

EVALUATING AND COMMUNICATING GEOGRAPHIC PATTERNS IN RADIATION ILLNESSES

R. Martin Jones, Ph.D.
Medical University of South Carolina

ABSTRACT

In cases of environmental health where known causes of adverse health effects have existed in the environment, shouldn't we see some corresponding geographic evidence of adverse health effects? The underlying logic of this question has driven public health investigations which lends impetus for exploratory data analysis of geographic patterns of radiation illnesses. An initial research effort has begun to examine innovative mapping programs to enable such a study. The early results are the identification of key mapping programs and sources of data. The initial conclusions are that there exist innovative methods and approaches applicable to such an investigation. Indeed one individual investigator is conducting such investigations. However, no one program is available for use by an outside investigator. The research continues with searches and will begin to focus on mapping of specific radiation illnesses on a national scale.

INTRODUCTION

The intent of this paper is to report on a research effort just begun by the author funded in part under cooperative agreement DE-FC02-96CH10849 between the U.S. Department of Energy (DOE) and the Waste Policy Institute and with funding from the Medical University of South Carolina. The use of collective pronouns in the presentation of the research is intended to acknowledge the input of the many contributors to this effort. It is not, however, intended to represent the views of anyone other than the author. The research is in the literature review stages of progress. Thus, the paper describes the overall rationale for the effort, its scope and focus and the initial findings from the review. Work is continuing on this research and a brief description of the next steps is included. The author welcomes any input to the research on the part of the reader.

RATIONALE

In cases of environmental health where known causes of adverse health effects have existed in the environment, shouldn't we see some corresponding geographic evidence of adverse health effects? This is simple logic regardless of its complexity in science. The arguments abound in the scientific literature. Nevertheless, it is the same underlying rationale that has driven much of the science and investigations of public health since the mid-1800's when John Snow mapped cases of cholera in London (1). The public health rationale historically begins with the adverse health effects and looks for a possible cause. The primary question being, "If we see cases of adverse health effects in a geographic region, could there be a related geographic cause?"

There are numerous cases in public health where the discovery of geographic patterns in health effects data through mapping has lead to the discovery of disease causation and transmission. More recently, there are cases where the discovery of geographic patterns has not resulted in uncovering new causes, but when combined with the biomedical

knowledge of causes, has resulted in health policy changes. (2) The outcome has been the same--to reduce the number of cases of adverse health effects in a particular region or with a particular at-risk population.

Therefore a reasonable objective, in general, for examining health effects data for patterns for *known* causes is to provide evidence in the formation of health policy. If we look and find patterns of health effects that could potentially be explained by the presence of known causes, then we have provided evidence to support pursuing regulation of the cause and exposure. If we look and don't find patterns, then we must question the significance of the environmental contamination as a risk factor in practice.

FRAMING THE INVESTIGATION

The investigation is limited by design to consider health effects where radiation is a known or suspected cause. Secondly, the investigation is focused on examining existing mapping programs with the intent of finding adequate methods and approaches. The term "mapping programs" refers to organizational efforts in the geographical mapping of health effects rather than computer software programs. We have also chosen to limit the investigation to mapping programs that are national in scale producing maps of health effects for at least the 48 contiguous states and the District of Columbia in the United States.

The investigation is focused initially on the *discovery* of patterns in health effects data. John W. Tukey (3) created a distinction between confirmatory data analysis and exploratory data analysis. To describe the distinction, Tukey utilized an analogy to criminal investigation. Exploratory data analysis is the detective work of searching for evidence. Confirmatory data analysis is the work of the judge and jury to evaluate the strength of the evidence found. The job of exploratory data analysis is to examine the data in many ways looking for meaningful patterns while imposing as few assumptions as possible upon the data and the phenomena underlying the data. Deming (4) provides numerous examples of how initially applying sound statistical procedures to data that is not statistically stable can be not only useless, but mislead one to the wrong conclusion about the phenomenon under examination. Therefore, mapping programs that are more exploratory in nature are preferable.

HEALTH EFFECTS DATA

A natural categorization of health effects data exists that is useful to surface in reviewing mapping programs. There exist routine and non-routine health effects data. An example of routine data is the collection of hospital discharge information and the assembly and reporting of this information by hospitals to a state on an annual basis. An example of non-routine data is the compilation of cancer incidence data for people living in counties surrounding government facilities. There are many collections of non-routine data concerning health effects of radiation. Some, such as the development of cancer registries mentioned above, may have pertinent information to support the discovery of particular patterns in health effects. In general, they are not comprehensive nor national in scope. Therefore, we will delay any detailed investigation of mapping programs built upon non-routine data.

