underlying statistical information (34). Although any map could be used after some study, only easy-to-use maps encouraged repeated use and exploration of the data. To follow up on these early findings, the NCHS Office of Research and Methodology initiated a research effort by an interdisciplinary team of statisticians, psychologists, and geographers to examine how users cognitively process mapped information (11).

Before this research, few empirical studies had been conducted to evaluate the disparate map styles recommended in the literature, and none of these studies considered maps as complex as those in a national small-area mortality atlas. NCHS research revealed that the typical atlas audience of epidemiologists and other public health professionals wanted to (a) read an approximate HSA rate from a map, (b) identify clusters of areas with similar rates and regional patterns on the map, and (c) compare patterns across maps by cause, race, or sex. NCHS conducted collaborative and in-house experiments to examine the effects of basic map style, color scheme, pattern combination, and legend design on the ability of users to perform these specific tasks.

Basic map style. Experiments compared performance and preference using maps where rates were represented by area shading (choropleth), symbols, dot density, and color-coded lines (isopleth). These experiments showed the clear advantage of classed (categorized rate) choropleth maps over competing map styles. A symbol map is not a feasible design for hundreds of small areas, and the map audience was unsure how to interpret lines and dot density for aggregated data mapped to variable-sized geographic units (35, 36). Map style preference differed somewhat by professional discipline (37) although performance did not (38, 39).

Attempts to convey more specific information about the distribution of mapped rates through the use of unclassed maps (for example, where "darkness" of color is proportional to the actual rate) failed. Users also rejected proportional legend designs where the height of the legend box reflects the width of the rate interval. Map readers were either confused by the unfamiliar design (40) or were unable to distinguish among similar shades on the map (35). Therefore, information about the rate distribution is separated from the legend in this atlas and shown instead as a density plot beneath the main map (figure 1a).

Color. Comparisons of color schemes confirmed cartographers' recommendations (33) that distinct hues or patterns facilitated reading a rate from the map but that a sequence of increasing "darkness" of a single hue associated with increasing rates facilitated identification of clusters or broad patterns (36, 41, 42). A double-ended scheme, where each of two hues represents rates above and below the median rate with levels of darkness increasing equally from the middle category to both extremes, can be used accurately for both tasks (43, 44). Therefore, a double-ended scheme was used for the age-adjusted maps (figures 1a, 1b), where identifying the value for a single HSA might be necessary, and a single-ended (monochromatic) scheme was used for the smoothed maps (figures 1d, 1e), where spatial pattern recognition is more important.

Final atlas map colors were chosen to avoid common color vision deficiencies and to balance levels of darkness (or lightness) so that no single color visually dominated a map (44). From a list of acceptable single and paired hues, colors for each of the three types of maps in this atlas were chosen so that no specific color appeared on more than one map and a hue was used consistently wherever it appeared (for example, reds for high rates and blues and greens for low rates).

Hatching. The addition of hatched lines over HSA's was found to accurately convey rate variance information to readers without hampering their ability to identify the underlying colors, and hence the patterns, on the maps (39, 45). Double-hatching with parallel white and black hatch lines allows visibility of the hatching over light and dark colors (figure 1a). Note that the map for white male heart disease (figure 1a) did not require hatching.

Regional rates. In several of the cognitive experiments, map users were asked to compare their estimates of average rates for several regions (35, 41). The variation of responses indicated that this was the most difficult of the questions posed to them. Therefore, to aid in evaluating broad spatial patterns, a rowplot (46) has been included to show confidence limits of model-based regional rates along with each map set (figure 1c).

Page layout. The final composite page layout for this atlas, with its combination of plots and several types of maps (figure 1), may initially seem complex. However, the variety of innovative presentation formats for each set of rates accommodates multiple uses of the maps and different backgrounds of the users. For example, the pattern of age-adjusted rates can be seen on the full page map, and approximate rates can be determined for HSA's (figure 1a). The map in figure 1b indicates where rates are significantly different from the U.S. rate. The smoothed maps illustrate the broad patterns in age-specific rates (figures 1d, 1e), and the graphic (figure 1c) allows comparison of modeled regional effects.

