Introduction to the 2006 Eastern North Carolina Atlas of Mortality

A tombstone in an eastern North Carolina church cemetery is inscribed:

In Memory of
James Bonner Foreman
Who was born
The 1st of December 1785
And died
The 22nd of December 1807
Aged 22 Years and 21 Days

Come view my Tomb as you pass by,
As you are now so once was I;
As I am now so must you be,
Therefore prepare to follow me.

Death is a personal event that we will all eventually experience. It is also something fundamentally empirical, recordable, and therefore measurable. The tradition and culture of recordkeeping varies throughout the world and in the west some countries have been compiling data on peoples' lives for centuries either for ecclesiastical or secular purposes. One extremely important secular purpose is the amassing of individual records over time and place into part of North Carolina's vital statistics collection. Eventually, every North Carolina resident shows up in the vital statistics registry "book" as a single data record, an abstraction, of a life once lived. Unlike Mr. Foreman's epigraph two centuries ago, more data and information pertaining to the circumstances of the mortal event are recorded. In addition to date of birth and death (i.e., age at death), these include the decedent's location at death, cause of death, race, sex, and residence. The data recording the circumstances surrounding people's deaths can be formed into a picture about the conditions of living in their period of time and their society when aggregated at various scales and dimension. The atlas format is an appropriate means of display and description of vital events such as mortality.

The present chapter is an introduction to the approach and concepts used in the current edition of the *Eastern North Carolina Atlas of Mortality.* Specifically addressed topics can be found using the following linked headings.

Overview of the Atlas

The 2007 edition of the *Eastern North Carolina Atlas of Mortality* is a narrated collection of such statistical pictures that describe the spatial and temporal facets—the descriptive geography--of death in the eastern-most 41 counties of North Carolina (ENC). Over the last three decades, this region has seen thousands of individuals dying in excess of what would be expected or experienced in other parts of the country. The underlying motivation for this work is to bring this ongoing tragedy to light and to show health professionals and policy makers where and on what problems need their attention. The information presented in this atlas will allow the reader to form a coherent image in his or her mind of the history and future of mortality in Eastern North Carolina. It is hoped that these statistical images will lead to not only an increased awareness of the conditions of life--and death--in ENC but that it will also stimulate thinking about hypotheses, research questions, policy, and strategies for making life better in our region.

In this work, the geographical distributions of mortality from leading causes are aggregated and portrayed for the years 2000 to 2004 (5 years) and chronicled over a 26-year time series beginning in the year 1979 an ending in 2004. From 2004, rate projections (linear best fit lines) are included. **Figure 1.1** portrays the 100 counties of North Carolina and delineates the major regions used in this Atlas. The regional focus is the eastern-most 41 counties whose western boundary is approximated by I-95 and extends to the coastline. ENC 41 also corresponds to the physiographic province of the Coastal Plain. The 41-county region is further divided into two sub-regions: ENC 29, comprised of the northeastern-most 29 counties of ENC 41, and a remaining southern 12-county region. ENC 29 corresponds spatially to the county service area of University Health Systems of Eastern Carolina. ENC 41 possesses North Carolina's greatest levels of poverty and ethnic diversity, while population and economic growth lags behind the remaining western 59 counties. To contrast and compare mortality rates with the rest of the state, the remaining 59 counties are grouped into two regions corresponding to the Piedmont (PNC) and the western mountain region (WNC). Over the last 30 to 40 years, PNC and WNC have experienced rather different population and economic trajectories than the east and this is reflected in their more favorable mortality outcomes.

The Atlas traces the spatial and temporal domains of ENC's mortality experience with the use of maps, tables, and time series charts. These three components of the Atlas are built on measures that summarize the population's mortality experience. Summary measures like mortality rates are calculated from several of the descriptive elements of the individual death record. The resulting rate calculations are then tabulated by county, region, and time period. In contrast to the simple table, maps are a 2-dimensional spatial ordering of mortality rates that describe a place's mortality experience and burden. Time series charts portray the temporal order of mortality rates for regions, counties, and their constituent population groups. These charts show general parallel, convergent, or divergent trends among regions and population groups. Relative and absolute mortality rate comparisons can be made from the maps, tables, and charts to determine progress toward the elimination of rate disparities and mortality burden over space and time.