This presentation of health effects data is brief. There are many complexities surrounding the details of health effects data. There are issues, for example, of how diseases are diagnosed and coded on forms. There are issues of patient confidentiality. This presentation is intended to give the reader who is not familiar with health effects data some basic terminology and understanding.

Only five root sources have been identified for routine health effects data that are publicly available. They are data compiled from death certificates, birth certificates, hospital discharge records, emergency room discharge records, and Medicare enrollee records. Mortality data are data about death and the causes or conditions of death. Morbidity data are data about the incidence of disease. Data compiled solely from death certificates would be mortality data by definition. Data on patients in hospitals will most likely describe both mortality and morbidity.

The National Center for Health Statistics (NCHS) of the Center for Disease Control and Prevention (CDC) is the national repository for vital statistics data which includes birth certificates and death certificates. Some record of death has been required from the time of the first English settlements. Since 1950, NCHS, or its predecessor agency, has made repeated efforts to improve the reliability and quality of mortality data. (5) Death Certificate data are collected in each state and submitted to NCHS. NCHS has made mortality data available in electronic form for the years 1968 through 1996. In some cases, they have also published electronic forms of mortality covering years as far back as 1950. (6) The data for 1968 through 1988 is published in a by county aggregation. From 1989 through 1996, in order to protect the privacy of individuals, no data is published for a geographic area with a total population less than 100,000. Electronic data from birth certificates (natality data) is published for the same years and follows the same patterns for protection. Mortality data from NCHS represents the most extensive, national data on health effects publicly available in electronic form.

There is no national repository for morbidity data. Within a state, hospitals are routinely required by state health departments to submit information on all hospital discharges. When a patient has been admitted and subsequently discharged from the hospital the hospital must submit a record to the state describing various aspects of the patient's hospital stay. Elements of the records include information about the patient such as age, place of residence, sex, and race. They include information about the primary and secondary diagnoses, the medical procedures performed, the attending physician and the cost and payer of the costs. State health departments will most often reduce the amount of information pertaining to each hospital discharge and create publicly available information that does not reveal private information about the patient, hospital, nor physicians. There is no uniform requirement of information from state-to-state. This makes it difficult to acquire individual state data sets and combine them into a national picture of morbidity.

The Agency for Health Care Policy and Research (ACHRP) initiated the Healthcare Cost and Utilization Project (HCUP) on a national scale. (7) As part of this project, they

conducted the National Inpatient Survey (NIS). Through cooperation with 19 state health departments, they created and implemented a standard reporting of hospital discharge data. Summary level data was provided in a consistent format and coding scheme. The sample is estimated to represent 50% of the hospital discharges in the nation. The NIS has continued and data is available electronically for the years 1988 through 1995. (7) A more detailed, yet uniform, record of discharge data is maintained by the participating states and is available by working directly with each of the states. Depending on the scale of the study involved, it may be relatively straightforward to acquire data from several individual states for more detailed analysis of geographic patterns.

The Health Care Finance Agency (HCFA) collects and publishes data on hospital discharges for Medicare beneficiaries. Whereas this represents a subset of the national population, the morbidity data in the HCFA system could be of value in instances where elderly patients are of primary concern. Certainly comparing the patterns of illness in this data for given age groups with patterns from mortality data could be of interest. Our search of this data is not complete. Early indications are that the data is readily available for the time period from 1990 through 1997. However, the available data may only cover some hospitals representing only part of the total in-patient stays. Also, the Medicare discharge files are coded using Diagnosis-Related Groups (DRGs). Hospital discharge and mortality data are coded using the International Classification of Diseases (ICD) codes. There is not direct translation of DRGs to the more detailed ICD codes.

For the purposes of this initial investigation, the sources of data most likely to reflect both a national scope and illnesses specific to radiation are the mortality data from NCHS, the partial morbidity data from the NIS and the Medicare data. The morbidity data would most likely serve as supporting evidence to any patterns discovered in mortality data. The natality data available does not appear to present enough information to be of value in assessing patterns of radiation illnesses.

