

4136076 PHBS/M : MAJOR : BIostatistics ; M.Sc.(BIostatistics)
 KEY WORDS : CASE-CONTROL STUDY / LOG-LINEAR MODEL

KRITTIKA SUWANRUNGRUANG: LOG-LINEAR MODEL APPROACH FOR CASE-CONTROL STUDY. THESIS ADVISORS: JUNYA PATTARAARCHACHAI, Sc.D., SUMALEE SINGHANIYOM, M.Sc. 94 P. ISBN 974-664-958-2

The purpose of this study was to determine the appropriate model, which could explain the relationship between risk factors and breast cancer. The population for this study was 252 cases aged less than 70 years with breast cancer and 504 controls within ± 5 years of their matched patient's age. Both cases and control groups were recruited from hospitals in the Northeast from November 1, 1998 to March 31, 2000. The study variables were age, number of pregnancies, age at menarche, age at menopause, age at first pregnancy, age at last pregnancy, oral contraceptive, body mass index, x-ray, family history and duration of breast feeding. This study was a case-control design and the data was analyzed by log-linear model.

Results showed that eight variables were found in six models. Among those, the best model was model 1, that could explain 87% ($R^2 = .87$) of the variability in the data. When the same variables were assessed by logistic regression, only R^2 equal to 1.26% was found. This indicated that log-linear model gave better explanation than logistic regression.

The variables included in the log-linear model were

$$\begin{aligned} \log(m_{ijkl}) = & 4.6303 - .5128 (\text{age below 45 years}) - .9431 (\text{case}) \\ & -1.5734 (\text{number of pregnancies two time or less}) \\ & -2.2289 (\text{age} < 45 \text{ years}) * 1.0236 (\text{age at menarche} < 15 \text{ years}) \\ & +2.1783 (\text{age} < 45 \text{ years}) * (\text{number of pregnancies} < 3) \\ & +.8209 (\text{age} < 55 \text{ years}) * (\text{number of pregnancies} < 3) \\ & +.3977 (\text{case}) * (\text{number of pregnancies} < 3) \end{aligned}$$

Recommendation as a result of this study: The sample size should be large enough for the potential major and minor effects being remain in the model. In order to give better explanation, more than 2 categories of classification should be used to express the effect for further studies.