Portraying Geographic Data

Maps are the most important feature of a geographical atlas. Along with other graphical means of communication, a wide range of topical literature has evolved that discuss the nature of maps and the geographic information and meaning that they portray from a variety of technical and philosophical both within and without the discipline of geography. A good discussion of the foregoing, which also includes Information Theory, can be found in Poore and Chrisman's *Order from Noise: Toward a Social Theory of Geographic Information* (Poore & Chrisman, 2006). The more salient and general points concerning maps and time series data found in this work are discussed below. For a more technical treatment of charts, with a strong emphasis on the proper construction of graphics that convey meaningful information from quantitative data, the reader is directed to the works of Tufte (Tufte, 1995; Tufte, 1997; Tufte, 2001; Tufte, 2006). Pragmatically, different aspects of various techniques and perspectives necessarily come together in the development of any atlas and how they come together may distinguish one atlas's approach from another. In this Atlas, our approach is one of description and chronicling in such a way that the reader can make meaningful geographical comparisons of the regional mortality experience.

One functional definition of geography considers both space and time as referential systems. Borrowing terminology from Werlen (Werlen, 1993), a space can be defined as a three dimensional container. This type of space orders events (an occurrence or areas with given attributes like mortality rates) by measuring their positional relationships (the *x* and *y* axes) and their sizes or magnitudes (the *z* axis). Another dimension can be added that orders those events temporally and therefore, sequentially. The 2-dimensional or 3-dimensional static map can be stacked or sequenced along a temporal axis to form a time series of maps. As long ago as 1964, Berry (Berry, 1964) described and operationalized a very similar concept as the geographical data matrix, where the matrix is the container of geographically referenced data—attributes/characteristics (or mortality rates) that are linked to places or areas. With some modifications, this prosaic and functional conceptualization describes how spatially referenced data are managed in modern Geographic Information Systems (GISs). With a GIS, these data can be stacked or sequenced in temporal order very quickly to create a moving picture of a geographic process. Because of space constraints, only the most current 5-year maps of mortality rates are provided in this Atlas, but they are accompanied by charts that show temporal trends among regions and population groups.

Geographical referencing and the binding together of attribute data over points in time or sequence of time periods are a means to the comparative study of trends in mortality processes.   In both spatial and temporal referential systems, there is a well-known tendency for objects within the system that are nearer to one another to be more alike than those more distant or, as stated in Waldo Tobler's first law of geography, "…everything is related to everything else, but near things are more related than distant things." (Tobler, 1970)   This notion of propinquity and similarity is important for understanding relationships among demographic, social, biological, and physical attributes of places. For example, a group of neighboring counties such as those found in Eastern North Carolina will tend to have similar age, race, and sex structures because they have had similar economic and demographic histories or, more generally, have experienced similar social relations and processes as well as live within similar spatial structures (Gregory & Urry, 1985). Since age is the greatest risk factor for mortality we would also expect a group of neighboring counties that share a similar age structure to have similar mortality rates. In varying degrees, these same counties may also have similarities in other known risk factors such as certain occupations, race, housing, and poverty. Within the spatial analytical line of inquiry, this well known propensity in geography is extremely useful for constructing hypotheses, modeling, and theory testing.

Maps can be thought of as models of real-world patterns and processes at a given point in time. They reduce reality to a set of graphical and geometric objects that have an *a priori* common meaning, which is necessary for interpretation and communication. This reality is not produced, reproduced, or experienced in exactly the same way by any two persons or reflected in individual death records but collectively similarities and patterns can emerge and be traced for population aggregates. A map as a representation allows a way for the user to apprehend a myriad of facts about places and order them both spatially and temporally into one coherent mental picture. Once geographic data have been integrated into a suitable level of coherency, assessment and analyses can begin with a certain set of well-grounded assumptions. These assumptions might include Tobler's first law of geography (the closer, the more similar) or considered in conjunction with certain risk factors such as age or diabetes with certain mortality outcomes. However, it should always be borne in the mind of the map user or analyst that these newly acquired understandings and cognitive models are ultimately based on a reduced reality—that is, in the time-worn phrase: the map is not the territory.

Finally, maps can be used either as arguments to make a case for further study into the etiology of the causes of mortality and morbidity or they can be used as propositions (or hypotheses) addressing potential causes of observed mortality and morbidity patterns (Koch, 2005). To illustrate, given the range of social and structural inequalities that exist among certain demographic groups in the US and particularly in the South, the Atlas provides evidence for the argument that differences in the underlying social fabric will manifest themselves in the observed patterns of mortality for Whites and Non-whites in eastern North Carolina and for all Eastern North Carolinians versus the rest of the state. The case can be made by employing maps, tables, and charts that permit comparisons among the race-sex groups at county, regional, and national scales. Maps of related demographic and socio-economic variables are either included or referenced in the Atlas as propositions about relationships underlying the observed mortality patterns. As a tool for integrating disparate data, either as argument or proposition, the Atlas can assist in developing research questions for topics on health disparities, health resources, and economic development.

