Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

LANTERN-01: AI model for Postoperative Complications Prediction after NSCLC Lung Resection: Prospective Multicentric Study and Extern Validation

Filippo Lococo
Catholic University of the Sacred Heart
Carolina Sassorossi
Sassorossi.caro@gmail.com

Catholic University of the Sacred Heart
Davide Dalfovo
University Hospital Carl Gustav Carus, TUD Dresden University of Technology
Annalisa Campanella
A. Gemelli University Hospital Foundation IRCCS
Luca Boldrini
Catholic University of the Sacred Heart
Filippo Gallina
Regina Elena National Cancer Institute
Mattia Bruschi
Regina Elena National Cancer Institute
Virginia Proietti
A. Gemelli University Hospital Foundation IRCCS
Akshaya Balamurugan
A. Gemelli University Hospital Foundation IRCCS
Esther G C Troost
OncoRay – National Center for Radiation Research in Oncology, University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Helmholtz-Zentrum Dresden- Rossendorf
Steffen Löck
University Hospital Carl Gustav Carus, TUD Dresden University of Technology
Róza Ádány
ELKH-DE Public Health Research Group, Faculty of Medicine, University of Debrecen
Núria Farré
Hospital de la Santa Creu i Sant Pau (IR-HSCSP)
Ece Öztürk

School of Medicine, Turkey and Koç University Research Center for Translational Medicine (KUTTAM), Sariyer, Koç University, , Turkey.

Edoardo Mercadante
Regina Elena National Cancer Institute

Marco Chiappetta
Thoracic Surgery Unit, Magna Graecia University

Alessandra Cancellieri
A. Gemelli University Hospital Foundation IRCCS

Rocco Trisolini
A. Gemelli University Hospital Foundation IRCCS

Emilio Bria
A. Gemelli University Hospital Foundation IRCCS

Antonio Gasbarrini
Catholic University of the Sacred Heart

Stefano Margaritora
Catholic University of the Sacred Heart

Research Article

Keywords: NSCLC, AI, post operative complications, personalized medicine

Posted Date: February 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8554813/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Introduction

: Artificial intelligence(AI) techniques may combine various omics datasets to create more accurate predictive models for lung cancer patients management. The aim of this study from the LANTERN project, is to develop an AI-based predictive model of post-operative complications after lung resection for NSCLC.

Methods

In the framework of LANTERN Consortium we prospectively collected data from 3/2023 to 12/2024 of patients who underwent curative surgery for Stage I-III NSCLC and were herein analyzed considering 80 preoperative clinical features and 43 spirometric variables to predict the occurrence of post-operative complications. Prediction models were developed on the basis of different feature selection algorithms and machine-learning models. An external dataset composed by a surgical series of 232 patients was used for validation.

Results

The final analysis was conducted on 231 surgical patients. Post-operative complications were observed in 37 patients (16%). AI-based models showed that pathologic score (AUC = 0.72, 95% CI [0.62–0.81]) and pre-opFEV1_TEOR (AUC = 0.77, CI [0.67–0.89]) was the most relevant variables. Testing the models on the external dataset, while the predictive value of Pathologic score alone was reduced, pre-opFEV1_TEOR and the combination of Pathologic score and pre-opFEV1_TEOR were confirmed to predict post-op outcome. Postoperative risk rises by about 6% per pathologic score level and 11–12% for each 10% decrease in FEV1_TEOR.

Conclusions

Combining the FEV1_TEOR and the pathologic score permits to predict the complications risk in significant way. This model could be tested in a further prospective cohort of patients to verify its effectiveness in order to improve the perioperative management of lung resection candidates.

INTRODUCTION

The application of artificial intelligence (AI) and machine learning (ML) into lung cancer treatment has shown potential in improving diagnostic accuracy, treatment personalization, and patient outcomes [1]. AI applications in thoracic surgery is revealing the potential to improve surgical precision, patient safety, and overall care efficiency [2]. Despite remarkable advances in lung cancer management, postoperative

complications continue to represent a major determinant of outcomes, as they may increase morbidity and mortality. In the era of novel multimodal strategies for resectable NSCLC, which may also affect the complexity of surgical procedures, the identification of patient subgroups at higher risk of postoperative complications is of paramount importance. Such stratification could enable the implementation of personalized perioperative strategies aimed at improving overall outcomes. Under this point of view, Advances in precision medicine offer hope to improve also surgical outcomes [3].

Previous studies investigated the possibility to individuate high risk patient in different setting (5), considering also the radiological or pre-operative features and using AI based models (5,6)

In this frame, the LANTERN (Lung cancer multi-omics digital human avatars for integrating precision medicine) project [4] arises, with the aims to deliver a novel approach for comprehensive lung cancer decision making solutions, based on predictive digital platforms powered by the integration multi-omics data. The current study, namely LANTERN1, has been led with the aim of creating a machine learning algorithm, able to analyze all the pre operative features of patients candidate to thoracic surgery, to predict the post operative complication risk.

In literature, different studies have been published with the aim of predicting post operative complications after non cardiac thoracic surgery. For example, He and coworkers [5] analysed a ICU (intensive care unit) cohort, to predict in particular respiratory complications and they used univariate and multilogistic regression to perform the analysis. Another interesting experience, from Nonomura and coworkers [6] tries to predict the risk of pulmonary complications on the base of pre operative radiological features. A very interesting experience is from Salati et al [7]. They already proposed an AI based model able to define the risk of cardiac and pulmonary complications in the early postoperative period for patients submitted to anatomic lung resection. By the way, as reported by the authors, their experience is a monocentric one and they analyzed the risk on the base of 50 preoperative features. Our study considered more than 100 variables and comprises an external cohort to validate the results. To the best of our knowledge, our analysis would be the first one with these characteristics, with the aim of predicting post operative complications, in patients candidate to surgery for NSCLC, with AI algorithms. This would improve the process of risk definition and therapy planning for patients undergoing lung surgery

With this study, we aimed to develop an AI based model for complication risk in a large cohort of patients who underwent curative surgery for NSCLC.