Statistical methods

The statistics mapped in this volume were computed by traditional methods and innovative statistical models. The new models permit examination of age-specific patterns, providing information that may be hidden by use of the traditional summary age-adjusted rate. All rates shown are death rates per 100,000 population. Age-adjusted rates were computed by the direct method (47) using the U.S. standard million population (table 2); these are mapped (figure 1a) and tested for significance compared to the U.S. rate (figure 1b).

Table 2. Standard million population used for age adjustment, proportional to total U.S. population in 1940

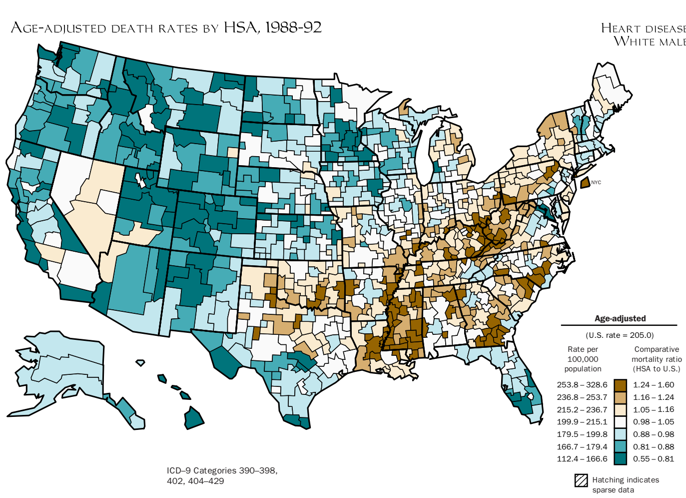
Age	*Standard Population*
0–4 years	80,061
5–14 years	170,355
15–24 years	181,677
25–34 years	162,066
35–44 years	139,237
45–54 years	117,811
55–64 years	80,294
65–74 years	48,426
75–84 years	17,303
85 years and over	2,770
Total	1,000,000

In addition, the age-specific numbers of deaths were modeled for each combination of race, sex, cause, and place using mixed effects generalized linear models (48). Briefly, logarithms of the age-specific rates were modeled as a function of age, allowing each HSA to have a random intercept and, where possible, a random slope, within its particular region. Model results were used to compute improved variance estimates for the age-adjusted rates compared to traditional methods, to estimate regional effects, and to produce smoothed age-specific maps that reflected the broad spatial patterns in the data. Further details of this modeling effort are provided in appendix II.

Statistical methods used for each component of the two-page layout (figure 1) are discussed below.

Age-adjusted death rates by HSA, 1988–92. The age-adjusted rate map (figure 1a) presents the directly adjusted death rates for each HSA. An HSA has an overlaid hatch pattern if its rate has a coefficient of variation at least 23 percent. (The coefficient of variation is defined as the standard error of the rate divided by the rate, then multiplied by 100 and expressed as a percentage.) These rates have a large standard error because they are based on sparse data, typically fewer than 20 deaths, and therefore should be interpreted with caution. Note that the variance used for this calculation is the standard binomial variance estimator for directly age-adjusted rates (49) corrected by the model-based dispersion estimator. Refer to appendix II for details.

The rates are categorized according to percentiles of the rate distribution; the seven categories include, from minimum to maximum rate, respectively, 10 percent, 10 percent, 20 percent, 20 percent, 20 percent, 10 percent, and 10 percent of the 805 rates. These exact distributional percentiles were adjusted for mapping, if necessary, so that the legends accurately list the ranges of rates in each category, rounded to the number of digits shown. For example, a

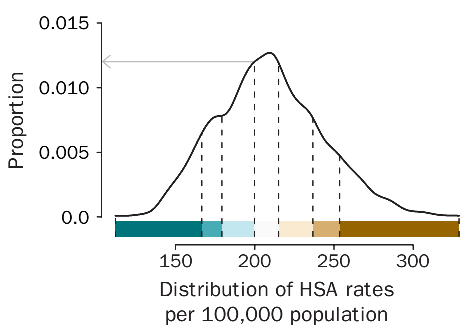


legend range of 5.2 to 10.3 includes all rates from 5.150 to 10.349.

