Leveraging machine learning techniques for predicting pancreatic neuroendocrine tumors grade using biochemical and tumor markers

Rui-Quan Zhou, Hong-Chen Ji, Qu Liu, Chun-Yu Zhu, Rong Liu

We thank for your work about this manuscript and we now answer the questions raised by reviewers and editors as follows:

Question 1: Please revise the manuscript according to the review report and my comments. And answer all of the reviewers' comments carefully.

Answer 1: According all the 3 reviews' comments, they were satisfied about the method, results, figures and tables. And there's some minor language polishing should be corrected and we have the paper language editing before and uploaded the English Editing Certificate to the editing system.

Question 2: Please provide and upload the approved grant application form(s).

Answer 2: We have replaced the support grant and upload associated files in editing system.

Question 3: Please write the article highlight section according to the guidelines listed below. Please don't copy from the main text.

Answer 3:

Research background

The incidence of Pancreatic neuroendocrine tumors (PNETs) has

increased rapidly, and to establish an predicting system of the tumor grade of PNETs defined by WHO is beneficial for the prognosis and treatment of PNETs. However, obtaining of the tumor grade by surgery or biopsy means a lot trauma, therefore, a simple and feasible method that could predict the grade of PNETs non-invasively would be very meaningful.

Research motivation

Machine learning (ML) algorithms have shown potential in improving the prediction accuracy using comprehensive data. We used four classical ML models in this article and we found machine learning could be a potential and feasible method to to predict the grade of PNETs by using routine clinical data. ML could be utilized effectively in solving some medical classification problems.

Research objectives

To provide a machine learning approach to predict the tumor grade of PNETs using clinical data and ML is effective in classifying the grade of PNETs by using the routine data obtained from the results of biochemical and tumor markers. This approach may be a promising method to predict the grade of PNETs non-invasively and has the potential to be used widely in clinical settings.

Research methods

The biochemical outcomes and tumor markers of 91 patients with histologically confirmed PNETs were collected, and a novel method of minimum p for the Chi-square test (MPCST) was used to divide the continuous variables into binary variables. Four classical supervised ML models, including logistic regression (LR), support vector machine (SVM), linear discriminant analysis (LDA) and multi-layer perceptron (MLP)

were trained by clinical data. The models were labeled with the pathological tumor grade of each patient. The performance of the different models was then evaluated. Finally, the weight of the different parameters in each of the models were calculated.

Research results

All the four models showed a potential performance in this classification task. Among them, LDA showed the best performance in predicting the grade of PNETs and MLP has the highest recall rate for grade 3 patients. These results proved that the models trained by the clinical data would give a feasible way to predict the pathological tumor grade of PNETs. However, there are still a few limitations in this study. Some parameters like tumor size and matastasis found form CT images were not included in this article. Because we think the two parameters may be not objective and may exist errors in data collection. In general, the result of our study provided a non-invasive way to judge the condition of PNETs and offers a reference for treatment.

Research conclusions

The following questions should be briefly answered:

What are the new findings of this study?

Machine learning is effective in classifying the grade of PNETs by using the routine data obtained from the results of biochemical and tumor markers.

What are the new theories that this study proposes?

Machine learning (ML) algorithms have shown potential in improving the prediction accuracy of classification of the grade of PNETs using

comprehensive data.

What are the appropriate summarizations of the current knowledge that this study provided?

There is still no effective way to determine the grade of PNETs non-invasively.

What are the original insights into the current knowledge that this study offered?

Machine learning (ML) algorithms have shown potential in improving the prediction accuracy using comprehensive data

What are the new hypotheses that this study proposed?
The combination of imageology and serological outcomes may improve the classification power of ML models.

What are the new methods that this study proposed?

A novel method of minimum p for the Chi-square test (MPCST) was used to divide the continuous variables into binary variables.

What are the new phenomena that were found through experiments in this study?

Patients of grade 3 (G3) showed more significant differences than grade 1 (G1) and grade 2 (G2)

What are the hypotheses that were confirmed through experiments in this study?

Machine learning is effective in classifying the grade of PNETs by using the routine data obtained from the results of biochemical and tumor markers

What are the implications of this study for clinical practice in the future?

Machine learning may be a promising method to predict the grade of PNETs non-invasively and has the potential to be used widely in clinical settings

Research perspectives

What experiences and lessons can be learnt from this study?

Some very simple and routine clinical data may play an important role in some medical classification tasks by using machine learning method.

What is the direction of the future research?

The combination of imageology and serological outcomes may improve the classification power of ML models. And more effective ML models could be utilized in this classification task.

What is/are the best method/s for the future research?

The combination of clinical data and experience and new ML models.

Question 4: Please write the article highlight section according to the guidelines listed below. Please don't copy from the main text.

Answer 4: We will upload the figures as separate electronic files in editing systems.