METHODS

Ethics

The protocol of the study has been drawn up in accordance with the Standards of Good Clinical Practice of the European Union and the current review of the Helsinki Declaration and has been approved by the Ethics Committee of this structure. All methods will be carried out in accordance with relevant guidelines

and regulations. The informed consent has been obtained from all subjects and/or their legal guardian(s). Approval of Fondazione Policlinico Universitario Agostino Gemelli IRCCS – Università Cattolica del Sacro Cuore Ethics Committee Number: 5420 – 0002485/23; Trial registration - clinicaltrial.gov: NCT05802771.

Study Design and Population

This study was designed as a prospective, multicenter observational trial followed by external validation, in accordance with the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) guidelines [8]. Data for 337 patients were collected prospectively between March 2023 and December 2024 within the framework of the LANTERN consortium [4]. Of these, 231 patients with Stage I-III Non-Small Cell Lung Cancer (NSCLC) underwent curative-intent surgery. Spirometry was performed for 176 of these patients (Fig. 1). To assess the generalizability of the final model, we used an independent, retrospective surgical cohort of 232 patients collected from a separate cancer research center..

Data Acquisition and Endpoints

We collected and categorized a comprehensive set of preoperative variables. The clinical dataset included 80 variables, covering among other things, patient demographics, clinical characteristics, and a panel of comorbidities (see Supplementary Table 1A). The second part of the dataset consisted of 43 spirometry variables derived from preoperative pulmonary function tests, including but not limited to Forced Expiratory Volume in one second (FEV1), Diffusing Capacity of the Lungs for Carbon Monoxide (DLCO), and other derived flow and volume parameters (see Supplementary Table 1B). The primary endpoint was the occurrence of any postoperative complication, graded according to the Clavien-Dindo classification [9].

We considered 20 adverse events as defined as Fernandez and coworkers reported [10] : air leakage, pleural effusion, subcutaneous emphysema, pneumonia, arrhythmia, hemorrhage, sepsis, wound dehiscence, abdominal collection, anemia, acute myocardial ischemia, deep venous thrombosis, fistula, infected wound, occlusion, pulmonary embolism, renal failure, infection of the urinary tract.

A further sub-analysis was performed to predict specific complications, grouped as cardiovascular events (defined as acute myocardial ischemia, arrhythmia, deep venous thrombosis pulmonary embolism) and respiratory events (defined as pleural effusion, air leakage, subcutaneous emphysema, pneumonia).

The cardiopulmonary as well as respiratory complications were analyzed as a binary outcome (complicated vs not-complicated patient). Only complications occurred within the first 30 days after the operation, or later if experienced within the same postoperative hospital stay were included in the present analysis.

Before proceeding with the AI analysis, the database underwent a preprocessing with specific data quality assessment procedures for optimizing data completeness (reported in Supplementary Table 1A and 1B) and consistency.

Feature Selection, Model Optimization and Evaluation

To identify preoperative predictors, we designed and implemented a machine learning workflow. First, we performed a feature selection protocol embedded within a 20-repetition bootstrapping framework to ensure the stability of feature rankings. Three distinct algorithms were applied: Minimum Redundancy Maximum Relevance (mRMR) [11], Lasso logistic regression [12], and Elastic-Net logistic regression [13]. Feature importance was aggregated across all repetitions to generate a stable ranked list, and the algorithm demonstrating the highest average Area Under the Receiver Operating Characteristic (AUROC) curve was chosen as the optimal method (Supplementary Figs. 1A and 1B).

Using the best ranked feature list, we then identified the optimal combination of signature length and machine learning model. We evaluated four supervised learning algorithms: logistic regression (LR), random forest (RF) [14], a support-vector machine (SVM) with a radial basis function (RBF) kernel [15], and eXtreme Gradient Boosting (XGBoost) [16]. For each algorithm, we incrementally constructed feature signatures of increasing length, from 1 (the most important feature) to 10 (the top 10 most important features). The performance of each algorithm and signature length combination was evaluated using a 5-repetition, 3-fold cross-validation scheme on the training data. Model-specific hyperparameters were tuned to maximize performance. The configuration that maximized the mean AUROC across all cross-validation folds was selected as optimal.

Finally, predictive models were trained on the entire training dataset using the optimal learner and feature signature length. This process was performed using clinical dataset only, spirometry dataset only, and integrating top features from both datasets. The performance of these final models was then assessed on the independent, external validation cohort by calculating the AUROC with 95% confidence intervals.

All data processing, statistical analysis, and machine learning model development were conducted using R (version 4.5.1). The feature ranking, signature determination, and final model development were supported by the “Fully Automated Machine Learning with Interpretable Analysis of Results” (FAMILIAR) framework [17].

The initial step in our modeling process was to identify the most stable and performant feature selection algorithm. We evaluated three distinct methods within a 20-repetition bootstrapping framework. The Elastic-Net algorithm consistently achieved the highest average Area Under the Receiver Operating Characteristic (AUROC) curve for both the clinical and spirometric datasets (Supplementary Fig. 1A and 1B).

RESULTS

Patient Characteristics and Outcomes

A total of 231 patients who underwent curative-intent surgery for NSCLC were included in the training cohort. The baseline clinical and pathological characteristics of this cohort are summarized in Table 1. The mean age was 68.6 years (SD ± 9.5), and the majority of patients were former or current smokers (184 patients, 79.7%). Postoperative complications of any grade occurred in 37 patients, corresponding to an overall complication rate of 16.0%. A total of 48 complications occurred in these 37 patients, as nine patients experienced two complications and one patient experienced three. Among the patients with complications, the majority experienced low-grade events (Clavien-Dindo grade I-II: 23 patients, 63.9%), with a mean time from surgery to complication onset of 7 days (SD ± 13). The characteristics of the complications are summarized in Fig. 2 and Table 2.

Model Development and Internal Performance

To identify the most stable and performant feature selection algorithm, we compared the performance of three methods. The Elastic-Net algorithm consistently achieved the highest average Area Under the Receiver Operating Characteristic (AUROC) curve for both the clinical and spirometric datasets (Supplementary Fig. 1A and 1B) and was therefore selected to generate the final feature importance rankings.Based on these Elastic-Net rankings (Supplementary Figs. 2 and 3), the machine learning workflow identified the Charlson Comorbidity index (also defines as Pathologic Score in text (see supplementary Table 1)) (Fig. 3A) and the ratio of preoperative Forced Expiratory Volume in one second to its theoretical value (FEV1_TEOR) (Fig. 3B) as the most influential variables for predicting postoperative complications. This was determined by evaluating four different machine learning models across various feature set sizes (signature lengths). The tables in Figs. 3A and 3B report the peak performance for each model. The selection of these features is the result of a cross-validated hyperparameter search that included optimizing the signature length for four distinct machine learning algorithms. The tables report the optimal cross-validated AUROC for each algorithm. A logistic regression model with a signature length of one demonstrated the highest predictive power for both clinical (AUROC of 0.674) and spirometric (AUROC of 0.738) datasets.

