Chapter X

A Decision Technology System to Advance the Diagnosis and Treatment of Breast Cancer

Guiseppe A. Forgione, Aryya Gangopadhyay and Monica Adya
University of Maryland Baltimore County

INTRODUCTION

Geographical variations in cancer rates have been observed for decades. Described spatial patterns and trends have provided clues for generating hypotheses about the etiology of cancer. For breast cancer, investigators have demonstrated that some variation can be explained by differences in the population distribution of known breast cancer risk factors such as menstrual and reproductive variables (Laden, Spiegelman, and Neas, 1997; Robbins, Bescianini, and Kelsey, 1997; Sturgeon, Schairer, and Gail, 1995). However, regional patterns also may reflect the effects of Workshop on Hormones, Hormone Metabolism, Environment, and Breast Cancer (1995): (a) environmental hazards (such as air and water pollution), (b) demographics and the lifestyle of a mobile population, (c) subgroup susceptibility, (d) changes and advances in medical practice and healthcare management, and (e) other factors. To accurately measure breast cancer risk in individuals and population groups, it is necessary to singly and jointly assess the association between such risk and the hypothesized factors.

Various statistical models will be needed to determine the potential relationships between breast cancer development and estimated exposures to environmental contamination. To apply the models, data must be assembled from a variety of sources, converted into the statistical models’ parameters, and delivered effectively to researchers and policy makers. A Web-enabled decision technology system can be developed to provide the needed functionality.

This chapter will present a conceptual architecture for such a decision technology system. First, there will be a brief overview of a typical geographical analysis. Next, the chapter will present the conceptual Web-based decision technology system and illustrate how the system can assist users in diagnosing and treating breast cancer. The chapter will conclude with an examination of the potential benefits from system use and the implications for breast cancer research and practice.

BACKGROUND

Environmental contaminants tend to be clustered geographically (Chen, 1996; Grimson and Oden, 1996; Oden, Jacquez, and Grimson, 1996). Studies that link breast cancer with environmental factors, then, have focused on geographical analysis.

In a typical geographic analysis, the researcher will utilize clustering techniques and other statistical methodologies to identify abnormal concentrations of breast cancer. Abnormality usually is defined as a concentration that deviates significantly from expected patterns. Sometimes regression and other multivariate statistical analyses will be used to isolate the independent and joint effects of environmental and other variables on breast cancer incidence and mortality. Customarily, the statistical analyses will be based largely on metric data and be single-equation in form. Other times, qualitative evaluations will be used to test the hypothesized relationships. Ordinarily, these evaluations will rely to some extent on the judgment and experience of the researcher and/or other experts.

Researchers and policy makers recognized that information technology could be used to facilitate the desired geographic analysis. Such recognition and subsequent lobbying efforts resulted in legislation that mandated the development of this technology. The National Institutes of Health’s National Cancer Institute (NCI) recently sought a geographic information system (GIS) to support breast cancer studies mandated by public law (Geographic Information System for the Long Island Breast Cancer Study Project, 1998).

Geographic Information System (GIS)

In the GIS sought by the NCI, researchers and other interested parties would utilize computer technology to process available inputs into desired outputs. Inputs would include a database that captures and stores spatial and attribute data for the geographic areas defined by the breast cancer studies. Spatial data, which includes longitude and latitude coordinates that are used to draw on study area maps, can be obtained from state and local government base map files, U.S. Postal Service ZIP Code files, U.S. Geological Survey hydrology data files, and U.S. Bureau of the Census TIGER (Topologically Integrated Geographic Encoding and Referencing) files (Fischer and Nijkamp, 1993). Attribute data, which, among other things, consists of pollution measures, population characteristics, and healthcare provider and patient statistics, can be obtained the U.S. Bureau of the Census, the U. S. Healthcare Financing Administration, the National Health and Nutrition Examination Surveys, state-specific public health data files, and state and federal survey data files on pollution-source emissions.

There is also a model base that contains statistical procedures and location formulas. Some statistical procedures are used to categorize attribute data within the study areas and to calculate summary statistics for the demographic, environmental, healthcare, and other pertinent variables within the areas. Other procedures analyze the data for spatial and temporal patterns and for space-time interactions. A third set of statistical procedures are used to perform selected health-related analyses, including cluster detection of breast cancer incident and mortality, assessment of disease risk from nearby environmental hazards, and genetic activity profiling. Location formulas are used to convert various geographic coordinate systems into alternative forms and to make corresponding map projections.

The decision maker (a scientist or public health official) uses available computer technology to organize the collected spatial and attribute data, structure the study area maps,
Ontology for Data Quality and Chronic Disease Management: A Literature Review
www.igi-global.com/chapter/ontology-for-data-quality-and-chronic-disease-management/115122?camid=4v1a