Constructing a Carcinogen Indicator Application Based on Statistical Analysis of Statewide Cancer Incidences in Relationship with Graphical Consumer Data

Asfari, Sarah
Asfari, Yousuf

Clinical studies that relate different cancers to various carcinogens can be quite costly, and may last anywhere from a couple of years to a few decades. This project is twofold, in that first, it was hypothesized that within a short period of time, via cost effective statistical methods, definite relationships could be deduced between statewide cancer incidences and common lifestyle and environmental factors, many of which are unrecognized as carcinogens. Second, based on conclusions drawn from analysis of the data, a computer application could be constructed which would calculate a user's increased risk of developing various types of cancer. Data for the project was obtained from online cancer-research websites, in addition to numerous web-based graphical representations of national consumer surveys and environmental studies. The data was translated into a spreadsheet, then manifold multiple regressions were performed, and factors that did not appear to significantly contribute to the linear models were gradually eliminated. The result was several linear equations, where each equation relates a precise category of cancer to certain lifestyle and environmental factors. Some of these factors, such as soda consumption and dry cleaning usage, are debated to be carcinogens. A JavaScript based computer program was then created which calculates the consumer's increased risk for acquiring different cancers, based on the user's daily habits and their general environment. This may help the consumer lead an improved way of life, and encourage individuals to petition their local governments to put more stringent rules on various industries.