KEY MAPPING PROGRAMS IDENTIFIED

Numerous programs of analysis of health effects data exist that use geographic mapping as a tool for communicating results. Only a few have been found that use mapping as an investigative tool on a national basis in general rather than specific to a particular investigation of the occurrence of a disease. The key programs identified so far that have innovation and strengths applicable to the exploratory data analysis objectives of this investigation are presented briefly.

National Center for Health Statistics and the National Cancer Institute

The National Cancer Institute (NCI) had an experimental mapping program in the 1970's that published two atlases of cancer mortality for the years 1950-1969, one for whites (8) and one for non-whites (9). The maps presented data by county for the nation. NCI's mapping program has, in essence, folded in with NCHS's mapping program. Many of the investigators involved in the pioneering NCI effort are now involved in the NCHS effort.

The National Center for Health Statistics (NCHS) has produced three atlases of mortality data. The first two extend the work of the NCI and are also atlases of cancer mortality, but cover the years 1950-1980. The first was published in 1987 reporting mortality of whites (6) with the second published in 1990 reporting non-white mortality (10). The geographic area for reporting statistics was chosen to be state economic areas which are, in general, aggregates of counties. The more recent atlas of mortality published by the NCHS represents significant effort by NCHS in many areas of methodology and presentation. (11) Maps and data are published for the 18 leading causes of death in the U. S. for the period 1988 – 1992. Data is presented by health service areas which, as in the case of state economic areas, are aggregations of counties.

In order to produce a better product, NCHS conducted extensive cognitive research on statistical maps. (12) The results of this research were incorporated in the design of the maps presented in the atlas. This research is worthy of consideration in any mapping program. Only part of the display of information about each leading cause of death is shown in Figure 1. This is the initial map of the basic data for one leading cause. There are other key maps and charts accompanying this map to communicate statistical information about each analysis performed.

