| | |
|---|---|
| 氏　　　　　　　　　名 | TRAN NHU NGOC QUE |
| 博士の専攻分野の名称 | 博　士　（　医　学　） |
| 学　位　記　番　号 | 医工農博４甲　第８５号 |
| 学 位 授 与 年 月 日 | 令和６年３月２２日 |
| 学 位 授 与 の 要 件 | 学位規則第４条第１項該当 |
| 専　　　攻　　　名 | 医学専攻 |
| 学 位 論 文 題 名 | A Machine Learning-Based Model to Predict In-Hospital Mortality of Lung Cancer Patients: A Population-Based Study of 523,959 Cases<br>(肺がん患者の院内死亡率を予測するための機械学習ベースのモデル: 523,959 例を対象とした集団ベースの研究) |
| 論 文 審 査 委 員 | 委員長　　教　授　　松川　隆<br>委　員　　准教授　　中村　貴光<br>委　員　　講　師　　加賀　重亜喜 |

# 学位論文内容の要旨

**Purpose:** In 2022, there were approximately two million new cases of lung cancer diagnosed in the U.S. In comparison to those of GLOBOCAN a decade ago, it can be considered an impressive step forward in lung cancer treatment with the increasing trend of survival. Nevertheless, that is not the case in low and middle-income countries, where lung cancer remains challenging as the mortality rate is approximately ninety percent of the incidence rate. There are some researches revealing the correlation between lung cancer and other factors such as sex, frailty, and brain metastasis. Despite that fact, there is no finding to stratify the in-hospital mortality risk of new lung cancer patients at current hospital admissions for diagnoses, only based on minimum epidemiological components. Subsequently, this individualized approach can assist healthcare providers to quantitatively strategize not only disease treatment but also end-of-life care in the first place, especially in countries with less accessible healthcare services.

**Methods:** 522,941 lung cancer cases with available data on the Surveillance, Epidemiology, and End Results (SEER) were analyzed for the predicted probability based on six fundamental variables including age, gender, tumor size, T, N, and AJCC stages. The patients were randomly assigned to the training (n = 115,145) and validation datasets (n = 13,017). The remaining cohort with missing values (n = 394,779) was then combined with the primary lung tumor datasets (n = 1018) from The Cancer Genome Atlas, Lung Adenocarcinoma and Lung Squamous Cell Carcinoma projects (TCGA-LUAD & TCGA-LUSC) for external validation and sensitivity analysis. Descriptive statistics of continuous and categorical variables were median (range) and the number of cases (percentage). We used Wilcoxon and chi-square tests to compare the difference in continuous and categorical features, respectively. The HI-VAE model was trained, using the original pipeline published by Nazabal et al. Other analyses were performed, using R software version 4.2.2 (The R Foundation, Vienna, Austria).

**Results:** Receiver Operating Characteristic (ROC) analyses showed high discriminatory power in the training and internal validation cohorts (Area under the curve [AUC] of 0.78 (95%CI = 0.78-0.79) and 0.78 (95%CI = 0.77-0.79), respectively), whereas that of the model on external validation data was 0.759 (95%CI = 0.757-0.761). We developed a static nomogram, a web app, and a risk table based on a logistic regression model using algorithm-selected variables. In particular, the nomogram visualizes the predictive power of each predictor compared to each other. Each 10 years of age increases the total risk by 2.5 points while each 5 cm of tumor size gives the total score 3 points. Categorical variables including gender (Male-1 point), AJCC stage (stage II-2.5 points; stage III-4.5 points; and stage IV-10.5 points), T stage (T2-0.7 points; T3-1.2 points; T4-2.5 points), and N stage (N2-3-1 point). We translated the risk calculation of the static nomogram into the risk table. These 2 tools have similar scoring system. On the other hand, dynamic nomogram is associated with the shiny web tool that is linked to the author's account (https:// lkhangkv1995. Shinyapps. io/LungCancer_In-hospitalMortality-nomogram). This application can yield the predicted probability of in-hospital mortality of the patient of interest as well as its 95%CI. However, it is noteworthy that the model may not yield an accurate prediction when the predicted probability is more than 50%, which can be observed in the calibration plots.

**Discussion:** From the statistical standpoint, our study shows the predictive value of the training model as the probability varies from 0 to 50% of in-hospital mortality since the initial diagnosis time-point. From the clinical standpoint, there are biases and confounding factors in the current study. First, in terms of diagnosis, the protocol for the differentiation of primary lung cancer from solitary lung metastasis, especially with unknown primary tumors, is challenging since the lung is considered a frequent metastatic location. Therefore, the model should be applied after the primary lung location is confirmed by careful examination. Second, the policymaking of lung cancer management varies from place to place. Because the model was built based on the SEER database, TCGA-LUAD, and TCGA-LUSC, it partly reflects the outcome of lung cancer management in developed countries. Furthermore, in spite of cutting-edge interventions related to immune-, chemo-, radiotherapy, etc., lung cancer is still the leading cause of mortality. Our model with its robustness could be useful for clinicians to adjust an individualized strategy depending on the real-time risk evaluation at every certain follow-up examination. This helps to improve the survival rate in the landscape of lung cancer. Regarding applications, we designed three platforms to put the model into practice for quantitative groundwork. The cut-off of 26 points for patient differentiation has not only robustness in terms of statistics but also clinical practicality that aids physicians in patient-centered planning. For instance, with respect to the function of rapid response teams (RRT) in the inter-association efforts for survival improvement. Our model is advantageous to leverage their role by focusing on high-risk groups in terms of end-of-life care enhancement. Consequently, in order to regulate possible delays, healthcare institutes can efficiently implement a system that pays attention to those vulnerable patients so that RRT can have timely activation.

**Conclusions:** Our model can stratify lung cancer patients into high- and low-risk of in-hospital mortality to assist clinical further planning. Three well-developed interfaces are friendly to both physicians and patients for prognosis-related conversations.

# 論文審査結果の要旨

１．学位論文研究テーマの学術的意義
　肺がん患者の予後に関して様々な検討がなされているが、決定的とされるものはない。
本論文では、肺がん患者の院内死亡率を予測するために、機械学習ベースモデル（523,959 症例）を対象とした集団ベースの研究を行った。

２．学位論文および研究の争点、問題点、疑問点、新しい視点等
＜背景＞
院内死亡が予見される新規肺がん患者を対象として本研究を行った。
＜方法＞
522,941名の原発性肺がん患者をデータベースから描出した。
6つの要素（年齢、性別、腫瘍の大きさ、TN stage、AJCC stage）を抽出し、機械学習ベースのモデルを作成し、それに基づいて予測を行った。
＜結果＞
ROC(Receiver, Operating Characteristic)分析によって、内的要因に基づいたデータCohortsでは有意差を認めたが、外的要因データでは有意差は無かった。
著者らは、アルゴリズム変数を用いたlogistic regressionに基づき、nomogramを作成し、院内死亡率の予測が可能であることを示すことが出来た。
＜結語＞
著者らの作成したモデルは、肺がんの院内死亡に関して、high riskとlow riskを予測可能であり、これにより、臨床における治療プランに非常に重要で有意義な情報を与え得ると考えられる。

３．実験およびデータの信頼性
　本研究において、研究のデザインや統計解析の方法などが明確に記載されており、これらの研究データの信頼性は十分であると判断した。

４．学位論文の改善点等
　今回提出された論文は、内容も様式も学位論文として全く問題が無く、改善点は無いと判断した。