Comorbidity in Cancer Mortality Analysis

Tsurko V.V.¹, Michalski A.I.¹

¹Institute of Control Sciences, RAS, Moscow, Russia (E-mails: v.tsurko@gmail.com, ipuran@yandex.ru)

Abstract

Investigation of relationships between health and mortality is actual because of longevity increase, which is observed in developed countries starting from the second part of the XX century. In order to release economic and social pressure due to ageing of population it is important to have the solution to following problems: obtaining reliable estimates for expected age structure of the population, gaining knowledge of factors responsible for “healthy aging” and understanding the impact of different diseases in cause-specific mortality. The last problem is known as mortality-comorbidity problem. This problem is especially important for old age groups in which the mortality is at the high level and several chronic diseases are presented.

Nowadays there is a great volume of statistical data for mortality and morbidity of the aged people. This data allows us to investigate the factors responsible for maintaining health in aging population, to evaluate the influence of heredity, environment and lifestyle.

Recent publications explain observed increase of human life expectancy by the reduction of mortality at middle age [1,4]. There is a hypothesis that people who are down in health have a high margin of “active longevity”, because enduring high health risk in young and middle ages gives an advantage in survival in old age. Effective adaptation of people with chronic diseases may serve as a biological basis for this phenomenon. If the hypothesis is correct, we need to focus preventive measures to ensure the "healthy aging" in the age groups of young and middle ages.

Relationship between the cause specific mortality and chronic diseases can be an indirect confirmation of the relationship between increased morbidity and reduced mortality. In the research the Multiple Cause-of-Death Public-Use Data for 2007 by the National Center for Health Statistics USA [3] is investigated. Distribution of associated diseases presented by the ICD10 codes among people who died of cancer (C00-C97) is compared with the same distribution among people who died of another disease. In order to select more “important” diseases associated with cancer mortality we solve a problem of contrasting the distributions. By the problem of contrasting we mean the selection of associated diseases for which we have the most distinguishable distributions.

We used symmetrized Kullback–Leibler divergence as a difference measure between the two distributions. For a set of associated diseases the symmetrized Kullback–Leibler divergence was estimated from the data as a half sum of mixed entropies corrected by a penalty term. This term takes into account both the amount of empirical data and the number of considered associated diseases. Two techniques for construction of such penalty term were used in the report: one is based on the Vapnik-Chervonenkis dimension and the second applies Rademacher complexity.

The results show that in a group of women at the age of 60-64 died of cancer not more than 8 of 24 considered classes of associated diseases are statistically linked with cancer mortality. Considering another group of women of the same age whose mortality cause was different from cancer and excluding of “metastasis” associated disease makes the class of chronic pulmonary diseases with high prevalence of asthma (ICD10 J45-J46) the most significant. This research validates the hypothesis that presence of asthma could decrease development and mortality of cancer [2].

Keywords: cancer mortality, distributions discrepancy, selection of associated diseases.

References