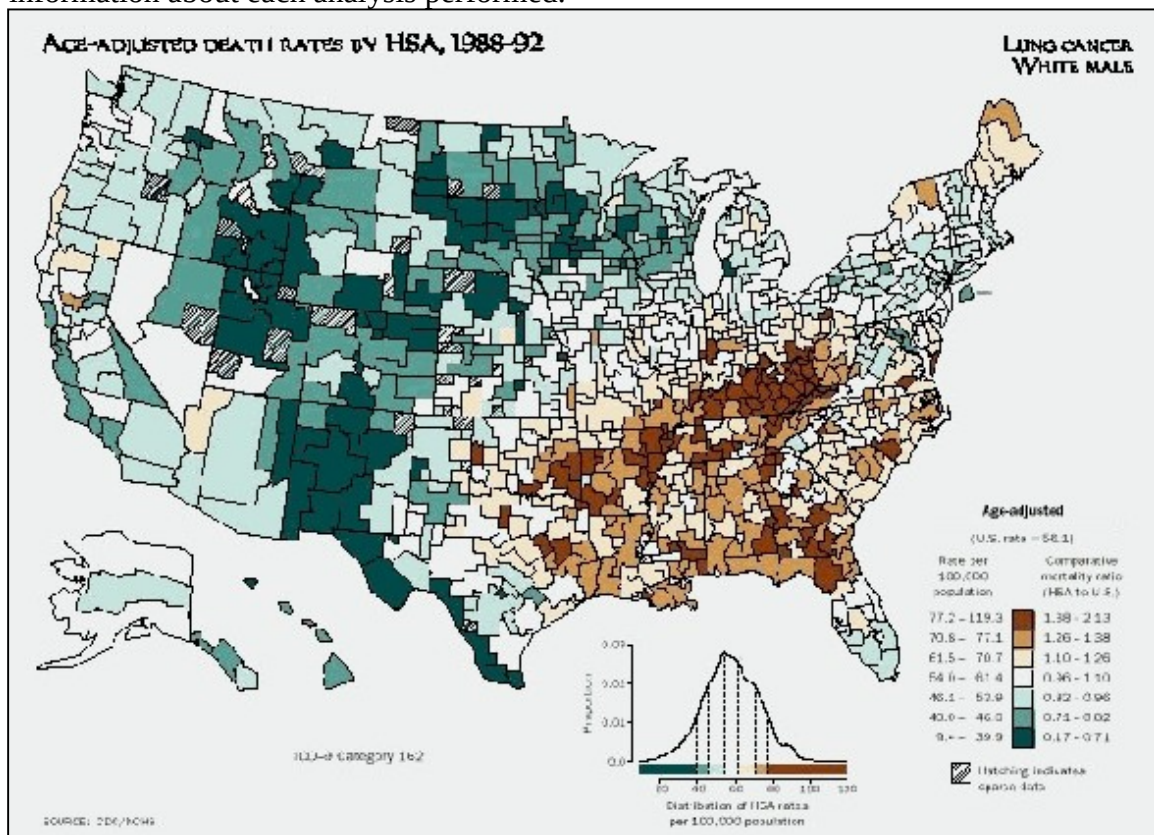


Figure 1. Example of mapping from NCHS Atlas of Mortality 1988-1992(11)

In addition to the cognitive research, NCHS conducted research on statistical methods and has implemented improvements in data quality (11). These are all factors of strength

and innovation to be included in the foundation of any exploratory data analysis effort considered.

RUSSELL A. BROWN

As an individual investigator, Russell A. Brown's work in mapping mortality data presents several innovations of importance. One of Mr. Brown's studies is recently published in a paper examining cases of thyroid cancer in relation to nuclear bomb testing. (13) It is this study and the collection of Mr. Brown's work that forms the impetus for the investigation underway. Mr. Brown's work most closely exemplifies the rationale of this investigation. In an unpublished paper on leukemia mortality (14) and in numerous presentations, Mr. Brown has used mapping of mortality data to explore the significance of hypotheses of radiation induced illnesses. This approach represents an important departure from the more confirmatory studies such as the recent National Institutes of Health study "Cancer in Populations Living Near Nuclear Facilities." (15)

In searching for geographic patterns, Mr. Brown departed from the early NCI and NCHS mapping programs by employing a volume preserving polygon-to-polygon interpolation algorithm (16) to convert mortality data from geopolitical areas to a uniform grid. This innovation in treatment of mortality data enabled him to produce topographical maps of the ratios of observed mortality rates to expected mortality rates. The topographical maps more easily portrayed geographic patterns of illness than the statistical choropleth maps of the early atlases. Figure 2 shows maps from his unpublished paper on leukemia mortality. NCHS employed a different algorithm to produce maps with a similar visualization in the more recent Atlas of Mortality (11). An example is shown in Figure3.