When performed on the entire training cohort, a model based on the pathologic score alone achieved an AUROC of 0.72 (95% CI [0.62–0.81]), while a model using only preoperative FEV1_TEOR demonstrated superior predictive power with an AUROC of 0.77 (95% CI [0.67–0.89]). By integrating these two key features, the combined model produced the highest performance on the training data, achieving an AUROC of 0.82 (95% CI [0.72–0.92]), indicating a strong predictive signal within the development dataset (Fig. 3C). To facilitate clinical application, this final two-feature model was translated into a nomogram, providing a graphical tool to estimate the probability of postoperative complications (Fig. 3D).

When fixing the pathologic score at 5, changes in FEV1_TEOR are associated with significant variations in postoperative risk: a value of 75% corresponds to an estimated risk of 26.4%, 65% to 37.0%, and 55% to 49.1%. Conversely, when fixing FEV1_TEOR at 65%, variations in the pathologic score also show a clear trend: a score of 4 corresponds to a 31.0% risk, 5 to 37.0%, and 6 to 43.3%. These findings suggest an approximate 6% increase in postoperative risk for each incremental step in pathologic score, and an increase of about 11–12% for every 10% decrease in FEV1_TEOR.

External Validation of Predictive Models

When the trained models were evaluated on the independent external validation cohort, their performance varied. The model based only on the pathologic score failed to generalize, showing no predictive ability beyond random chance (AUROC = 0.50, 95% CI [0.43–0.58]). In contrast, the preoperative FEV1_TEOR was confirmed as a robust and generalizable predictor (Fig. 4A), with the spirometry-only model achieving an AUROC of 0.72 (95% CI [0.66–0.79]) on the external dataset. The combined model, integrating both pathologic score and FEV1_TEOR, also demonstrated valid predictive performance (Fig. 4B), with an AUROC of 0.68 (95% CI [0.61–0.75]), although this was slightly lower than the performance of the FEV1_TEOR-only model.

Analysis of Specific Complication Types

To investigate the prediction of specific adverse events, the modeling pipeline was re-applied to develop distinct models for cardiovascular and respiratory complications. Using the previously identified optimal features, we trained and validated models specifically for these endpoints. In the external validation cohort, similar to the overall complication analysis, the pathologic score alone showed no predictive power for these specific endpoints. Whereas, model using only preoperative FEV1_TEOR (Supplementary Fig. 4A) proved to be effective for predicting both cardiovascular events (AUROC = 0.76, 95% CI [0.64–0.87]) and respiratory events (AUROC = 0.69, 95% CI [0.61–0.77]). Finally, the combined model, incorporating both FEV1_TEOR and pathologic score (Supplementary Fig. 4B), also remained predictive for both sub-endpoints, though its performance was somewhat attenuated, achieving an AUROC of 0.69 (95% CI [0.56–0.81]) for cardiovascular and 0.62 (95% CI [0.53–0.71]) for respiratory complications. To facilitate clinical use, nomograms derived from these combined models are provided for estimating the probability of cardiovascular and respiratory events (Supplementary Fig. 4C).

DISCUSSION

The present analysis, based on LANTERN data, focused on a still not well investigated topic, the prediction of post operative complications in patients candidate to thoracic surgery for NSCLC with curative intent with ML algorithms.

The LANTERN project was ideated [4] with the purpose to create personalized models of lung cancer treatment through the collection of a great amount of patients data (big data) in digitalized human

avatar (DHA). Using an artificial intelligence–based approach, we were able to include a large number of variables (over one hundred), which are typically excluded from analyses relying on traditional statistical methods due to limited data collection or intrinsic methodological constraints.

Only few studies already described the AI application for post-operative complication prediction. Matar and coworkers [18], trained a machine learning model to create a nomogram to predict osteonecrosis after fibula free flap reconstruction in oral cancer patients. In their model, pre-operative radiation therapy, post-operative wound infection, plate exposure, and surgical re-exploration were the most influential features in model predictions. Another interesting experience comes from Wang and colleagues [19]. They focused on a machine learning-based model to predict of gastroparesis risk following complete mesocolic excision. They found this complication to be associated with advanced age, prolonged surgeries, extensive intraoperative blood loss, surgical techniques, low serum protein levels, anaemia, diabetes, and hypothyroidism and the strength of their results stays in the fact that they validated their results with an internal validation.

Also in pediatric surgery ML algorithms have been studied. Liu et al. [20] presented an integrated machine learning model delayed graft function prediction in pediatric renal transplantation from deceased donors.

While the aforementioned studies successfully apply specific ML algorithms, our work highlights the value of a structured machine learning workflow. We systematically tested combinations of feature selection techniques and ML learners to ensure the final model represents the most robust and predictive configuration for identifying patients at risk of postoperative complications.

Concerning instead the prediction of complication specifically in thoracic surgery, some studied formulated analysis based on traditional statistics to create nomograms. For example, Wang et al. [21] created indeed a nomogram, based on four objective and easily assessed factors. It demonstrates excellent predictive performance for pediatric postoperative pulmonary complications after one-lung ventilation, enabling early risk assessment and targeted interventions to improve patient outcomes. Zhou and colleagues [22] led a prospective bicentric study for the elderly patients submitted to thoracic surgery. They developed and validated of a nomogram for predicting postoperative pulmonary complications in this cohort of patients. This nomogram could identify patients at risk for complication within 30 days. Liu and coworkers developed a very similar experience [25] creating a nomogram to predict complications in older patients. They identified preoperative presence of chronic obstructive pulmonary disease (COPD), elevated leukocyte count, higher partial pressure of arterial carbon dioxide ($PaCO_2$) level, surgical site, thoracotomy, intraoperative hypotension, blood loss > 100 mL, surgery duration > 180 min, and malignant tumor to be good prognosticator of complications. A very interesting experience comes from Shixin and colleagues [24], who focused their analysis on post operative complication prediction in patients surgically treated for NSCLC. The population of interest is the same we considered but they developed a nomogram through traditional statistic approach.

