癌症資料庫存活率線上分析系統

OTSAS: an Online Analytic System for Tumor Survival Analysis

中文摘要

在近代應用醫學統計中存活分析（survival analysis）扮演相當重要的角色，尤其在醫院的臨床研究中，存活分析在探討癌症病患的病程進展和治療的評估方式上，成爲不可或缺主要工具之ㄧ。

癌症成爲臺灣地區十大死因之首已經多年，癌症病患資料庫內資料量日益增加，電腦可以輔助臨床醫護人員及研究人員進行即時大量運算，提供快速又正確的統計分析，因此本文將存活率計算與 Web-Base 結合，發展一套癌症資料庫存活率線上分析系統，提供便利的動態查詢方式，幫助醫護人員有效率、且輕易完成資訊的結構分析，以存活分析模型呈現給查詢者，又可縮短資料萃取、篩選、分類、彙總的時間，增加臨床研究效率。

本文使用的癌症資料爲台北榮民總醫院癌病中心-癌症登記資料庫（Cancer Register Data Management System），內容爲 1993 年至 2002 年癌症個案，資料筆數共 50500 筆，包含人口學特性資料庫及癌症特性資料庫。

利用存活分析中三種常見的方式：對存活率估算的 Kaplan Meier、兩組存活曲線比較的 Logrank Test 及多種預後因子之存活分析的 Cox Proportional Hazards Model。

英文摘要

The study of longitudinal analysis plays an important role in contemnor medicine. The results may address the way to improve the efficacy and the efficiency of therapy. In most of problems, time occurs as an important factor for analyzing the data. Survival analyses are therefore to treat time as outcome. In survival analyses, probabilities are calculated not just for groups but individuals in a group. For clinical work, it is a great sufficiency for us to create the beneficiary of medical cares. Cancer has become the first leading cause of deaths in Taiwan since 1982. Although the medicine and technology have significantly improved, there are still bottlenecks in the cancer treatment areas which need to be broken. Analyzing of cancer patient's data can be a useful way to increase the effectiveness of cancer treatments. Hence that we are going to build a web-based survival analysis, OTSAS (abbreviated from Online Analytic System for Tumor Survival Analysis), offering the convenient dynamic query, that is able to assist the medical researchers efficiently and easily analyze the structure of information by using three famous survival statistical analyses. This thesis uses the data of CRDMS (Cancer Registry Data Management System) from VGH Taipei (Taipei Veterans General Hospital), which contains over

50500 cancer cases from year 1993 to year 2002, including the demography properties database and cancer properties database. This thesis provides three famous survival analysis methods： First is the Kaplan Meier method for generating the survival curve; second is the Logrank test for generating and comparing the survival curves of two groups, the last one is the Cox Proportional Hazards Model for estimating the risk ratio of some prognosis factors.