

Application of the rough sets theory to evaluate prognostic factors in breast cancer patients subjected to mastectomy*

Renata Szoszkiewicz¹, Jerzy Załuski¹, Jerzy Krysiński², Szymon Wilk³, Bartłomiej Prędko³

¹Chemotherapy Ward, Wielkopolska Oncology Center, Poznań, ²Faculty of Pharmacy, Nicolaus Copernicus University Collegium Medicum in Bydgoszcz, ³Institute of Computing Science, Poznań University of Technology, Poznań, Poland.

Background: the paper presents analysis of relationship between variables describing breast cancer patients and therapy results. The method based on the rough sets theory and induction of decision rules is applied to perform the analysis. Rough sets are a method of dealing with domains characterized by inconsistent and incomplete information. Proceeding in this way, they formulate some indications, which may be helpful in making decisions referring to the treatment of breast cancer patients.

Material and methods: the data set contains 718 breast cancer patients described by 21 variables (factors) and divided into two classes: patients who did not experience cancer recurrence and patients who had cancer recurrence. In the years 1992-1994, those patients were subjected to mastectomy and underwent chemotherapy at the Chemotherapy Ward of the Wielkopolska Oncology Centre in Poznań. The observation period was equal to 10 years (2002-2004). The whole group of patients was divided into two sets: a learning set and a testing one.

Results: in the first phase of the analysis, the rough sets based approach was applied to determine variable importance for the patients' classification. The set of selected variables, which ensured high quality of the classification, was obtained. Then, the decision rules were generated from the learning set by means of the algorithm inducing the minimal cover of the learning examples. The testing set was a base to evaluate prognostic potential of the generated decision rules. Total accuracy of prognosis (classification) for the decision rules was equal to 70.3 %. In the case of the patients who had had cancer recurrence the prognosis accuracy was 76.3%, and for the patients who had had no recurrence of cancer it was 60.7%. The prognosis accuracy is described as a ratio of number of test cases for which the rules correctly indicated cancer recurrence or lack of recurrence to the total number of test cases.

Conclusions: the obtained decision rules provide guidelines which may be helpful in making decisions referring to treatment of breast cancer patients as well as evaluating their prognosis.

*The research was supported by grant no. 3T11E04226 from State Committee for Scientific Research (KBN).