Taking into account the prediction of complications in thoracic surgery for NSCLC, with the application of AI, to the best of our knowledge, only 2 experiences are reported. A similar analysis to ours, was reported by Salati et al. [7], but they considered patients submitted to surgery for both malignant and non-malignant conditions and analysed a lower number of variables (n = 50). Furthermore, their one was a retrospective experience, instead ours is a prospective one with the validation on an external cohort. Another one comes from Wand and colleagues [25], who considered patients operated for NSCLC, but then took into account for the analysis, only pneumonectomies. They applied ML models, as multiple machine-learning models (Logistic regression, eXtreme Gradient Boosting, Random forest, Light Gradient Boosting Machine and Naïve Bayes). Multivariate logistic regression analysis revealed that age, surgery duration, intraoperative intercostal nerve block, postoperative patient-controlled analgesia, bronchial blocker use and sufentanil use were independent predictors of cardiovascular and neurologic complications.

Our analysis brings up some novelties if compared to literature. To our knowledge, this is the first study to evaluate a cohort of patients surgically treated for NSCLC with curative intent, encompassing all types of resections performed in this setting—both anatomical and non-anatomical—according to the best available evidence. Furthermore, one interesting aspect, is that more than a hundred variables were considered, of which nearly a half about spirometric parameters. Our results show the weight of the variables in predicting overall complications and the different type of complication with several clinical implications. In particular, pre operative FEV1_TEOR and the combination of Charlson comorbity index and pre operative FEV1_TEOR are the most powerlful factors in predicting post operative complications both in our series and in the external cohort, even if in this cohort with a lower AUROC. Charlson comorbity index alone could predict post operative complication only in our cohort. These results were also confirmed when analysing in particular cardiovascular and respiratory complications.

The external cohort used for validation consisted of surgical patients, and therefore the indications for surgery were comparable to those applied at our center. However, the patient populations differed, as did the distribution of lobectomies and segmentectomies and the overall rate of postoperative complications. This represents a strength of our study, since the reproducibility of the results in an independent external cohort supports the robustness of the proposed model and demonstrates that it can be trained on heterogeneous patient populations while maintaining satisfactory performance.

The observed differences between the two cohorts may be partly attributable to the ongoing process of data reporting within the dataset, which relies on human input and may introduce variability related to the subjective assessment of complications. Indeed, a well-recognized limitation of AI-based analyses is the challenge of reproducibility, as data codification, interpretation, and reporting are inherently dependent on human oversight and may vary across centers, even when the ontology and definitions are prospectively standardized. In addition, population characteristics may differ between institutions, despite our efforts to select cohorts with comparable features in terms of ethnicity, lifestyle habits, and geographical background. From this perspective, the implementation of cloud-based infrastructures enabling continuous and automated data updating will be of paramount importance, as it would allow

models to be continuously retrained, thereby reducing variability, improving reproducibility, and ultimately enhancing their generalizability.In light of the considerations outlined above, the proposed model is not yet ready for routine clinical implementation. In accordance with the objectives of the LANTERN project, prospective validation in our patient cohort will be required once enrollment is completed.

The use of a machine learning approach enabled the integration of a broad range of variables that are frequently excluded from conventional statistical analyses because of data quality issues or methodological limitations. This represents a clear advantage over traditional regression models, as it allows the simultaneous evaluation of multiple clinical features that have been insufficiently explored in relation to postoperative complications, largely due to the constraint of limiting the number of covariates in standard multivariate analyses.

The ultimate goal is the development of an online application (e.g., for smartphones and tablets, already widely adopted in clinical practice) capable of estimating the overall risk of postoperative complications, as well as the specific risk of cardiovascular and respiratory events, based on patient data, including at least the most clinically relevant variables.

In the near future, such a tool could support personalized perioperative management by enabling targeted preventive strategies. For instance, a predicted increased risk of postoperative atrial fibrillation could prompt the initiation of preoperative beta-blocker therapy. More broadly, postoperative intensive care unit (ICU) utilization could be anticipated according to the estimated surgical risk, potentially reducing both healthcare costs and ICU-related complications.

Although cardiovascular and respiratory complication categories encompass events requiring different clinical interventions (e.g., antibiotic therapy for pneumonia versus surgical sealants for air leaks), identifying patients at increased risk for a specific complication group may heighten clinical awareness and facilitate the implementation of tailored preventive measures, such as prophylactic sealant use, early antibiotic treatment, or intensified respiratory physiotherapy.

Overall, this modeling approach has the potential to enhance clinical practice by informing training strategies, optimizing diagnostic and therapeutic decision-making, supporting interventional planning, and improving patient monitoring and follow-up.

CONCLUSIONS

In the field of AI application in surgery, the implementation of a model including preoperative data, in particular, pre-operative FEV1_TEOR and the pathologic score permits to predict post-operative complications, selecting ad hoc strategies for patients management.

Declarations

The study was led in accordance to Helsinki Declaration

Ethics approval and consent to participate: The informed consent has been obtained from all subjects and/or their legal guardian(s). Approval of Fondazione Policlinico Universitario Agostino Gemelli IRCCS – Università Cattolica del Sacro Cuore Ethics Committee Number: 5420 – 0002485/23; Trial registration - clinicaltrial.gov: NCT05802771.