Representational data used in the construction of maps are of two distinct classes. The first data class is made up of a limited set of geometrical objects that are used to represent a large range of real-world features on a map. The most basic of these data is the geometric point that is located on a geometric plane. The point can represent an event, institution, or place, for example. On this same plane an additional point will define a line and a series of lines can represent features such as road networks, stream systems, or social relationships and connections. Three or more points will define a polygon and can represent real-world entities such as counties or urban areas. In some maps polyhedra or solids defined by four or more polygons can be constructed to represent specific types of features. These geometrical representations (or features) have some measurable quality or attribute assigned to them, which provides the basis for making comparisons and discerning patterns.

Points, lines, and polygons can be assigned an attribute, quality, or quantity that describes map features. This second class of data can be partitioned into three categories: nominal, ordinal, and interval/ratio (Earickson & Harlin, 1994). Nominal data refers to the binary presence/absence of a quality or one or more types of a given feature, such as vegetation cover or soil. Ordinal data are ranked in ascending or descending order and can be used to describe a hierarchical system of, for example, health states or levels of care quality measured as poor, fair, good, or excellent. Finally, interval/ratio scale (or metric scale) data measure quantities like mortality rates, dentist to population ratios, or disease prevalence. For interval data the difference between any pair of values is always the same no matter where they are located along the metric scale. There is a small but important distinction when considering either interval or ratio data. Interval data can include values that are less than an arbitrarily defined zero, such as temperature or elevation. However, unlike elevation, one cannot speak of a temperature being twice as cold or hot as another. These data are strictly interval in nature. Ratio data are interval data that can be compared meaningfully. For example, one could make the statement that the mortality rate for female breast cancer in county A is 33% greater than the rate in county B. Interval data can be evaluated as “twice as much,” “half as great,” or as some percent or proportion of one value in relation to another.

Data Sources

The predominant types of data employed in this Atlas are polygons bound or joined to interval/ratio data attributes. Polygons are used to represent counties, which are the basic units of analysis and are the building blocks for larger multi-county regions. County-level polygon data (i.e., boundary files) are obtainable from the geography page of the US Census website. These data are available in several formats and are ready for use with most GIS packages. Because boundary files have unique county identifiers, they are also ready to "join" or link to attribute data.

A wide variety of county-level attribute data are employed in this work. Demographic and socio-economic data can be obtained from the American FactFinder section of the US Census website and the NC State Data Center. In the Atlas, mortality rates by leading causes of death are calculated from two sources. The North Carolina source is located at the University of North Carolina's Odum Institute, which provides the most up to date vital statistics for the state. Mortality data for the nation and other areas of the county are calculated from data found in the Compressed Mortality File (CMF) series produced by the National Center for Health Statistics. These data tend to be 3 to 4 years behind the latest year for North Carolina.

Mapping with GIS Software

Today, nearly all data required for GIS and mapping exist in a digital form. Many printed tabular data sources, collected in more remote periods of time, have been archived either on paper or microforms. These data sources can be scanned or imaged into formats suitable for optical character recognition (OCR) programs or other software tools that will transform the printed character or numeral into a digital rendition. Once obtained, the data need to be stored in some type of database. Storage can be in a large relational enterprise level database such as MS-SQL® or Oracle® with member tables distributed according to function anywhere on the globe or data storage can simply be in a spreadsheet "database" residing on a desktop PC. In Microsoft's Excel®, one or more data ranges (i.e., columns × rows) described in a worksheet can behave as individual database tables within a workbook. These data ranges and tables loosely correspond to Berry's geographic data matrices. (Berry, 1964)

Using a small set of basic database functions in Excel, it is possible to link and match records (table rows) in a way similar to what is done in a true relational database. In order to match records, there must be a field serving as an index. An index field contains rows of unique identifiers and is common to all tables that will be linked or joined. In this Atlas, we use either the unique county name within the state or the Federal Information Processing Standard (FIPS) code that uniquely identifies any county among the more than 3,000 counties in the US. These same identifiers are used to match attribute data to county polygons prior to mapping in a GIS.

Map-making today is largely done using GIS software that integrates a wide variety of disparate data sources and data types. The construction of maps is actually one of many functions a modern GIS can perform. Other functions include spatial querying, spatial analyses, modeling, as well as layering and combining spatial objects and their attribute data to develop new data. For the purposes of descriptive spatial epidemiology and ultimately the comparisons that will be made, the Atlas here employs the primary and more basic functions of a GIS which manage geographically referenced data and quickly generate map layers with accompanying cartographic elements.

