

RESEARCH

Open Access



The role of the diaphragm in prediction of respiratory function in the immediate postoperative period in lung cancer patients using a machine learning model

Radomir Vesovic^{1,2*}, Milan Milosavljevic³, Marija Punt⁴, Jelica Radomirovic^{3,4}, Slavisa Bascarevic^{1,2}, Milan Savic^{1,2}, Vladimir Milenkovic^{1,2}, Marko Popovic² and Maja Ercegovic^{1,2}

Abstract

Background The prediction of postoperative respiratory function is necessary in identifying patients that are at greater risk of complications. There are not enough studies on the effect of the diaphragm on postoperative respiratory function prediction in lung cancer surgical patients. The objective of this study is to estimate the precision of machine learning methods in the prediction of respiratory function in the immediate postoperative period and how diaphragm function contributes to that prediction.

Materials and methods Our prospective study included 79 patients who underwent lung cancer surgery. Diaphragm function was estimated by its mobility measured both ultrasonographically and radiographically and by non-invasive muscle strength tests. We present a new machine learning multilayer regression metamodel, which predicts FEV1 for each patient based on preoperative measurements.

Results The proposed regression models are specifically trained to predict FEV1 in the immediate postoperative period and were proved to be highly accurate (mean absolute error in the range from 8 to 11%). Predictive models based on resected segments give two to three times less precise results. Measured FEV1 was $44.68\% \pm 14.07\%$, $50.95\% \pm 15.80\%$, and $58.0\% \pm 14.78\%$, and predicted postoperative (ppo) FEV1 was $43.85\% \pm 8.80\%$, $50.62\% \pm 9.28\%$, and $57.85\% \pm 10.58\%$ on the first, fourth, and seventh day, respectively. By interpreting the obtained model, the diaphragm contributes to ppoFEV1 13.62% on the first day, 10.52% on the fourth, and 9.06% on the seventh day.

Conclusion The machine learning metamodel gives more accurate predictions of postoperative lung function than traditional calculations. The diaphragm plays a notable role in the postoperative FEV1 prediction.

Keywords Diaphragm, Respiratory function, Machine learning, Lung cancer, Prediction

*Correspondence:

Radomir Vesovic

rvesovic@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Preoperative assessment of postoperative respiratory function is routinely performed in patients with moderate to severe chronic obstructive pulmonary disease (COPD) who are being prepared for lung cancer surgery [1]. The interaction between COPD and lung cancer is common among patients considering that smoking is a great risk factor for both lung cancer and COPD, and that COPD is an important independent risk factor for lung cancer [2]. Postoperatively, COPD patients have lower survival outcomes than non-COPD patients [3]. The assessment of postoperative respiratory function is recommended in identifying patients at greater risk of complications [1]. Predicted postoperative forced expiratory volume in 1 s (ppoFEV1) is most commonly used in postoperative respiratory function and risk factor estimation. PpoFEV1% calculation methods based on removed segments, subsegments, or functional segments are still widely used given their simple application in clinical practice [1, 4–6].

However, there is often a certain difference between predicted and postoperatively measured respiratory parameters. It has been shown that there is a strong correlation between ppoFEV1% and measured FEV1% in the 3- to 6-month period after surgery, while in the immediate postoperative period, when most of the complications arise, the correlation is weak. In the first postoperative days, the difference between predicted and measured FEV1% can reach 30% [7].

The role of the diaphragm in postoperative respiratory function prediction and in postoperative respiratory complications in lung cancer surgical patients has not been significantly studied in literature [8]. There are indirect data, from meta-analysis, about the potential significance of inspiratory muscles strength, which is mostly diaphragm strength, in reducing postoperative pulmonary complications [9]. The potential significance of the diaphragm is also shown by the fact that preoperative diaphragm dysfunction contributes to respiratory complications in cardiac surgery patients, and that it was associated with prolonged mechanical ventilation after lung transplantation [10, 11].

Recently, machine learning models have been used in clinical practice to estimate outcomes and predict postoperative lung function and risk of complications after lung surgery [12–14]. Previous studies have shown that machine learning models potentially have higher accuracy than conventional statistical methods [12].

The aim of this study is to assess the precision of machine learning methods in the prediction of postoperative lung function in the immediate period after surgery and how mobility and strength of the diaphragm contribute to the aforementioned prediction.

Materials and methods

Study design and participants

The prospective cohort study included 79 patients. They had resection performed for primary non-small cell lung cancer at the Clinic for Thoracic Surgery of the University Clinical Center of Serbia from January 2015 to October 2016. Lung resection was done by muscle-sparing thoracotomy with extension towards complete posterolateral thoracotomy when it was necessary.

