

# Construction of Statistical Models for Hospital Management

Yuko Tsumoto\*<sup>1</sup> and Shusaku Tsumoto\*<sup>2</sup>

1 Department of Fundamental Nursing,  
Shimane University, School of Nursing  
9-1 Enya-cho, Izumo 693-8501 Japan

`tsumotoy@med.shimane-u.ac.jp`

2 Department of Medical Informatics,  
Shimane University, School of Medicine  
89-1 Enya-cho, Izumo 693-8501 Japan

`tsumoto@med.shimane-u.ac.jp`

## Abstract

It has passed about twenty years since clinical information are stored electronically as a hospital information system since 1980's. Stored data includes from accounting information to laboratory data and even patient records are now started to be accumulated: in other words, a hospital cannot function without the information system, where almost all the pieces of medical information are stored as multimedia databases[1]. Especially, if the implementation of electronic patient records is progressed into the improvement on the efficiency of information retrieval, it may not be a dream for each patient to benefit from the personal database with all the healthcare information, "from cradle to tomb". However, although the studies on electronic patient record has been progressed rapidly, reuse of the stored data has not yet been discussed in details, except for laboratory data and accounting information to which OLAP methodologies are applied. Even in these databases, more intelligent techniques for reuse of the data, such as data mining and classical statistical methods has just started to be applied from 1990's[2, 3].

Human data analysis is characterized by a deep and short-range investigation based on their experienced "cases", whereas one of the most distinguished features of computer-based data analysis is to enable us to understand from the different viewpoints by using "cross-sectional" search. It is expected that the intelligent reuse of data in the hospital information system provides us to grasp the all the characteristics of university hospital and to acquire objective knowledge about how the hospital management should be and what kind of medical care should be served in the university hospital. This paper focuses on the following two points for analysis. One is what kind of knowledge (statistical models) can be extracted by statistical methods from the datasets stored for about twenty years in Chiba University Hospital. The other is how these pieces of knowledge are useful for the future hospital management and decision support. For construction of statistical models, we applied R to large hospital data because

R gives a wide variety of statistical model construction with nice visualization interface..

The analysis gives interesting results: (1) malignant neoplasm is the first major category which determines the profitability of Chiba University Hospital, which is stable for twenty years. (2) In a global view, the length of stay is the principle factor for the revenue of the hospital, whose distribution follows the log-normal distribution. (3) Treatment method may be a secondary factor to determine the distribution of the length of stay for each disease, which may be correlated with the property that the length of stay follows log-normal distribution for each minor division in total. (4) Treatment without a surgical operation should be more examined by additional information, which is also important to evaluate the profitability of the university hospital.

## References

- [1] Institute of Medicine Committee on Improving the Patient Record (1997) The Computer-based Patient record: An Essential Technology for Health Care. National Academy Press, Washington DC.
- [2] Tsumoto S (2000) Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic. *Inf. Sci*, 124(1-4): 125-137.
- [3] Tsumoto S (2001) Chapter G5: Data mining in medicine, In: Kloesgen W, Zytkow J, editors. *Handbook of Data Mining and Knowledge Discovery*, pp.798-807, Oxford University Press, Oxford.