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request

Consent for publication: not applicable

Funding: This project was supported by the Ministry of Health under the frame of ERA PerMed JTC2022

Authors' contributions: FL:Conceptualization, ; visualization: writing-review and editing, surpervision, methodology CS: Conceptualization, writing- original draft preparation, visualization, writing-review and editing , DD and AC: data curation and formal analsysis LB, FG. MB, VP, AB, ET, SL, RA, NF, EO, EM: resources, MC: visualization, Alessandra Cancellieri: resources, investigation RT: resources, EB: resources, validation AG: supersivion, validation, Stefano Margaritora: supervision

Acknowledgements: none

References

1. Bonci EA, Bandura A, Dooley A, Erjan A, Gebreslase HW, Hategan M, Khanduja D, Lai E, Lescaie A, Nitescu GV, Ramalho S, Thiha A, Bertolaccini L. Artificial intelligence in NSCLC management for revolutionizing diagnosis, prognosis, and treatment optimization: A systematic review. Crit Rev Oncol Hematol. 2025;216:104929. Epub ahead of print. PMID: 40967442.
2. Aleem MU, Khan JA, Younes A, Sabbah BN, Saleh W, Migliore M. Enhancing Thoracic Surgery with AI: A Review of Current Practices and Emerging Trends. Curr Oncol. 2024;31(10):6232–44. 10.3390/curroncol31100464. PMID: 39451768; PMCID: PMC11506543.
3. Lococo F, Ghaly G, Chiappetta M, Flamini S, Evangelista J, Bria E, Stefani A, Vita E, Martino A, Boldrini L, Sassorossi C, Campanella A, Margaritora S, Mohammed A. Implementation of Artificial Intelligence in Personalized Prognostic Assessment of Lung Cancer: A Narrative Review. Cancers (Basel). 2024;16(10):1832. 10.3390/cancers16101832. PMID: 38791910; PMCID: PMC11119930.
4. Lococo F, Boldrini L, Diepriye CD, Evangelista J, Nero C, Flamini S, Minucci A, De Paolis E, Vita E, Cesario A, Annunziata S, Calcagni ML, Chiappetta M, Cancellieri A, Larici AR, Cicchetti G, Troost EGC, Ádány R, Farré N, Öztürk E, Van Doorne D, Leoncini F, Urbani A, Trisolini R, Bria E, Giordano A, Rindi G, Sala E, Tortora G, Valentini V, Boccia S, Margaritora S, Scambia G. Lung cancer multi-omics digital human avatars for integrating precision medicine into clinical practice: the LANTERN study. BMC Cancer. 2023;23(1):540. doi: 10.1186/s12885-023-10997-x. Erratum in: BMC Cancer. 2023;23(1):1082. 10.1186/s12885-023-11606-7. PMID: 37312079; PMCID: PMC10262371.

5. He X, Dong M, Xiong H, Zhu Y, Ping F, Wang B, Kang Y. Prediction models for postoperative pulmonary complications in intensive care unit patients after noncardiac thoracic surgery. BMC Pulm Med. 2024;24(1):420. 10.1186/s12890-024-03153-z. PMID: 39210309; PMCID: PMC11360767.
6. Nonomura R, Tabata T, Sasaki T, Mitomo H, Ishibashi N, Sugawara T, Metoki H. Prediction of Postoperative Respiratory Complications after Lobectomy in Lung Cancer Patients with COPD by Quantitative Image Analysis: A Historical Cohort Study. Ann Thorac Cardiovasc Surg. 2022;28(6):411–9. 10.5761/atcs.oa.22-00133. Epub 2022 Oct 7. PMID: 36216576; PMCID: PMC9763719.
7. Salati M, Migliorelli L, Moccia S, Andolfi M, Roncon A, Guiducci GM, Xiumè F, Tiberi M, Frontoni E, Refai M. World J Surg. 2021;45(5):1585–94. 10.1007/s00268-020-05948-7. Epub 2021 Feb 16. PMID: 33594578. A Machine Learning Approach for Postoperative Outcome Prediction: Surgical Data Science Application in a Thoracic Surgery Setting.
8. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med. 2015;13:1.
9. Dindo D, Demartines N, Clavien P-A. Classification of Surgical Complications: A New Proposal With Evaluation in a Cohort of 6336 Patients and Results of a Survey. Ann Surg. 2004;240:205–13.
10. Fernandez FG, Falcoz PE, Kozower BD, et al. The societyof thoracic surgeons and the european society of thoracic sur-geons general thoracic surgery databases: joint standardization ofvariables definitions and terminology. Ann Thorac Surg. 2015;99:368–76.
11. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005;27:1226–38.
12. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Ser B Stat Methodol. 1996;58:267–88.
13. Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. J R Stat Soc Ser B Stat Methodol. 2005;67:301–20.
14. Tin Kam Ho. Random decision forests. in Proceedings of 3rd International Conference on Document Analysis and Recognition vol. 1 278–282. (IEEE Comput. Soc. Press, Montreal, Que., Canada, 1995).
15. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–97.
16. Chen T, Guestrin C, XGBoost:. A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794ACM, San Francisco California USA, (2016). 10.1145/2939672.2939785
17. Zwanenburg A, Löck S. familiar: End-to-End Automated Machine Learning and Model Evaluation. 1.5.0 https://doi.org/10.32614/CRAN.package.familiar (2022).
18. Matar DY, Mackert GA, Matar AY, Chen AC, Panayi AC, Knoedler L, Knoedler S, Yang R, Mady LJ, Kao HK. Artificial Intelligence Outperforms a Nomogram for Osteoradionecrosis Prognostication Following Fibula Free Flap Reconstruction in Oral Cancer Patients. J Stomatol Oral Maxillofac Surg. 2025 Oct 1:102584. doi: 10.1016/j.jormas.2025.102584. Epub ahead of print. PMID: 41043770.

