

Analysis of Effect of SSM on 152,989 Cancer Patients

Shuichi Shinmura

Seikei Univ., Dept. of Economics

Kichijoji Kitamachi 3-3-1

Musashino, Tokyo 180-8633 Japan

shinmura@econ.seikei.ac.jp

Kazumi Iida, Toshiyuki Hirai, Atsushi Nagazumi

The Research Institute of Vaccine Therapy for Tumors and Infectious Diseases,

Nihon Medical School

Sendagi 1-1-5, Bunkyo, Tokyo 113-8706 Japan

k-iida@nms.ac.jp

1. Introduction

At the Research Institute of Vaccine Therapy for Tumors and Infectious Diseases, treatment using SSM for cancer began in 1964. Since 1979, 152,989 patients administered SSM have been registered and data on them have been collected. In this paper, we analyze this data using the decision tree as well as other statistical methods in order to prove the effectiveness of SSM.

2. Data and Important Variables

Table 1 shows the frequencies of important events in the lifetime of the patient. The first event (e1) is the year of diagnosis. After that, according to the situation of the patient, e2 shows the year in which surgery took place, e3 shows the starting year of SSM treatment and e4 shows the ending year of SSM treatment. There is much missing data in

	e1	e2	e3	e4
-1980	771	16849	16356	7714
1981-1985	6113	20307	48002	29342
1986-1990	21536	19561	38086	28091
1991-1995	18866	15152	29970	19533
1996-2000	10719	7474	20548	15756
Sub Total	58005	79343	152962	100436
Miss. Val.	94984	73646	27	52553
Total	152989	152989	152989	152989

three of the four variables. And in most cases, the year of death is not clear. For this reason, we have used e4 instead of year of death.

In this research, long term survivors are distinguished from short term survivors by the number of their survival days using decision tree analysis, thereby creating a “long” group and a “short” group. By finding out what conditions are related to their survival, we will be able to indicate who are good candidates for SSM. This information can be useful for doctors and patients in the future. Our objective in this research is to find out if our classification is valid for the administration of SSM.

There are three dependent variables as follows. $y1(=e4-e1)$ is the number of days from the day of

diagnosis to the ending day of SSM therapy. $y_2 (=e_4-e_2)$ is the number of days from the surgery to the starting day of SSM therapy. $y_3 (=e_4-e_3)$ is the number of days from the surgery to the ending day of SSM therapy. There are over 30 predictor variables, among which $x_1(=e_3-e_1)$ and $x_2(=e_3-e_2)$ are particularly important.

The numbers of missing values of y_1 , y_2 , y_3 , x_1 and x_2 are 119,125, 105,548, 52,909, 105,253 and 87,449 respectively. Therefore, we mainly analyze the combination of (y_2, x_2) and (y_3, x_2) . We survey these variables by Box-whisker-plot and categorize 6 in intervals for the decision tree analysis as follows: interval ' 3' (non-positive), interval ' 0' (less than 30), interval ' 1' (greater than 31 and less than 90), interval ' 3' (greater than 91 and less than 360), interval ' 12' (greater than 361 and less than 1080), interval ' 36' (greater than 1081).

3. Analysis and Results

We analyze this data by SPSS, AnswerTree and Statistica. Initially, the entire data is analyzed, then 30 groups. There are 9,092 cases of stomach cancer, which are segmented by 29 terminal nodes. Four nodes are related by the variable x_2 and 500 cases are included in these four nodes.

An analysis in detail about gullet cancer will be announced verbally.

4. Conclusion

In this research, we analyze 118,301 that are composed of 30 groups of cancer. Although we have utilized various statistical methods, our results in this paper are based on decision tree analysis. Some famous doctors have said that SSM has no effect on cancer and it is same as water. If true, there is no reason to prescribe SSM for patients after surgery or diagnosis as soon as possible. Therefore, let us adopt the null-hypotheses model that SSM is water. But, many results by decision analysis show that x_1 and x_2 are important variables for segmentation compared to other variables. These facts imply that null-hypotheses model is not valid. There are many outputs from the decision tree, which we are considering disclosing on the World Wide Web.

REFERENCE

Shinmura, S. and Iida, K. and Maruyama, C. (1987). Estimation of the Effectiveness of Cancer Treatment by SSM using a Null Hypothesis Model. *Medical Informatics* 7(3), 263-275.

RESUME

Au centre de recherche sur le traitement par vaccin du Nippon Medical School, le traitement par SSM a débute depuis 1964. Depuis 1979, nous avons collectionné les données sur 152,989 patients. Dans cet article, nous analysons ces données sur base de la méthode arbre de décision aussi bien que d'autres méthodes statistiques pour prouver l'effectivité du SSM.