Inclusion criteria in the study were as follows: full cooperation of a patient while measuring diaphragm movements, proven primary lung cancer and complete assessment of functional status, and overall cardiorespiratory risk.

Exclusion criteria were as follows: poor cooperation of a patient during the required measurements, prior surgery in abdomen and thorax, definitive histopathological findings that indicate another disease, existence of neuromuscular and skeletal diseases, resection of chest wall deemed to be necessary, the presence of massive adhesions observed radiographically preoperatively, and the presence of a ventral hernia of abdominal wall.

Measurements of respiratory function and respiratory muscles strength

Preoperative lung function was measured upon admission to the hospital and classified according to the GOLD criteria [15]. Postoperative respiratory function was measured on the first, the fourth, and the seventh day after surgery. Measurements were taken three times, out of which the highest result was used in the analysis.

Preoperative spirometry and body plethysmography were done with a pneumotachometer (MasterScreen Pneumo, Viasys Healthcare, Germany), while postoperative spirometry was done with a SpiroPRO portable pneumotachometer (Viasys, Germany).

Tests of respiratory muscles strength, maximal inspiratory pressure (P_Imax), maximal expiratory pressure (P_Emax), and sniff nasal inspiratory pressure (SNIP) were completed with MicroRPM (Care Fusion, San Diego, CA, USA) according to ATS/ERS guidelines [16].

Postoperative measurements were taken under the maximal analgesia (NSAID and tramadol-chloride), and analgetics were administered in equal intervals during the first 48 to 72 h and after that upon the patient's request.

These patients had both hemidiaphragms movements measured radiographically and ultrasonographically preoperatively along with respiratory function.

Radiographic measurement

Preoperative radiography was done in the radiology department with a patient in the upright position.

Using standard chest radiography, the distance between the inferior margin of the second rib posteriorly and a horizontal tangent line to diaphragm dome was measured in maximum inspiration (distance a) and maximum expiration (distance b). Preoperative diaphragm movement amplitude (A) was calculated by subtracting the aforementioned distance in expiration (b) from the same distance measured in inspiration (a): $A = a - b$ [8].

The measurement was performed separately for the right and left hemidiaphragms.

Ultrasonographic measurement

Ultrasound measurement of diaphragm movement was done using the Nemio XG MK1 apparatus (Toshiba, Japan).

Once a patient is in a supine, 45° semi-recumbent position, a 3.75-MHz convex transducer was symmetrically placed subcostally between the mid-clavicular and mid-axillary line to obtain a sagittal plane of the hemidiaphragm during every respiratory phase. Upon identification of the right and left hemidiaphragm, two-dimensional (2D) scans were taken, by using a real-time gray scale technology in the sagittal plane that included maximal renal bipolar length. Using 2D images, the hemidiaphragm location was determined relative to the renal pelvis. The point was marked where the hemidiaphragm was observed during craniocaudal excursion at the end of deep expiration. The other point was recorded at the end of the maximum inspiration with the diaphragm lying at the same depth from the transducer on the ultrasound scan. The distance between these points represents the diaphragm mobility measured by ultrasound technique [8].

These measurements were taken three times, for each hemidiaphragm, out of which the best value was used in the analysis.

Estimation of respiratory function

Predicted postoperative FEV1% (ppoFEV1%) was calculated as follows:

- By using the formula developed by Nakahara (N) and associates [5]:

$$PpoFEV1\% = [1 - (n - a)/(42 - a)] \times \text{preoperative FEV1\%} \tag{1}$$

where (n) is the number of resected subsegments in the lobe, that is, 6, 4, and 12 for right upper, middle, and lower lobe and 10 for left upper and lower lobe, while (a) is the number of subsegments obstructed by a tumor.

- By using the Juhl-Frost (JF) formula [4]:

$$PpoFEV1\% = \text{preoperative FEV1\%} \times (1 - 0.0526 \times \text{number of resected lung segments}) \tag{2}$$

which means for upper right lobe 3 segments, middle 2 segments, left upper lobe 4 segments, and lower lobes 5 segments.

- ERS/ESTSa guidelines formula based on removal on functional segments (FS) [1]:

$$PpoFEV1\% = \text{preoperative FEV1\%} \times (1 - a/b) \tag{3}$$

where (a) is the number of obstructed segments that are to be resected and (b) is the total number of unobstructed segments.

In pain level assessment, the visual analogue scale was used.

The study was conducted in accordance with the World Medical Association Declaration of Helsinki and in accordance with the relevant guidelines and regulations. This study protocol was reviewed and approved by the Ethics Committee of the University of Belgrade, Faculty of Medicine, approval number 29/XII-10. All subjects gave written informed consent before participation. The authors received no financial support for this research.

