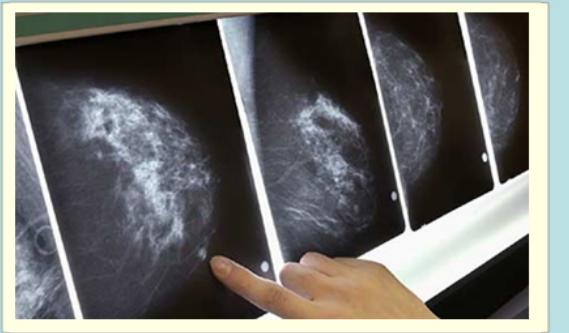


Abstract

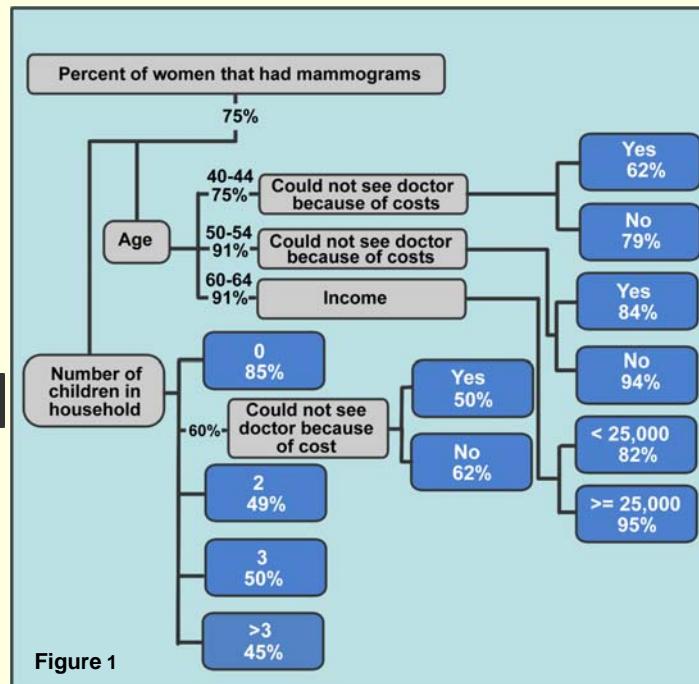
Breast cancer is a major chronic disease and early identification is necessary for treatment. Mammograms are clinical tests that increase the detection of breast cancer. Funded by the Centers for Disease Control and Prevention and the Missouri Department of Health and Senior Services, this preliminary work identifies profiles of women's breast cancer screening practices based on socio-economic factors using a data mining technique called CHAID. Major factors include age, cost of access, income and number of children in household.

Introduction

Breast cancer is one of the most common type of cancer among women in the United States[1]. Incidence of breast cancer and 5-year survival rates is known to be disparate among various subgroups of women[2]. Early identification of breast cancer through screening is vital for increasing the survival rates of all women. However, screening rates are lower than those desired under Healthy People 2010[3]. This raises the question as to the reasons for low screening rates and how these rates are different between subgroups of women based on socio-economic factors.



Chi-Squared Automatic Interaction Detection (CHAID) Breast Cancer Behavioral Risk



Methodology

Funded by the Centers for Disease Control and Prevention (CDC) and the Missouri Department of Health and Senior Services (MoDHSS), this research is at its preliminary stages and uses data mining to identify women having mammograms.

Chi-Squared Automatic Interaction Detection (CHAID) is a type of decision tree method[4] that goes over all possible variable combinations in the data to partition the node to explore relationships between a dependent variable and a series of predictor variables. CHAID was used to analyze behavioral risk data from 3331 women respondents to telephone surveys conducted as part of surveillance and reported by CDC. All profile relationships are significant at least with $p=0.01$.

Results

Seventy five percent of women have had a mammogram. More importantly, when analyzing the demographic variables, the most significant predictors of respondents ever having a mammogram were age and number of children in household. Figure 1 shows the various profiles based on these factors and the percentage of women that have had mammograms in each of the classifications. Each branch shows only the significant splits.

Conclusion

Age, number of children in household, cost of access to a doctor and income are among the major socio-economic factors in the profile of women having mammograms. This provides specific groups of women to target for intervention programs to improve screening practices. These profiles can also be simulated to study trends in public health.

References

1. Jemal A, Murray T, Ward E et al. Cancer statistics, 2005. *CA Cancer J Clin* 2005; 55(1):10–30.
2. Jatoi I, Becher H, Leake CR. Widening disparity in survival between white and African-American patients with breast carcinoma treated in the US Department of Defense Healthcare system. *Cancer* 2003; 98(5):894–899.
3. Aldridge M, Daniels J, Jukic A. Mammograms and healthcare access among US Hispanic and non-Hispanic women 40 years and older, 2006. *Fam Comm Heal* 2006; 29(2):80–88.
4. Kass GV. 1980. An Exploratory Technique for Investigating Large Quantities of Categorical Data. *App Stat* 29: 119-127.