Cartographic elements include the legend or map key derived from data and feature classification and symbology. Data in an atlas of mortality are typically rates and percentages (interval/ratio data). A GIS is able to partition and classify a data distribution with a choice of automated default methods (e.g. quantile, equal interval, natural breaks, or statistical) or the user can classify the data manually. The choice of method is based on the purpose of the map (e.g., statistical description, proposition, or argument) and the intended audience of map readers (Wilson & Buescher, 2002). The GIS also provides color palettes for selecting a hue for each theme. A hue can be further divided into a series of graded shades with hue saturation corresponding logically to category ranges. Analysis proceeds by examining the resulting patterns of categorized rates represented as shades: do counties with more saturated shades tend to cluster together? Or are they more dispersed, demonstrating no real comprehensible pattern? Such basic analyses can yield ideas for the development of hypotheses or intervention strategies if something is known about the processes that created them. Different ideas about presentation of map data and experimentation with categories can proceed quickly with a GIS. What took several days to produce by hand as recently as twenty years ago today only takes several minutes. The maps in this Atlas were created in ESRI's ArcGIS 9.1 and 9.2.

Maps in the Atlas

The Atlas is organized in a way that invites the assessment of patterns in both the spatial and temporal domains. Maps show the distribution of categorized county rates of mortality for the years 2000 to 2004. Mortality rates are, in effect, measures of density. They measure the density of events (deaths by selected causes) in relation to the population producing those events. Both crude and age-adjusted rates are employed for making regional comparisons in those maps depicting total deaths by cause, while only age-adjusted rates are used for making county and regional comparison by race-sex groups. Crude rates are constructed by dividing the number of events (or case mortality by cause) in a county by that county's total population, and then multiplying the result by 100,000, which has the effect of reducing in a certain time period the number of decimal places and thereby making the rate more easily understood. A crude rate is the actual rate and is useful for measuring the burden of disease mortality in an area and time period. However, making comparisons among counties with crude rates is problematic because the differences in their respective age-structures can confound interpretation. For example, knowing that increased age is the greatest risk factor for dying in a given time period, a county with a larger proportion of elderly (e.g., retirees) will naturally produce a greater crude rate than a county where there are larger proportions of college-age students or individuals stationed on military bases.

To make meaningful comparisons, a county's age structure (the numbers of people in each previously defined age group) must be adjusted. Essentially this adjustment is a re-weighting of a county's population that produces an expected, as opposed to actual or observed, number of deaths for that population. The weights are based on an external or synthetic population structure known as the *standard million population*. Age-specific death rates based on the weights are calculated for each group in the age structure and then summed to produce an age-adjusted rate (Buescher, 1998). An age-adjusted county rate is the rate a county would have if it had the same age structure as the external or standard million population and renders this county's rate comparable to any other county using the same standard million population. It should be emphasized that age-adjusted rates used in making comparisons are not the actual observed rates but are the rates that would be expected if each county and region had the same age structure. The external population used in this work is the *US Standard Million* for the year 2000. Knowing which standard million population is used is extremely important when comparing rates calculated from mortality data from different states and time periods, otherwise the rates are simply not comparable.

Time Series Charts in the Atlas

Time series graphs for the years 1979 to 2004 provide a synoptic view of mortality trends for regions and race-sex groups. Age-adjusted rates are used to make comparisons among the 41 counties of ENC, the remaining 59 counties of the state (RNC), North Carolina, and the US (1979 to 2002). Time series plots for four ENC race-sex groups (male and female Whites and Nonwhites) are provided on an additional graph. Best-fit lines are incorporated into the time series plots for both regional and population charts so that the user can assess differences and trends. How well the trend and projection line fits the data is described by the coefficient of determination, R^2. (R^2 is a statistic with values 0.0 to 1.0; the closer to 1.0, the better the fit.) For some leading causes of death there are Healthy People 2010 goals, which are age-adjusted target rates for the year 2010 (U.S. Department of Health and Human Services, 2000). Where applicable, target values are included in the chart and can be used in conjunction with the projected trend lines. This permits the user to make comparisons among regions and population groups in terms of the amount of progress that is being made against a nationally recognized standard.