Model training and analysis

A dataset $(X_i, y_i), i = 1, 2, \dots, N$ where N is the number of patients who underwent surgery is available. The vector X_i represents the M -dimensional vector of features, measured preoperatively, while the variable y_i is the measured value of the postoperative FEV1% which is standard indicator of respiratory function.

Given that we have the set $(X_i, y_i), i = 1, 2, \dots, N$, we will create the best possible predictive model within the machine learning methodology.

$$\hat{y}_i = F(X_i) \tag{4}$$

Since y is a continuous variable, (4) is actually a general nonlinear regression model. The first problem to be solved is identification of mapping $F(X_i)$ that gives the highest possible accuracy, measured by the criterion of mean absolute error, MAE:

$$MAE = \sum_{i=1}^N |\hat{y}_i - y_i|. \tag{5}$$

When the regression model (4) is sufficiently accurate, it is possible to estimate the contribution of each of the input variables of the model to the output prediction

of FEV1% [17]. It is interesting to note that this issue has gained actuality within the newly emerging field of study called explainable artificial intelligence [18].

Postoperative FEV1% was measured on the first, the fourth, and the seventh day after surgery. These variables were selected for output variable y in this model (4). This means that a separate regression model is trained for each output variable, so a total of three models. The first 25 characteristics from Table 1 are used as input features. The model selection and construction were executed using the scikit-learn library [19] in the Python programming language. For data analysis, we used the NumPy and Pandas libraries to gain insights into the patients' characteristics [20, 21].

In order to measure the accuracy of our model, we will compare the obtained regression predictive models with existing methods of calculating ppoFEV1% based on preoperative measurements. We limited ourselves to three basic segment counting methods easily available in clinical practice given by (1), (2), and (3), respectively.

Finding the best model was divided into two phases. In the first phase of research, we examined the error of regression models based on individual basic regression algorithms. Table 2 shows the MAE of basic regression models with the default (recommended) hyperparameters for predicting postoperative FEV1% on the seventh day after surgery. Error was calculated by fivefold cross-validation. A lower value represents a better result. The default model hyperparameters are chosen according to the documentation of the used scikit-learn library [22, 23].

In the second phase of the research, regression meta-models were examined. By meta-models, we mean the combination of individual base models in the form of layered structures of a certain depth. The model architecture is shown in Fig. 1. It consists of a combination of previously described basic regression models in two layers, so depth of this meta-model is 2. This step proved to be justified, since the meta-model MAE is lower than that of each individual model.

The MAE of our meta-model is $7.98\% \pm 1.51\%$. This meta-model was also used in prediction of remaining output variables.

Feature importance

For a trained model, it is significant to understand the importance of each feature on the model's accuracy. One of the most developed techniques, which successfully answers this question, is the method based on calculation of the so-called SHAP (SHapley Additive explanations) values for each feature, first introduced in the work of Lundberg and Lee [17].

Benefits of this method are as follows:

- SHAP values can be calculated for a wide class of models, either by exact or approximate methods.
- Global interpretability, which consists in the fact that aggregated SHAP values show how much each feature (predictor) affects output variable (prediction), either positively or negatively

The procedure for calculating SHAP values for a given model consists of the following steps:

1. Form a cross-validation structure of given data.
2. Train the model on the current fold and calculate SHAP values of all features on corresponding test fold.
3. At the end of cross-validation, find mean value for SHAP values of each feature.
4. In order to obtain the cumulative effect of every feature associated with the diaphragm, the SHAP values of those features are summed up (mobility of left and right diaphragm measured radiographically, mobility of left and right diaphragm measured by ultrasound, preoperative PImax%, preoperative Snip%). This step is justified by the additivity of SHAP values.

Results

Patient attributes are shown in Table 1. Most patients were male, 45 (56.96%), while 34 (43.04%) of them were female. Age group was 60.24 ± 7.31 . Sixteen (20.25%) patients had a BMI greater than 30. The majority of the patients, 59 (74.68%), had a lobectomy performed, while 15 (18.99%) patients underwent pneumonectomy, and 5 (6.33%) had bilobectomy. Most of them, 63 (79.74%), were in cancer stages I and II, while 15 (18.99%) were in stage IIIa, and only one patient was in stage IIIb. According to GOLD criteria, 46 (58.22%) patients had COPD, with 39 of them (49.37%) having mild COPD and 6 (7.59%) moderate, and one patient had severe COPD.

Table 3 shows the MAE of the meta-model and traditional calculation methods for ppoFEV1%.

By comparing our machine learning model FEV1% predictions with predictions obtained by the JF model, as being the most precise out-of-segment counting methods, the following results are obtained. The machine learning model is inaccurate $8.24\% \pm 0.93\%$ on the first postop day, $10.56\% \pm 0.87\%$ on the fourth, and $7.98\% \pm 1.51\%$ on the seventh. JF method is inaccurate: $25.87\% \pm 1.79\%$ on the first postop day, $20.03\% \pm 2.32\%$ on the fourth, and $14.32\% \pm 1.64\%$ on the seventh. Error of N and FS methods is shown in Table 3. Comparing the MAE of our meta-model with the most accurate JF method, we can notice that it is lower by two to three times (Table 3).