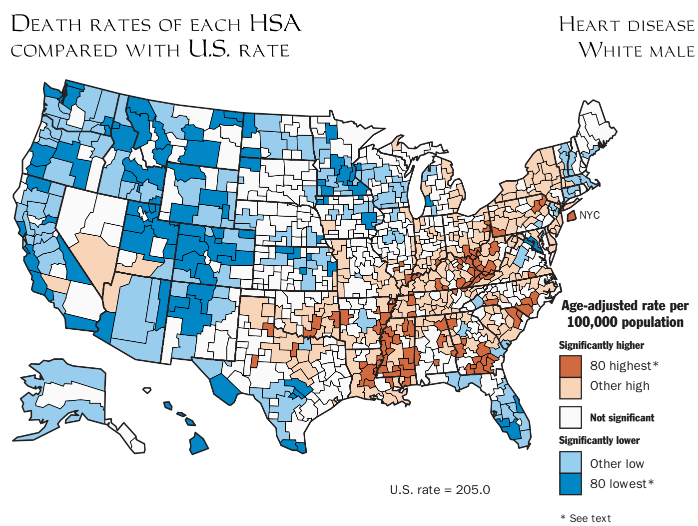
In instances where over 10 percent of the 805 rates are zero (no deaths occurred), all HSA's with zero rates are assigned to the darkest green category. The percentage in the next category is reduced to reflect only nonzero rates up to the next percentile cutpoint. If more than 20 percent (or 40 percent) of the 805 rates are zero, all HSA's with zero rates are assigned to the lowest category as above; but the second (and third, if necessary) lowest category is labeled "No HSA" in the legend to indicate that this color category does not appear on the map. The exact percentage of zero rates is shown in a density plot for each age-adjusted rate map.

The legend also shows ranges for the comparative mortality ratio, defined as the HSA age-adjusted rate divided by the U.S. age-adjusted rate. For example, for an HSA rate range of 100 to 150 per 100,000 population and a U.S. rate of 100 per 100,000 population, the comparative mortality ratio range would be 1.00 (100/100) to 1.50 (150/100), indicating rates that are at least equal to, but no more than 50 percent greater than, the U.S. rate. Although the ratio ranges on the right side of the legend may appear to overlap, this is just the result of rounding after dividing the rate ranges by the constant U.S. rate.

Distribution of HSA death rates. The distribution of HSA death rates is shown graphically below the age-adjusted rate map. This density plot, which may be interpreted as a smoothed histogram, provides the proportion of the 805 rates with a particular value. The area under the curve sums to 1.0. In the example shown below, 1.2 percent (or 10) of the HSA rates are approximately equal to 200. The color bar below this graph shows the correspondence of the mapped rate categories (figure 1a) to the density plot (for example, the endpoints of each segment of the color bar correspond to the cutpoints of the legend categories). For causes with extremely high outliers, the highest category on the plot is truncated at the 99th percentile of the distribution, and the actual maximum rate is indicated above the rightmost color bar. This was done so that the reader could see the shape of the distribution for every cause of death among blacks and for HIV among whites. For causes of death with no observed deaths in some HSA's, the proportion of zero rates is indicated by the height of a vertical line at zero (or an arrow to indicate that this proportion is beyond the scale of the graph).



Death rates of each HSA compared with the U.S. rate. This map (figure 1b) indicates whether each HSA rate is significantly different from the U.S. age-adjusted rate (α=0.05), which is shown below and to the left of the map legend. The significantly high rates are further subdivided into the highest 80 rates and other significantly high rates; significantly low rates are similarly subdivided. Note that the variance used for this hypothesis test is the standard binomial variance estimator for directly age-adjusted rates (49) corrected by the model-based dispersion estimator. Refer to appendix II for details.