19. Wang W, Yan Z, Zhang Z, Zhang Q, Jia Y. Machine learning-based prediction of gastroparesis risk following complete mesocolic excision. Discov Oncol. 2024;15(1):483. 10.1007/s12672-024-01355-9. PMID: 39331201; PMCID: PMC11436699.
20. Liu XY, Feng RT, Feng WX, Jiang WW, Chen JA, Zhong GL, Chen CW, Li ZJ, Zeng JD, Liu D, Zhou S, Hu JM, Liao GR, Liao J, Guo ZF, Li YZ, Yang SQ, Li SC, Chen H, Guo Y, Li M, Fan LP, Yan HY, Chen JR, Li LY, Liu YG. An integrated machine learning model enhances delayed graft function prediction in pediatric renal transplantation from deceased donors. BMC Med. 2024;22(1):407. 10.1186/s12916-024-03624-4. PMID: 39304842; PMCID: PMC11415997.
21. Wang L, Xiao T, Du Z, Chen T, Pei D, Qu S. Development and validation of a nomogram to pediatric postoperative pulmonary complications following thoracic surgery. BMC Anesthesiol. 2025;25(1):255. 10.1186/s12871-025-03122-x. PMID: 40399778; PMCID: PMC12096645.
22. Zhou Y, Wang H, Lu D, Jiang T, Huang Z, Wang F, Yao Y, Gu Y, Wei W. Development and validation of a nomogram for predicting postoperative pulmonary complications in older patients undergoing noncardiac thoracic surgery: a prospective, bicentric cohort study. BMC Geriatr. 2025;25(1):169. 10.1186/s12877-025-05791-2. PMID: 40082767; PMCID: PMC11905546.
23. Liu J, Xue D, Wang L, Li Y, Liu L, Liao G, Cao J, Liu Y, Lou J, Li H, Yang Y, Mi W, Fu Q. Development and validation of a nomogram for predicting pulmonary complications in elderly patients undergoing thoracic surgery. Aging Clin Exp Res. 2024;36(1):197. 10.1007/s40520-024-02844-1. PMID: 39368046; PMCID: PMC11455794.
24. Ma S, Li F, Li J, Wang L, Song H. Risk factor analysis and nomogram prediction model construction of postoperative complications of thoracoscopic non-small cell lung cancer. J Thorac Dis. 2024;16(6):3655–67. 10.21037/jtd-24-113. Epub 2024 Jun 12. PMID: 38983183; PMCID: PMC11228728.
25. Wang Y, Xie S, Liu J, Wang H, Yu J, Li W, Guan A, Xu S, Cui Y, Tan W. Predicting postoperative complications after pneumonectomy using machine learning: a 10-year study. Ann Med. 2025;57(1):2487636. Epub 2025 Apr 7. PMID: 40193241; PMCID: PMC11980193.

Tables

Tables 1 and 2 are available in the Supplementary Files section.

Figures

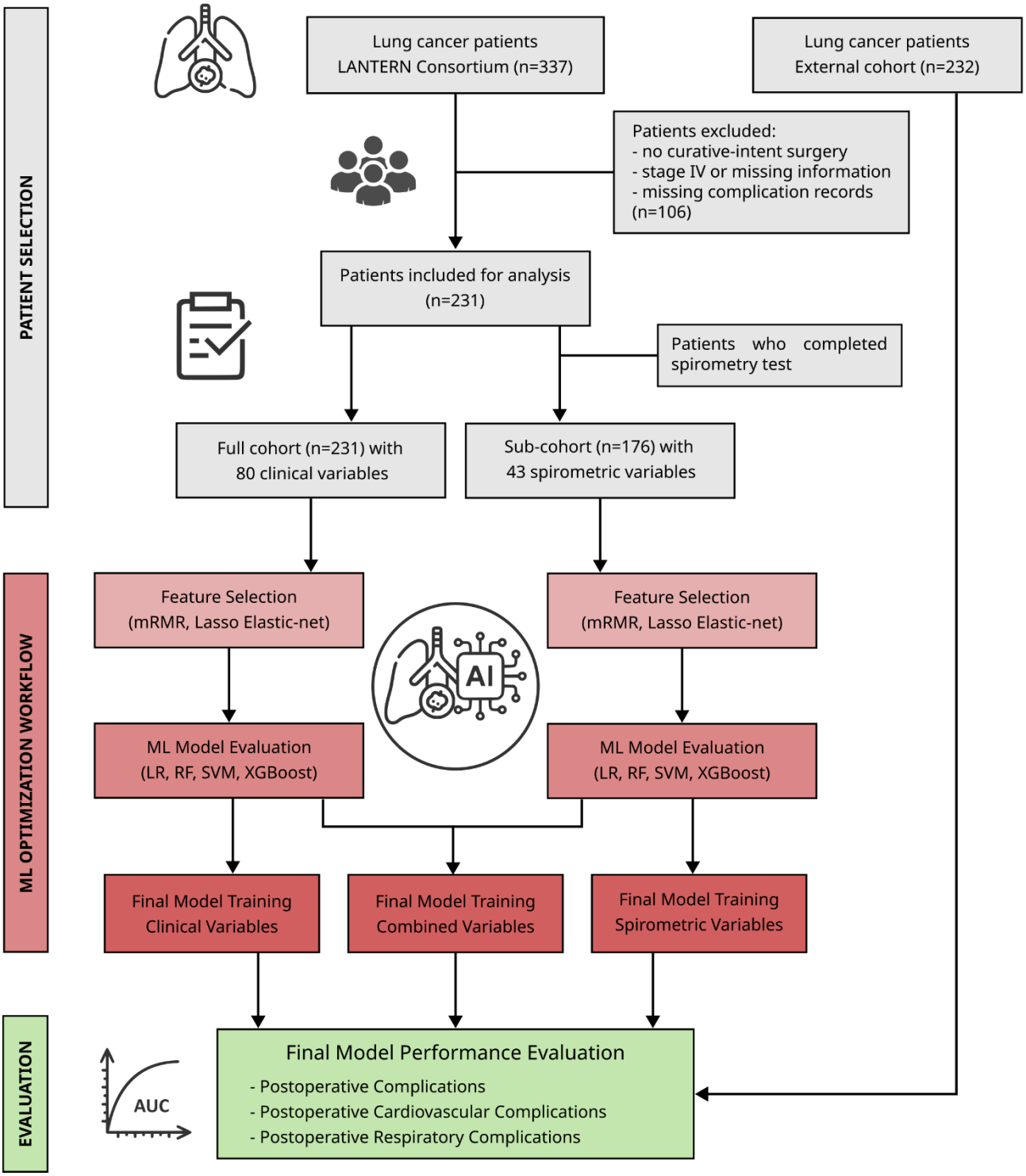


Figure 1

Workflow for Patient Selection and Machine Learning Model Development and Evaluation. The process involves data collection, preprocessing, feature selection, training of different machine learning models and evaluation on an external cohort.