Table 1 Patient characteristics

No	Characteristic	Values (mean ± std. or frequency (%))
1	Age	60.24 ± 7.31
2	Type of operation on the right lung	Operation is not on the right lung 32 (40.51%) Right upper lobectomy 18 (22.78%) Right bottom lobectomy 14 (17.72%) Right middle lobe lobectomy 3 (3.80%) Bilobectomy superior 2 (2.53%) Bilobectomy inferior 3 (3.80%) Right pneumonectomy 7 (8.86%)
3	Type of operation on the left lung	Operation is not on the left lung 47 (59.49%) Left upper lobectomy 15 (18.99%) Left bottom lobectomy 9 (11.39%) Left pneumonectomy 8 (10.13%)
4	BMI ($\frac{kg}{m^2}$)	Underweight 1 (1.27%) Normal weight 33 (41.77%) Overweight 29 (36.71%) Obesity 16 (20.25%)
5	Type of respiratory function	Normal respiratory function 33 (41.77%) Mild (stage I) 39 (49.37%) Moderate (stage II) 6 (7.59%) Severe (stage III) 1 (1.27%)
6	COPD index = (FEV1% + FEV1%/FVC)/100	1.66 ± 0.22
7	Preoperative FEV1%-preoperative forced expiratory volume in the first second %	94.53 ± 15.63
8	Preoperative VC %-preoperative vital capacity %	109.23 ± 15.87
9	Preoperative FVC %-preoperative forced vital capacity %	107.86 ± 15.04
10	Preoperative VCin %-preoperative vital capacity in inspiration %	105.51 ± 14.20
11	Preoperative FEV1%/FVC	71.32 ± 9.41
12	TLC %-total lung capacity %	116.27 ± 14.41
13	RV %-residual volume %	137.13 ± 34.03
14	FRC (ITGV) %-functional residual capacity %	133.67 ± 27.32
15	RV/TLC (% predicted)	109.49 ± 20.25
16	FRC (ITGV) % (% predicted)	107.51 ± 18.53
17	Mobility of the right hemidiaphragm measured radiographically	4.16 ± 1.41
18	Mobility of the left hemidiaphragm measured radiographically	4.08 ± 1.39
19	Mobility of the right hemidiaphragm measured by ultrasound	68.25 ± 10.28
20	Mobility of the left hemidiaphragm measured by ultrasound	62.58 ± 11.10
21	Preoperative PImax %-preoperative maximal inspiratory pressure %	109.63 ± 35.76
22	Preoperative PEmax %-preoperative maximal expiratory pressure %	92.19 ± 18.04
23	Preoperative snip %-preoperative "sniff" inspiratory pressure	91.83 ± 23.26
24	The number of functional segments removed by the operation	3.41 ± 2.02
25	The number of total functional segments in the lungs	17.23 ± 1.07
26	Sex	Male 45 (56.96%) Female 34 (43.04%)
27	Cancer stage	Ia 14 (17.72%) Ib 13 (16.45%) IIa 26 (32.91%) IIb 10 (12.66%) IIIa 15 (18.99%) IIIb 1 (1.27%)

Table 1 (continued)

No	Characteristic	Values (mean ± std. or frequency (%))
28	Pain after operation	1st day 29.41 ± 10.86
29		4th day 18.85 ± 10.46
30		7th day 11.70 ± 8.02
31	Measured FEV1% after operation	1st day 44.68 ± 14.07
32		4th day 50.95 ± 15.80
33		7th day 58.01 ± 14.78

Continuous variables are represented as mean value ± standard deviation, and categorical variables are represented as frequency and relative frequency

Table 2 Mean value and standard deviation of MAEs models for predicting postoperative FEV1% on the seventh day after surgery, obtained by fivefold cross-validation

No	Model	Values (mean ± std. or frequency (%))
1	Extra tree	8.64 ± 1.43
2	Random forest	9.05 ± 1.47
3	SVM linear	9.05 ± 2.06
4	Lasso	8.05 ± 2.01
5	XGBoost	10.09 ± 2.37
6	Ridge regression	9.98 ± 1.54
7	KNN	9.63 ± 1.69
8	MLP1 (7, 3, 2)	12.23 ± 1.36
9	SVM.RBF	11.38 ± 1.51
10	MLP2 (3, 2)	9.85 ± 2.00
11	LightGBM	8.90 ± 1.02
12	AdaBoost	9.45 ± 1.35

The machine learning model for predicting FEV1% on the first, the fourth, and the seventh postoperative day demonstrates almost identical mean values compared to the measured postoperative FEV1% on the equivalent days (Table 4, Fig. 2). The difference between the two is -0.83% on the first day, -0.33% on the fourth, and -0.16% on the seventh. Discrepancies in standard deviations are indicated in the same Table 4.

