Секция «Биофизика и нанобиотехнологии»

Machine learning analysis of clonogenic survival data for radioresistance evaluation in cancer cell lines

Научный руководитель – Гольцов Алексей Николаевич

Ваннус Маджед

Студент (магистр)

МИРЭА - Российский технологический университет, Институт международного образования, Москва, Россия

E-mail: majdwannous0098@gmail.com

Effectiveness enhancement of radiotherapy (RT) and its combination with chemo- and targeted therapies is determined basically, by the development of the new RT units with improved beams quality and optimized RT plans, but on the other hand, by including molecular and radiomics-based biomarkers of radiosensitivity (RS) of patients' tumors into the RT plans. Currently, for optimization of the RT plans and prediction of RT results, machine learning (ML) and artificial intelligence (AI) are being developed based on the analysis of a large amount of experimental and clinical data on the molecular and structural characteristics of cancerous tumors. One of the main integral characteristics of cancer cells is related to their therapeutic response to ionizing radiation (IR), which is determined by the dose dependences of clonogenic cell survival. The problem of using these data in RT plans for the selection of the radiation dose and its fractionation is associated with difficulties in extracting reliable quantitative characteristics from the dose-effect data.

In this work, we developed a computational method for determining the RS of cancer cells based on the analysis of clonogenic survival data using machine learning. The method consists in clustering the characteristics of cancer cells in order to determine clusters of RS and radioresistant (RR) cancer cells. To determine the parameters of cell survival, the experimental dose-effect data were approximated using the equation of a mathematical model used in radiobiology: the linear-quadratic (LQ) model. In the process of training of the developed ML model, we used a set of published experimental data on over 130 human cancer cell lines with known RS. As a result of experimental data approximation by the LQ model, two parameters α and β were determined, the ratio of which is traditionally used in radiobiology to evaluate the radiosensitivity of cells. A high value of the α/β ratio is characteristic of RS cells, which have a low ability to repair damage after IR. A low α/β value is characteristic of RR cells and reflects the high ability of cells to repair.

The obtained α/β values of the studied cells ranged from 0 Gy (PC3 0.5%) to 259.4 Gy (FamPac). In most cancer cells, the α/β ratio was comparable to the value of non-malignant HPDE cells (α/β = 4.88 Gy), which makes it difficult to discriminate between healthy and cancer cells. The classification of RS and RR cells based on the α/β ratio did not allow the cells to be reliably separated into two groups (clusters), including either RS or RR cells.

In order to increase the reliability of discrimination between RS and RR cells according to clonogenic survival data, cell clustering was performed based on the principal components analysis (PCA) in the parameter space (α, β) along with the K-means clustering method and hierarchical clustering. As a result, a reliable separation of RS and RR cells into two clusters was achieved.

This method was implemented in Python language and included a module for determining the parameters α and β for the LQ model based on experimental dose dependences of clonogenic survival.