Over the course of many years, mortality rates will ebb and flow with small annual perturbations deviating from the general trend. A larger view over many decades may show gradually decreasing (the ebb) or increasing (the flow) trends for chronic diseases and intermittent spiking for epidemics during that period of time when communicable and infectious diseases were predominant causes of death (see **figure 1.2**). Long term directional changes and pattern shifts in mortality rates are known as secular trends.* These trends are both responding and contributing to the underlying long term shifts in demographic, socio-economic, and environmental processes. One of the best examples is the nearly complete decline of mortality due to infectious diseases in the early part of the twentieth century. Infectious diseases tend to carry off larger proportions of susceptible young as well as those in the older age groups. Socio-economic and environmental processes such as improved access to better food and nutrition, improved sanitation, and generally better living conditions resulted in fewer deaths of the young as a result of contagion. In turn, a gradual shift in demographics occurred: more children survived into adulthood and into later life. This demographic shift—the result of more individuals now surviving into the older age groups-- is a major influence on the rise of the crude mortality rate from cardiovascular disease (with the exception of stroke) in the early-mid twentieth century. These kinds of changes are described in Omran's work on the epidemiologic transition (Omran, 2005).

The long term mortality trends resulting from different causes of death may not all be the same. Generally, mortality rates over the long term trace curvilinear patterns. As these patterns are examined more closely, parts of the curve begin to take on a more linear form. To simplify and give a general snapshot of recent trends, the mortality time series depicted in this Atlas models the data linearly. The benefit to this is that it provides easily understandable summary measures of mortality events occurring over three decades. However, the reader is cautioned to examine the general pattern of the entire series, giving more weight to events that have occurred later in the series than earlier.

The maps, tables, and charts found in the *Eastern North Carolina Atlas of Mortality* form an armamentarium for understanding, integration, and synthesis of the region's mortality burden and experience. Singly, an individual's death, like the one found in an obscure corner of a church cemetery may appear to be a random event. However, when lone events like these are amassed into numerators and then rates, meaningful pictures about the conditions of life in a place can be created. In the end, it should always be kept in mind when gazing upon the abstract representation of the mortality map that ultimately it was the lives rather than the deaths of people that generated the observed patterns.

---

* The term secular as used here refers to a characteristic pattern for a given age or time period in population history. For example, until the First World War in the United States infectious and communicable diseases had a much more prominent role in observed mortality patterns than they do today. The last several decades of the twentieth century has seen a gradual decline for certain chronic diseases like those of the heart and some cancers.

The next chapter addresses general mortality. In this chapter, the leading causes of death for ENC 41 are delineated for the 5-year period, 2000 to 2004. Discussion of the spatial and temporal distributions of mortality from all causes (i.e., general mortality) follows, including a more in-depth treatment of rates and measures in light of the observed data. Subsequent chapters address the 10 leading causes of death for the region and will generally follow the pattern of discussion found in the chapter on general mortality.

References

Berry, B. J. L. (1964). Approaches to regional analysis: A synthesis. *Annals of the Association of American Geographers, 54*(1), 2-11.

Buescher, P. A. (1998). *Age-adjusted death rates*. Raleigh, North Carolina: North Carolina Center for Health Statistics.

Earickson, R., & Harlin, J. M. (1994). *Geographic measurement and quantitative analysis*. New York :Macmillan ; Toronto; New York: Maxwell Macmillan Canada; Maxwell Macmillan International.

Gregory, D., & Urry, J. (1985). *Social relations and spatial structures*. New York: St. Martin's Press.

Koch, T. (2005). *Cartographies of disease : Maps, mapping, and medicine* (1st ed.). Redlands, Calif.: ESRI Press.

OMRAN, A. R. (2005). The epidemiologic transition: A theory of the epidemiology of population change. *Milbank Quarterly, 83*(4), 731-757.

Poore, B. S., & Chrisman, N. R. (2006). Order from noise: Toward a social theory of geographic information. *Annals of the Association of American Geographers, 96*(3), 508-523.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography, 46*(Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods), 234-240.

Tufte, E. R. (2006). *Beautiful evidence*. Cheshire, Conn.: Graphics Press.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, Conn.: Graphics Press.

Tufte, E. R. (1997). *Visual explanations : Images and quantities, evidence and narrative*. Cheshire, Conn.: Graphics Press.

Tufte, E. R. (1995). *Envisioning information* (5th printing, August 1995 ed.). Cheshire, Conn.: Graphics Press.

Werlen, B. (1993). *Society action and space : An alternative human geography* [Gesellschaft, Handlung und Raum.] . London ; New York: Routledge.

Wilson, J. L., & Buescher, P. A. (2002). *Mapping mortality and morbidity rates*. Raleigh, North Carolina: North Carolina Center for Health Statistics.