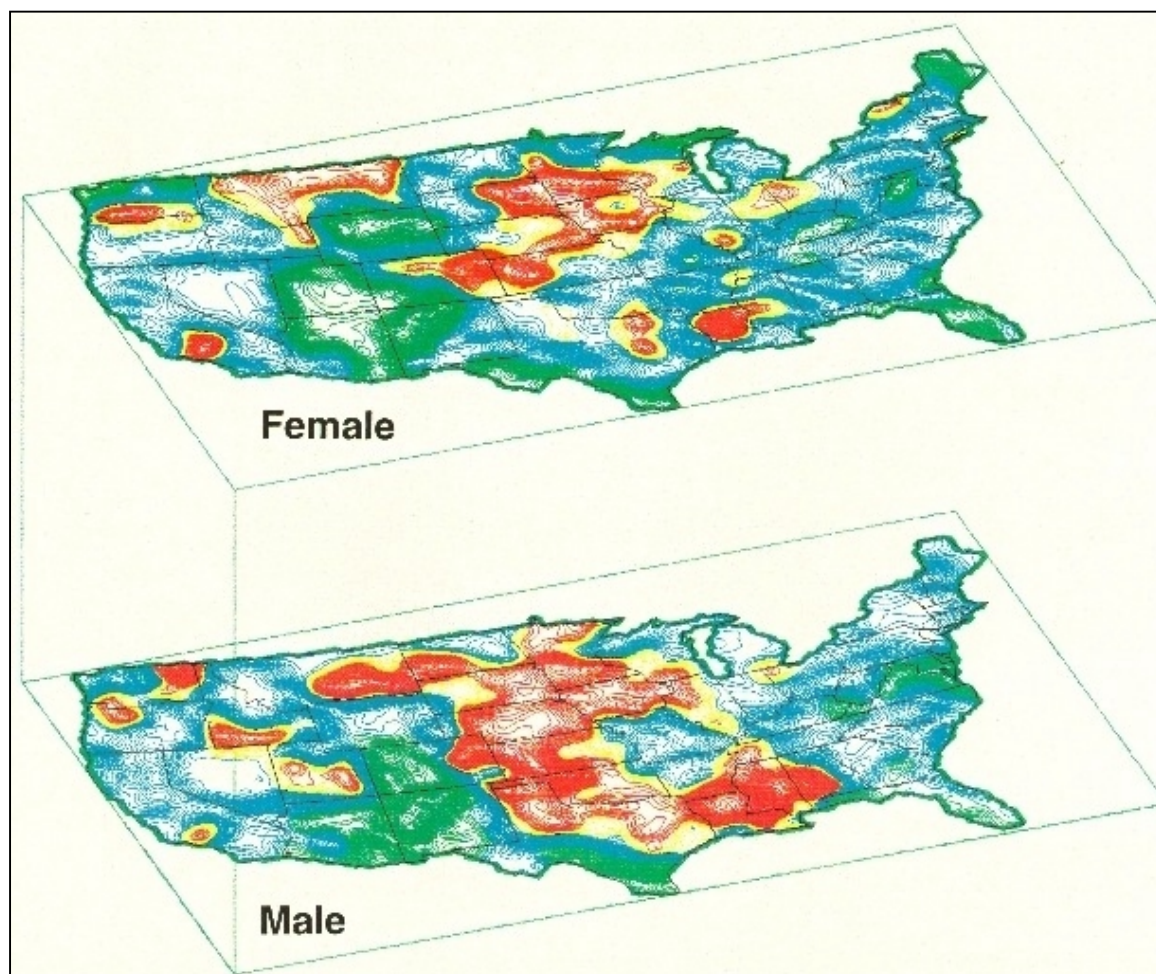


Figure 2. Russell A. Brown's Topographical Maps of Leukemia Mortality for Whites for the years 1950-1980 (14). (Reprinted with permission from the author.)

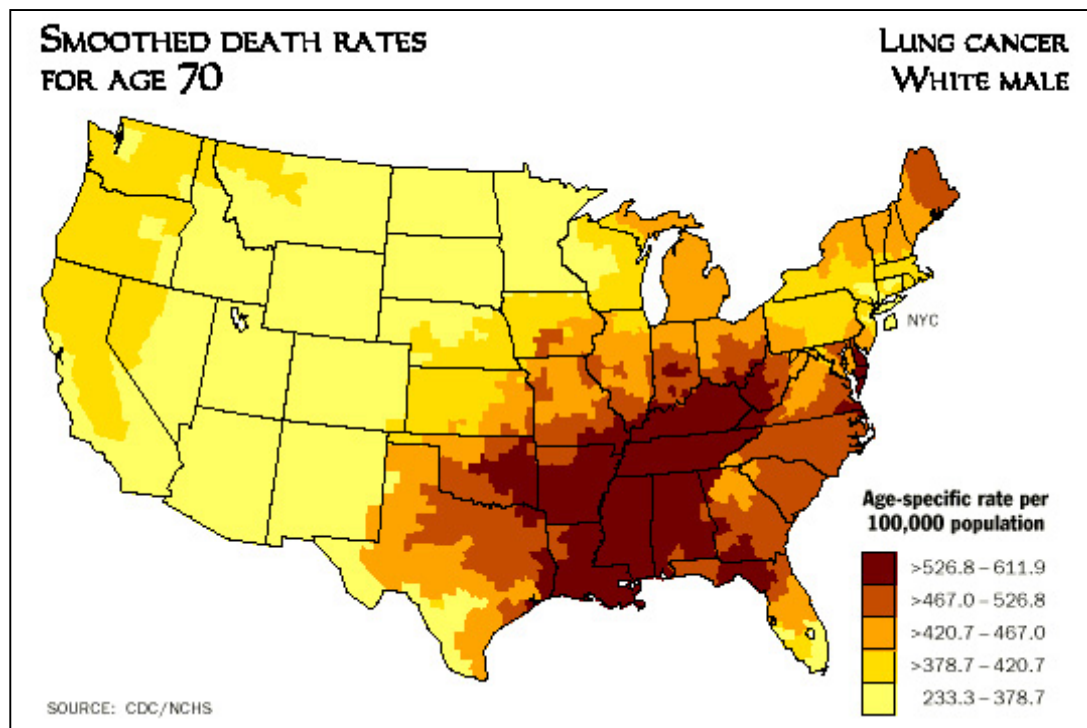


Figure 3. Expanded Image of Smoothed Map from the NCHS Atlas of Mortality (11).