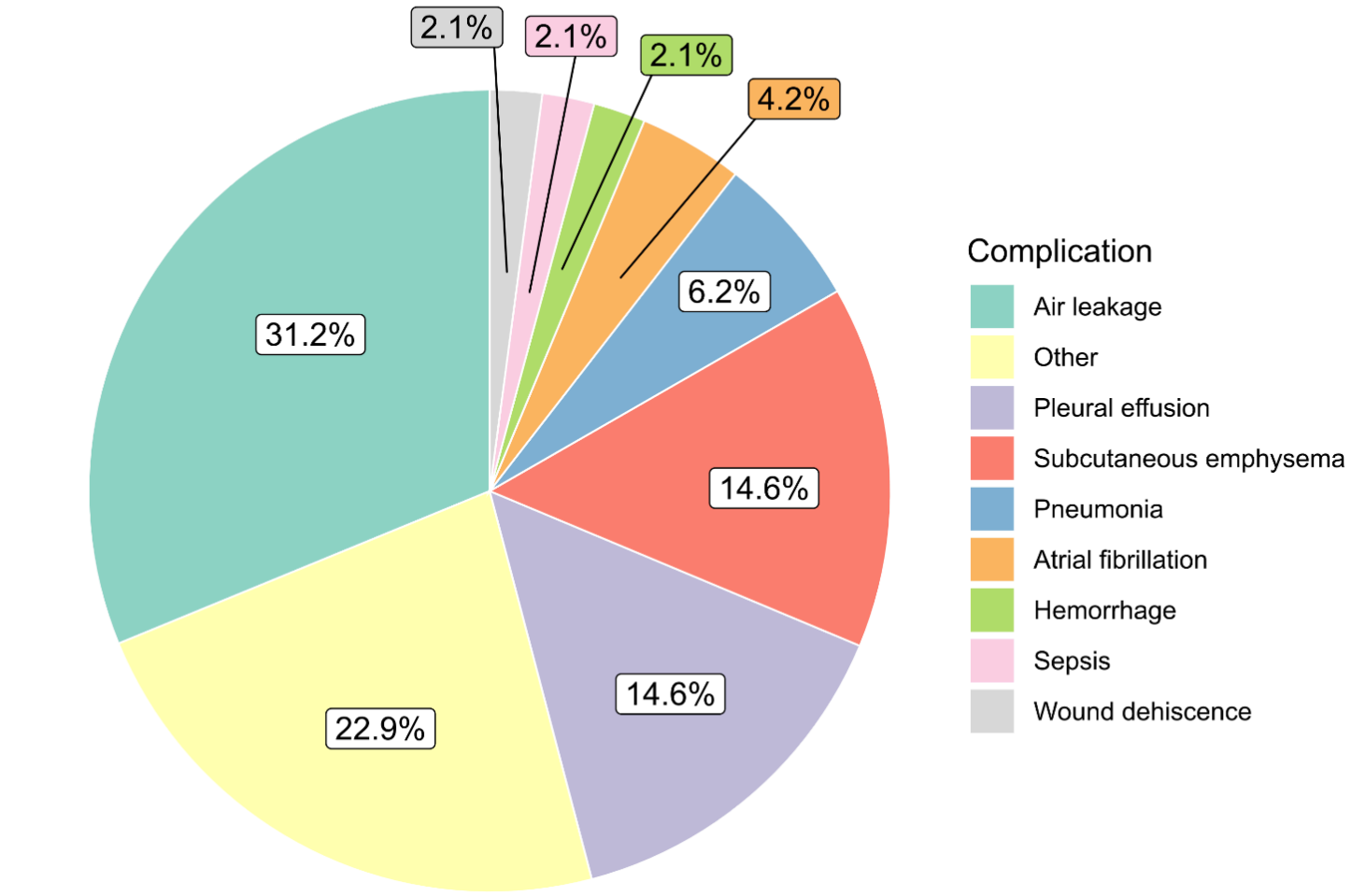
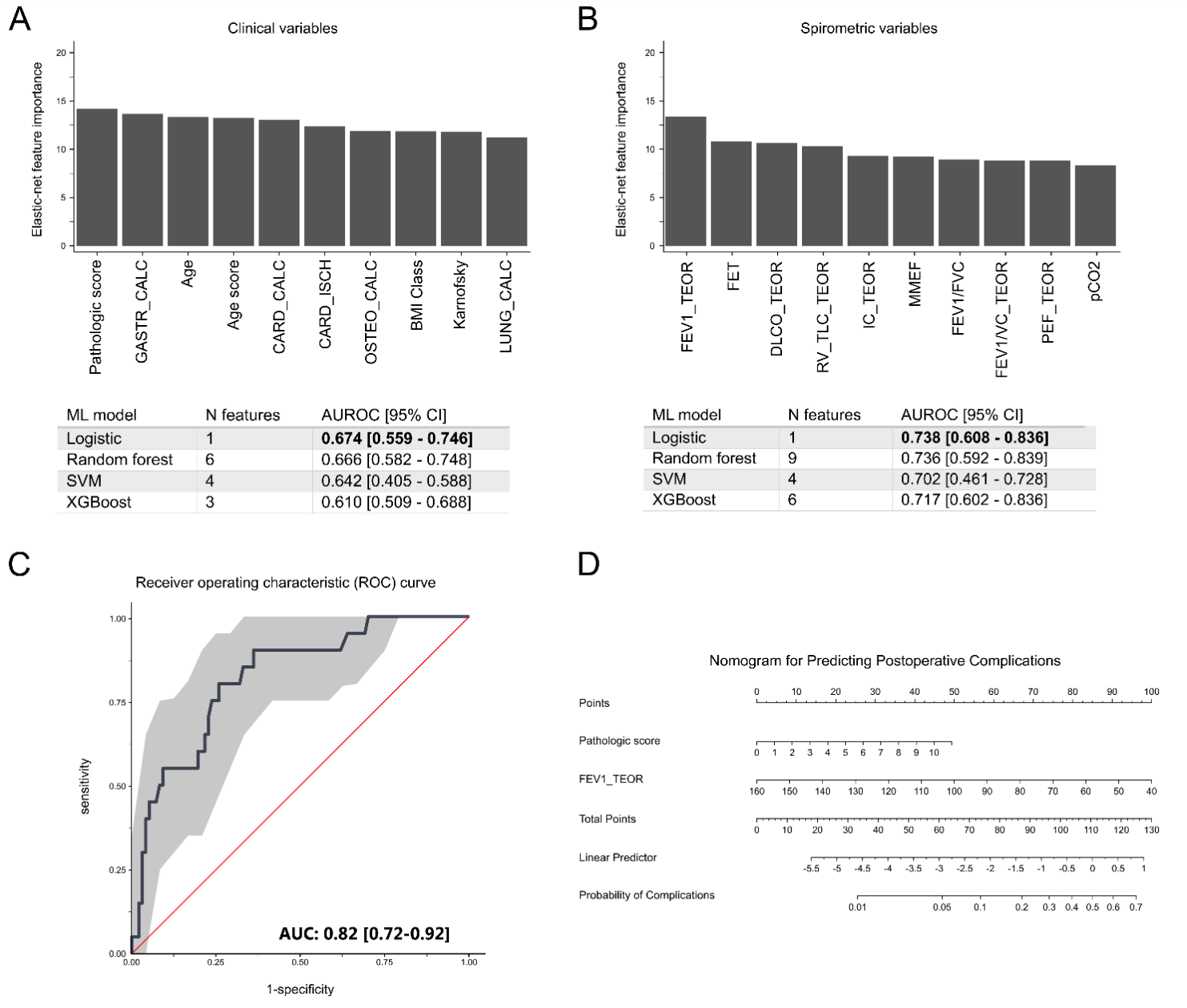


Figure 2

Distribution of post operative complications. The pie chart illustrates the relative frequency of all 48 postoperative complication events.