We also used segment counting methods that are widely applied in clinical practice, in measuring the difference between the predicted and measured FEV1% in order to be able to compare them with the aforementioned results obtained by our machine learning model. The results are the following: on the first postoperative day, the difference between the predicted and actual FEV1% is 31.23% measured by FS method, 25.23% by JE, and 29.33% by N. On the fourth postop day, those differences are 24.96%, 18.96%, and 23.06%, while on the seventh postop day, they are 17.9%, 11.9%, and 16%, measured by FS, JE, and N methods, respectively. These results suggest that the FS method is the least precise, while the JE is the most precise. Going forward from the first to the seventh day, the FEV1% is recovering, and the measured values are approaching the predicted ones (Table 4, Fig. 2).

Following the steps described for calculating the SHAP values, we get the feature importance representing the impact of various preoperative factors on the ppoFEV1% (Fig. 3).

On the first postoperative day, the diaphragm has the largest impact in the FEV1% prediction, and its mobility has an influence of 9.80%. Combined with the effects of muscle strength tests, PImax 1.0%, and SNIP 2.82% (not shown in figures separately), the total influence of the diaphragm on ppoFEV1% is 13.62%. Other most important factors are spirometry parameters; see Fig. 3a.

On the fourth postoperative day, the role of the diaphragm in ppoFEV1% is slightly less significant, and the

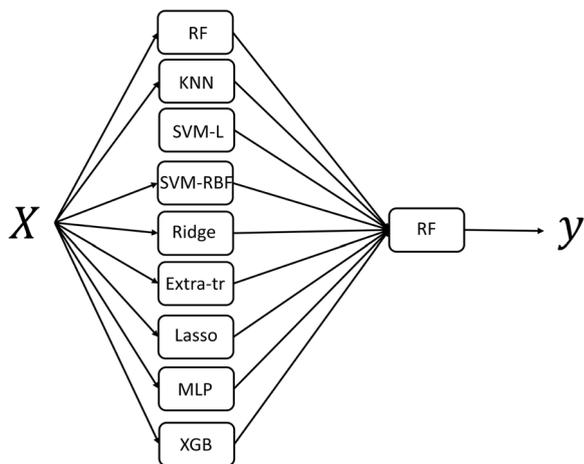


Fig. 1 The architecture of the regression metamodel of depth 2

Table 3 MAE of the model for ppoFEV1%

MAE ppoFEV1%	1st day after surgery	4th day after surgery	7th day after surgery
Our metamodel	8.24 ± 0.93	10.56 ± 0.87	7.98 ± 1.51
Functional segments	31.30 ± 2.44	25.40 ± 2.51	18.86 ± 1.01
Juhl Frost	25.87 ± 1.79	20.03 ± 2.32	14.32 ± 1.64
Nakahara	29.57 ± 2.06	23.65 ± 2.40	17.11 ± 1.25

Table 4 Measured and predicted values of FEV1%

	Mean ± std
FEV1% 1st day	44.68 ± 14.07
FEV1% 4th day	50.95 ± 15.80
FEV1% 7th day	58.01 ± 14.78
ppoFEV1% 1st day	43.85 ± 8.80
ppoFEV1% 4th day	50.62 ± 9.28
ppoFEV1% 7th day	57.85 ± 10.58
Functional segments	75.91 ± 17.40
Juhl Frost	69.91 ± 17.50
Nakahara	74.01 ± 17.21

impact of diaphragm mobility is 6.53% PImax test 2.13% and SNIP test 1.86% (not shown in figures separately), contributes 10.52%, and is the second most important factor in prediction. Respiratory function parameters are the most significant, among the top 10 factors, in

predicting FEV1% on the fourth postop day, as shown in Fig. 3b.

On the seventh postoperative day, the effect of diaphragm mobility is 5.63%, PImax 1.17%, and SNIP 2.26% (not shown in figures separately) which makes the total contribution of the diaphragm 9.06% and is the third most important factor in prediction. The spirometry parameters lead in importance; see Fig. 3c.