THE DARTMOUTH ATLAS OF HEALTH CARE

The Center for the Evaluative Clinical Sciences of Dartmouth University recently published The Dartmouth Atlas of Health Care 1998 (17) as a follow-up to the first Dartmouth atlas publication (18). The innovation of most interest to this investigation is the form of publication of the mapping program. Dartmouth teamed with the geographic information systems (GIS) software developer ESRI, to publish the data and maps in a mapping package that enables the user to conduct some investigation of the geographic data. The user can select time frames and variables and geographic regions of interest. There are data analysis tools that enable the user to develop algorithms of analysis. This method of packaging empowers the user with some limited exploratory and confirmatory data analysis capability to pursue hypotheses.

There are limited health effects data incorporated in the database for analysis. The morbidity and mortality data comes from the Medicare discharge information provided by HCFA. The purpose of the Atlas is focused in areas that involve the use of some disease incidence data but not as a primary focus. Figure 4 presents the on-line image produced for health referral regions which are aggregates of health service areas.

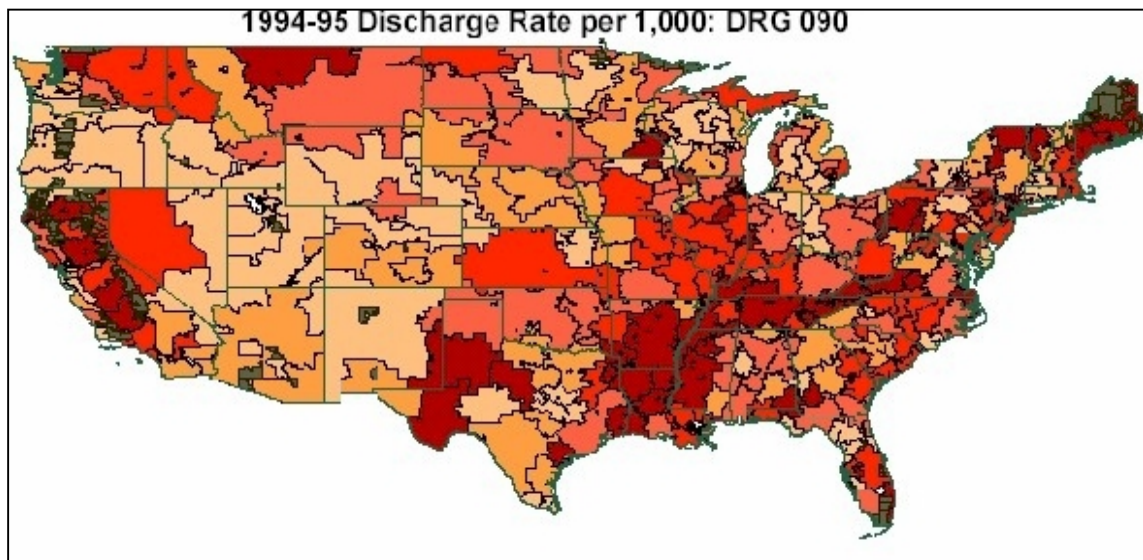


Figure 4. Screen Image from the Dartmouth Atlas Data Viewer Presenting Incidence of Myeloproliferative/Lymphoma/Leukemia Diagnoses Other Than Chemotherapy for Medicare Beneficiaries for the Years 1994-1995. (17)

CONCLUSIONS FROM THE INITIAL INVESTIGATION

Of the mapping programs surveyed, these three represent important innovations and strengths for consideration. Russell A. Brown's work forms the overarching model for the need for examining health effects data in efforts to consider risks from environmental contamination. The significant impacts resulting from the NCI and subsequent NCHS mapping programs signal the potential value of engaging in mapping of health effects on a policy level. The Dartmouth initiative provides a glimpse of the potential to package the results in a manner that enables investigators to conduct exploratory work.