ML model	N features	AUROC [95% CI]
Logistic	1	**0.674 [0.559 - 0.746]**
Random forest	6	0.666 [0.582 - 0.748]
SVM	4	0.642 [0.405 - 0.588]
XGBoost	3	0.610 [0.509 - 0.688]

ML model	N features	AUROC [95% CI]
Logistic	1	**0.738 [0.608 - 0.836]**
Random forest	9	0.736 [0.592 - 0.839]
SVM	4	0.702 [0.461 - 0.728]
XGBoost	6	0.717 [0.602 - 0.836]

Figure 3

Machine learning workflow and model performance on the training cohort. (A) The bar plot shows the top 10 most important clinical features ranked by elastic net feature importance. The table below displays the peak cross-validated AUROC for four ML models, identifying the Pathologic Score as the single best predictor (AUROC = 0.674) using Logistic Regression. **(B)** The bar plot shows the top 10 most important spirometric features. The corresponding table reports the peak cross-validated AUROC for each model, identifying FEV1_TEOR as the single best predictor (AUROC = 0.738), also via Logistic Regression. **(C)** AUROC for the final model, which combines the Pathologic Score and FEV1_TEOR. The model achieves an AUROC of 0.82 (95% CI [0.72-0.92]). The shaded area represents the 95% confidence interval, and the red diagonal line indicates a random classifier (AUROC = 0.50). **(D)** A nomogram developed from the final logistic regression model to provide a graphical tool for calculating the probability of postoperative complications based on a patient's Pathologic Score and FEV1_TEOR.

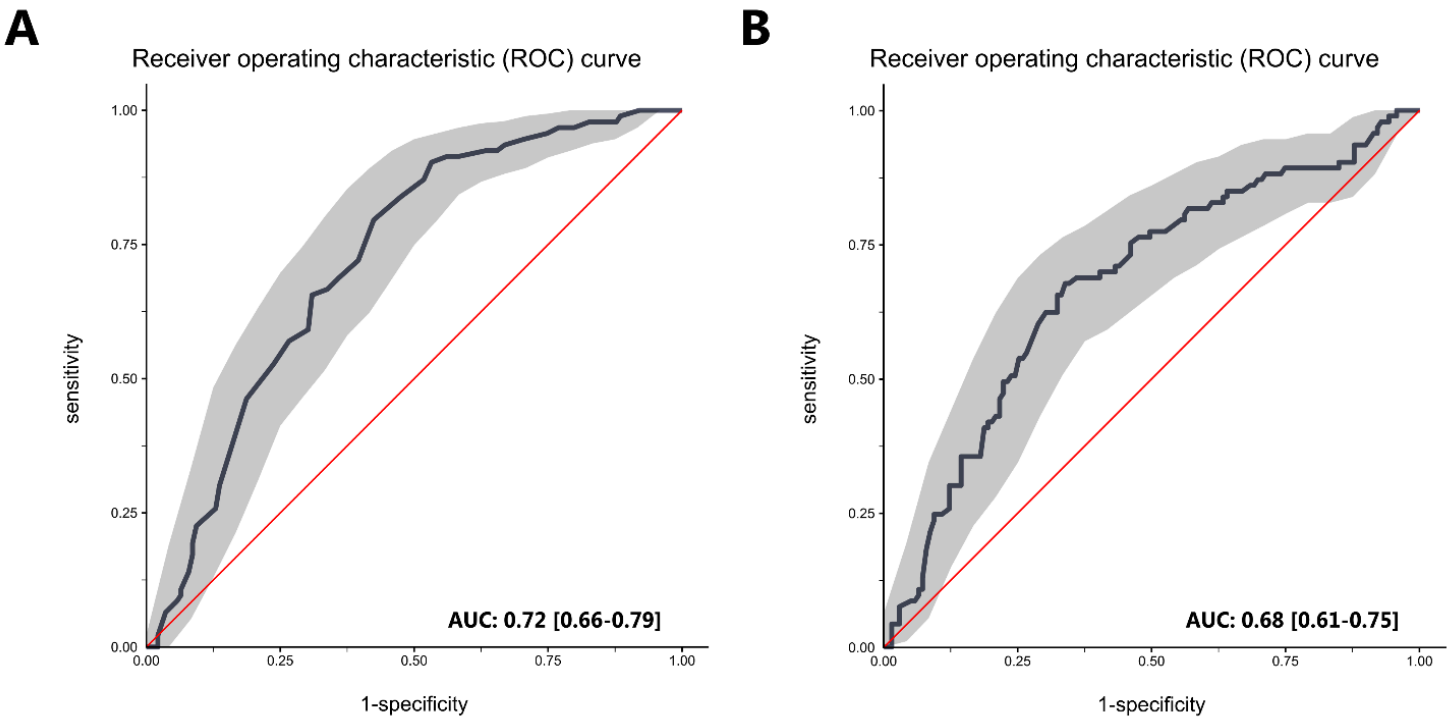


Figure 4

Model performance in the external validation cohort to predict postoperative complications. The panels show the performance, evaluated by Area Under the Receiver-Operator Characteristic (AUROC) curves, of models using: **(A)** FEV1_TEOR alone; and **(B)** the combined Pathology Score and FEV1_TEOR. For all curves, the light grey shading represents the bootstrapped 95% confidence interval, and the red line indicates the performance of a random classifier (AUC = 0.50).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryfiguresBMC.docx
- SupplementarytablesBMC.docx
- Tables.docx