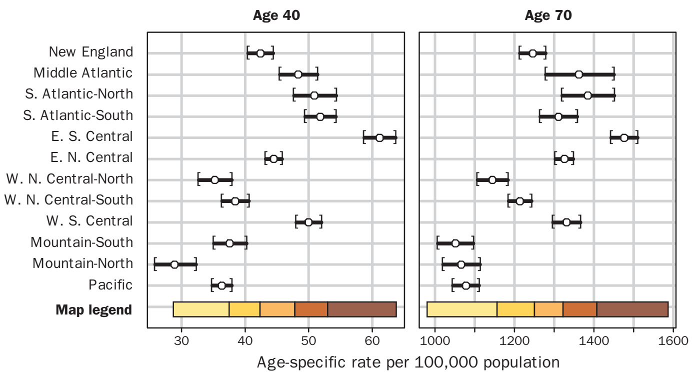


Smoothed rate maps and graphs. The remainder of the second page in each set (figures 1c, 1d, 1e) presents results from the statistical models (see appendix II). The geographic hierarchy included in the models provides estimates of the age-specific rates for each region and HSA. Separate graphs (figure 1c) or maps (figures 1d, 1e) are shown for two representative age groups. Ages 40 and 70 years are shown for natural causes of death, and ages 20 and 70 years are shown for external causes of death, which generally have higher rates for younger adults.

Predicted regional rates for smoothed rate maps. This plot provides the point estimates and 95-percent confidence limits for the predicted age-specific regional rates. As was done for the age-adjusted rate map, a color bar is included reproducing

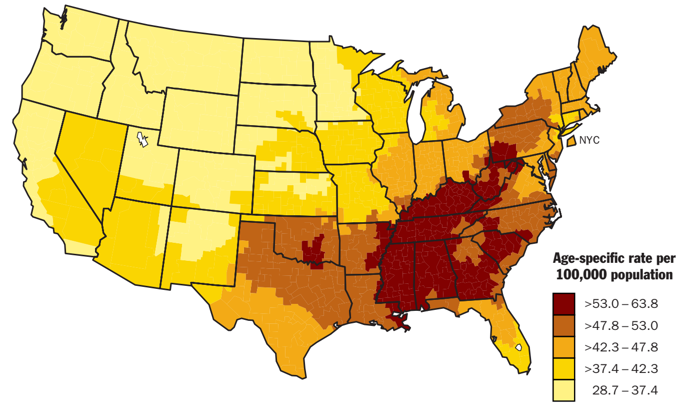
PREDICTED REGIONAL RATES FOR SMOOTHED RATE MAPS

HEART DISEASE WHITE MALE



SMOOTHED DEATH RATES FOR AGE 40

HEART DISEASE WHITE MALE



the legends of the corresponding smoothed age-specific maps (figures 1d, 1e). In instances where the maximum mapped HSA rate (figures 1d, 1e) is over four times the maximum predicted regional rate (figure 1c), this category bar is truncated on the graph and the actual maximum mapped HSA rate is shown inside the color bar.

Smoothed death rates for age 20, 40, or 70 years. These maps illustrate the broad spatial patterns in the age-specific death rates. HSA rates predicted by the models for the two representative ages (ages 40 and 70 years for natural causes and ages 20 and 70 years for external causes) were further smoothed using a two-dimensional weighted median smoothing algorithm and then categorized into quintiles of the rate distribution. Unlike the age-adjusted rate maps, these cutpoints were not adjusted to permit accurate reading of an HSA's rate, because the purpose of these maps is to show broad patterns. Instead, the legend ranges are shown as, for example, >12.2–14.0, where the upper limit of the next lower quintile is 12.2 (rounded to one decimal place). Because the rates are color coded to show the relative ranking of the 805 HSA rates, the reader is cautioned to examine the legends carefully so as not to be misled by the usually great differences in the levels of rates for the younger and older age groups.

The original implementation of the smoothing algorithm (50) has been shown to retain important features of the data pattern better than competitive methods (51). The modification to include inverse standard error weights (52) gives more weight to rates based on large numbers of deaths, so that reliably estimated rates are less likely to be “smoothed out” of the map, even when they differ from rates in surrounding areas. Conversely, rates based on few deaths are more likely to be modified by the algorithm to appear similar to surrounding area rates. Thus, the smoothed map rates may not reflect the observed age-specific rate in a particular HSA. The reader is cautioned that although these maps accurately depict the expected level of age-specific rates in broad areas, they should not be used to estimate a rate for a single HSA.