Discussion

PpoFEV1% is a parameter that is widely used in postoperative mortality and morbidity assessment and is recommended according to ERS/ESTS guidelines [1]. The loss of pulmonary function is the greatest in the first postoperative week and improves afterwards in the following 6 months when it stabilizes at a lower level than it was preoperatively. After surgery, reduced respiratory function is not only the result of lung volume loss but also impairment in chest wall compliance and lung compliance which is related to the accumulated bronchial

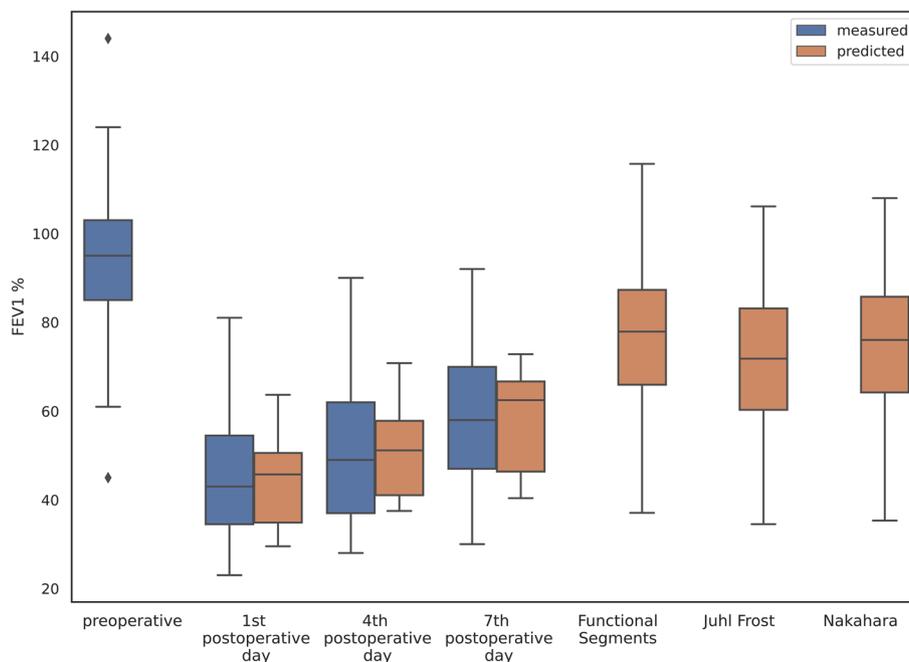


Fig. 2 A boxplot of FEV1% representing the range of measured and predicted values