The NCHS mapping effort is limited in that the recent Atlas considered only the 18 leading causes of death. For example, leukemia was not one of these. Therefore, a full exploratory study of radiation illnesses would have to acquire the mortality data and apply NCHS's methods to that data to generate maps. The implementation of Mr. Brown's methods is individually compiled and reasonably available only to Mr. Brown. Therefore, an independent investigator would have to do a similar development and implementation of the computational and mapping algorithms in order to explore the data. The Dartmouth Atlas presents an example of packaging that is directionally appealing. The specific health effects data are too limited to be of significant value in exploring radiation illnesses on a national basis. The tools are also limited to the specific goals and objectives of the Dartmouth Atlas. Further investigation would be needed to determine if a set of exploratory tools could be packaged in a similar fashion. As GIS software improves in usability, it may no longer be necessary to package data in this manner for the investigator, but packaging in this manner may present a useful tool for the practitioner.

NEXT STEPS

There are two basic efforts continuing at this moment. First, we are continuing with the literature search for both data and mapping programs. Secondly, based on the evaluations of those programs identified, we are exploring software approaches and acquiring some test data to evaluate the time and effort required to establish an exploratory data analysis initiative focused on the objectives of the investigation. We expect to collect and/or produce maps of patterns in radiation illnesses from a variety of sources in the Spring of 1999.

REFERENCES

1. Snow, J. (1855). "On the mode of communication of cholera" (2nd ed.). New York: The Commonwealth Fund.
2. Maher, R. J. (1995). A history of the influence of statistical maps on public policy. In L. M. Pickle and D. J. Herrmann (Eds.), *Cognitive aspects of statistical mapping* (Cognitive Methods Staff Working Paper Series, No. 18, pp. 13-18). Hyattsville, MD: National Center for Health Statistics.
3. Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley Publishing Company.
4. Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Center for Advanced Engineering Studies, Massachusetts Institute of Technology.
5. Hetzel, A. M. (1977). "History and organization of the vital statistics system". Hyattsville, MD: National Center for Health Statistics.
6. Pickle, L. W., Mason, T. J., Howard, N., Hoover, R., & Fraumeni, J. F., Jr. (1987). "Atlas of U.S. cancer mortality among whites: 1950-1980" (DHHS Publication No. (NIH) 87-2900). Washington, DC: U.S. Government Printing Office.
7. "Healthcare cost and utilization project (HCUP), 1988-95" [On-line publication]. (1998). Rockville, MD: Agency for Health Care Policy and Research. Available <http://www.ahcpr.gov/data/hcup-pkt.htm>
8. Mason, T. J., McKay, F. W., Hoover, R., Blot, W. J., & Fraumeni, J. F., Jr. (1975). "Atlas of cancer mortality for U.S. counties: 1950-1969" (DHEW Publication No. (NIH) 75-780). Washington, DC: U.S. Government Printing Office.
9. Mason, T. J., McKay, F. W., Hoover, R., Blot, W. J., & Fraumeni, J. F., Jr. (1976). "Atlas of cancer mortality for U.S. nonwhites: 1950-1969" (DHEW Publication No. (NIH) 76-1204). Washington, DC: U.S. Government Printing Office.
10. Pickle, L. W., Mason, T. J., Howard, N., Hoover, R., & Fraumeni, J. F., Jr. (1990). "Atlas of U.S. cancer mortality among nonwhites: 1950-1980" (NIH Publication No. 90-1582). Washington, DC: U.S. Government Printing Office.
11. Pickle, L. W., Mungiole, M., Jones, G. K., & White, A. A. (1996). "Atlas of United States mortality". Hyattsville, MD: National Center for Health Statistics.
12. Pickle, L. M., & Herrmann, D. J. (Eds.). (1995). "Cognitive aspects of statistical mapping" (Cognitive Methods Staff Working Paper Series, No. 18). Hyattsville, MD: National Center for Health Statistics.

13. Brown, R. A. (1997). "Bomb fallout and thyroid cancer—statistical sheep in real wolves' clothing". Focus on Environmental Communications, Miami, Florida: ASME Symposium.
14. Brown, R. A. (1990). "The topography of leukemia in America". Unpublished paper.
15. Jablon, S., Hrubec, Z., Boice, J. D., Jr., & Stone, B. J. (1990). "Cancer in populations living near nuclear facilities" (NIH Publication No. 90-874). Bethesda, MD: National Cancer Institute.
16. Lam, N. S. (1983). "Spatial interpolation methods: a review". The American Cartographer, 10 (2), 129-149.
17. "Dartmouth atlas of health care in the United States (1998 ed.)". (1998). Hanover, NH: Dartmouth College.
18. "Dartmouth atlas of health care in the United States". (1996). Hanover, NH: Dartmouth College.