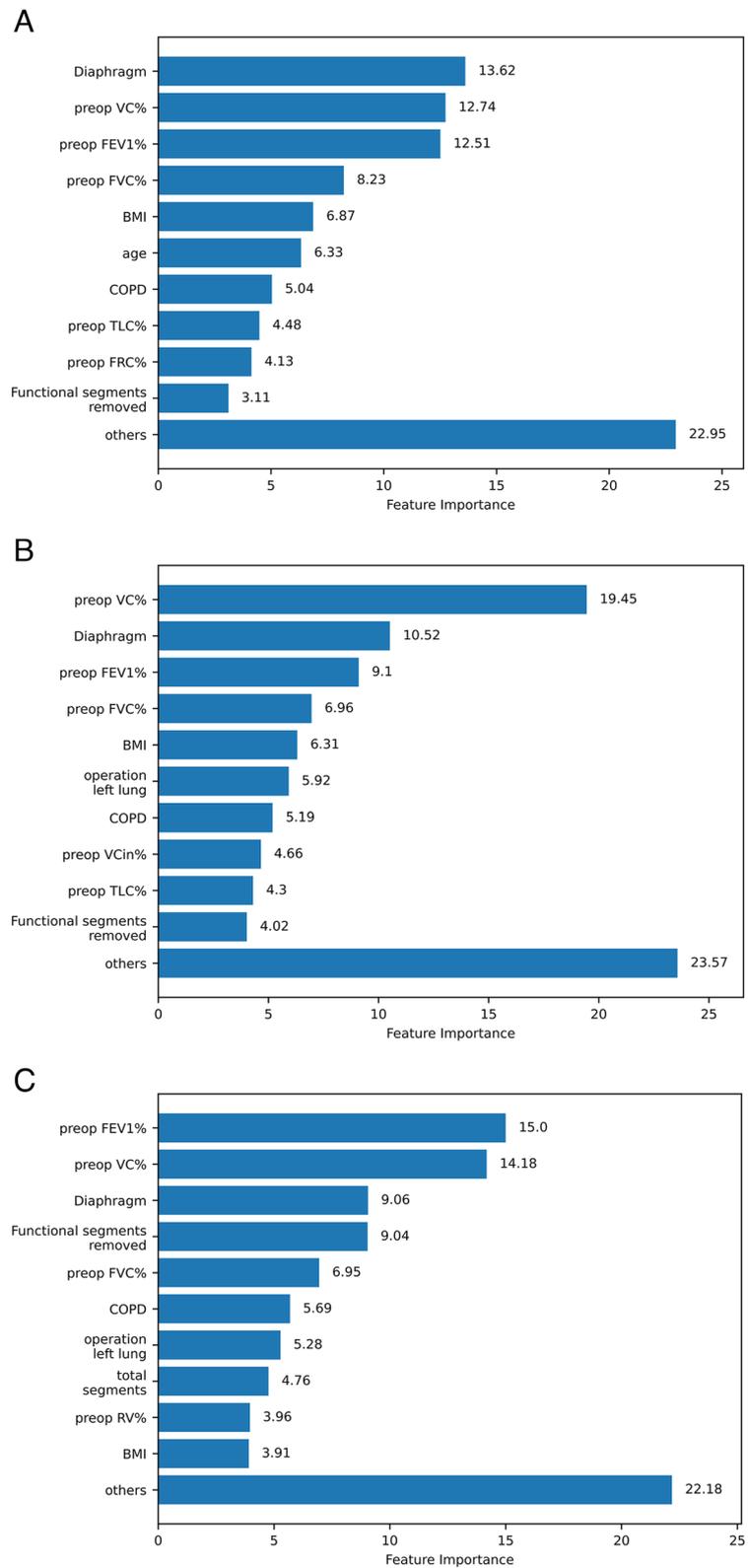


Fig. 3 The top 10 most important features for the metamodel and the output variable ppoFEV1%. **a** The 1st, **b** the 4th, and **c** the 7th day after surgery, in descending order of significance calculated from the SHAP values

secretion, bronchial hyperreactivity, microatelectasis, increased lung water, and reduced surfactant activity. Also, the diaphragm function is eroded. The aforementioned impairments gradually improve after surgery with time [24].

It is presumed that the respiratory function recovery within 30 days of surgery is also a consequence of chest wall surgical injury healing and alleviation of surgical pain [6]. It is previously shown that the lower the pain, the higher the postoperative FEV1% [7]. The machine learning model in our study suggests that surgical pain had no influence on postoperative measured FEV1%.

Almost every potential cardio and respiratory complication occurs immediately after surgery [25–27], and therefore, predicting FEV1% in the immediate postoperative period is important because it is shown that the measured FEV1% on the first postoperative day is more significant in assessing risks of postoperative complications than ppoFEV1% [27].

The measured FEV1% in our patients immediately after surgery correspond to the values published in earlier literature, and their values recover from the first to the seventh postoperative day, and they differ from the ppoFEV1% obtained by segment counting methods [7, 25].

Segment counting methods for predicting FEV1% overestimate the measured FEV1% in the first postoperative days [7], which was noted in our study as well.

Machine learning models have been used in clinical studies to estimate outcomes, predict postoperative lung function and risk of complications after lung surgery, and are shown to be more precise than standard statistical methods [12–14]. The machine learning model that was applied in this study predicts the postoperative FEV1% significantly better (two to three times, measured by MAE) in the immediate postoperative period in comparison to segment counting methods that are widely used in clinical practice [6]. The proposed regression models are specifically trained to predict FEV1% for the 1st, 4th, and 7th day after surgery and were proven to be highly accurate (with a mean absolute error in the range between 8 and 11%).

Meta-analysis that included 17 studies showed that prediction of FEV1% after lung surgery is more precise when computed tomography volume and density measures were combined. However, in every study included in this analysis, respiratory function prediction is performed for the period at least 3 months after surgery, and it is concluded that every available method is imprecise [28].

Methods of postoperative prediction of respiratory function do not take into account the impact of the diaphragm. Several meta-analyses indicated that

preoperative respiratory muscle training could contribute to larger respiratory muscle strength postoperatively, and that would significantly reduce the occurrence of respiratory complications. This benefit is especially observed in older patients, higher risk patients, and thoracic surgery patients [9, 29, 30]. The aforementioned findings indicate the potential significance of inspiratory muscles, the diaphragm primarily, in postoperative prediction of respiratory function.

We demonstrated that the diaphragm plays a big role in postoperative respiratory function prediction in lung cancer surgical patients.

We used noninvasive methods easily available in clinical practice in assessment of the diaphragm function.

Diaphragm function is presented in our study by diaphragm mobility measured both by ultrasound and chest radiography and noninvasive muscle strength tests. Considering that ultrasound and chest radiography capture different aspects of diaphragm mobility, ultrasound measures mobility of its posterior parts, while chest radiography measures mobility of its dome, anterior parts, and diaphragm mobility is represented as the sum of these values. Our machine learning model allowed us to do that by recognizing them as different independent variables in output prediction. It has been shown previously by ultrasound that diaphragm mobility is greater posteriorly than anteriorly and greater laterally than medially which was confirmed by dynamic magnetic resonance imaging at deep breathing [31, 32].

Considering that inspiratory muscle strength tests, PImax, and especially SNIP present diaphragm muscle strength to the largest extent, their values combined with diaphragm mobility represent diaphragm function as a single appropriate parameter [16].

Our study shows that the diaphragm has a great role in the prediction of respiratory function in the immediate postoperative period in lung cancer surgical patients. The diaphragm function impact gradually decreases from the first to the seventh day, while at the same time, the impact of preoperative respiratory function increases, which could be a consequence of chest wall and lung compliance recovery [24].

This prediction model of respiratory function has advantages because it is equally applicable in prediction after lobectomy and pneumonectomy, while both the Nakahara and functional segment methods are not suitable in prediction after pneumonectomy [8].

Our prospective study has several limitations. The main limiting factor in the design of the regression metamodel is the relatively small training set. With its increase, the possibility of more precise adjusting hyperparameters of the base models opens up, which further improves the accuracy of the entire model.

In the analysis, diffusing lung capacity for carbon monoxide (DLCO) was not taken into account. It is recommended to be used in current preoperative patient assessment according to guidelines [1]. Future studies are to include it.

The vast majority of patients in our study had normal respiratory function or mild COPD. Since the study had only one patient with severe COPD, future studies are called for to estimate the results in patients with moderate to severe COPD who need studies like this the most.

Conclusion

The obtained results indicate that the accuracy of the prediction using the metamodel is significantly higher compared to traditional calculations. This is one of the rare analyses that study the effect of the diaphragm on postoperative respiratory function prediction in lung cancer surgical patients. Our analysis unequivocally established a correlation between diaphragm and respiratory function. Using our model and its results, it can be concluded that the diaphragm cannot be ignored and plays a notable part in FEV1% prediction in the immediate postoperative period.

A special contribution of our model is that its improved prediction of respiratory function could contribute to better quality of surgical patient selection.

Acknowledgements

This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, contract number: 451-03-47/2023-01/200103.

Authors' contributions

R. V.: Conceptualization, Methodology, Validation, Investigation, Resources, Writing-Original Draft, Supervision, Project Administration. M. M.: Methodology, Software, Validation, Formal Analysis, Data Curation, Writing-Original Draft, Supervision. M. Punt.: Methodology, Software, Validation, Formal Analysis, Data Curation, Writing-Original Draft, Supervision. J. R.: Methodology, Software, Validation, Formal Analysis, Data Curation, Writing-Original Draft, Visualization. S. B.: Investigation, Resources, Writing - Review & Editing. M. S.: Investigation, Resources, Writing - Review & Editing. V. M.: Investigation, Resources, Writing - Review & Editing. M. Popovic: Investigation, Resources, Writing - Review & Editing. M. E.: Resources, Supervision. All authors have audited and approved the final version of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The datasets used and/or analyzed in the current study are available through the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the World Medical Association Declaration of Helsinki and in accordance with the relevant guidelines and regulations. This study protocol was reviewed and approved by the Ethics Committee of the University of Belgrade, Faculty of Medicine, approval number 29/XII-10. All subjects gave written informed consent before participation.

Competing interests

The authors declare no competing interests.

Author details

¹Faculty of Medicine, University of Belgrade, Dr Subotica 8, 11000 Belgrade, Serbia. ²Clinic for Thoracic Surgery, University Clinical Center of Serbia, Koste Todorovica 26, 11000 Belgrade, Serbia. ³Vlatacom Institute of High Technology, Bulevar Milutina Milankovica 5, 11000 Belgrade, Serbia. ⁴School of Electrical Engineering, University of Belgrade, Bulevar Kralja Aleksandra 73, 11000 Belgrade, Serbia.

Received: 29 September 2023 Accepted: 9 December 2023

Published online: 22 December 2023

References

- Brunelli A, Charloux A, Bolliger CT, Rocco G, Sculier J-P, Varela G, et al. ERS-ESTS clinical guidelines for evaluating fitness for radical treatment (surgery and chemoradiotherapy) in patients with lung cancer. *Eur Respir J*. 2009;34(1):17–41.
- Young RP, Hopkins RJ, Christmas T, Black PN, Metcalf P, Gamble GD. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur Respir J*. 2009;34(2):380–6.
- Zhai R, Yu X, Shafer A, Wain JC, Christiani DC. The impact of coexisting COPD on survival of patients with early-stage non-small cell lung cancer undergoing surgical resection. *Chest*. 2014;145(2):346–53.
- Juhl B, Frost N. A comparison between measured and calculated changes in the lung function after operation for pulmonary cancer. *Acta Anaesthesiol Scand Suppl*. 1975;57:39–45.
- Nakahara K, Monden Y, Ohno K, Miyoshi S, Maeda H, Kawashima Y. A method for predicting postoperative lung function and its relation to postoperative complications in patients with lung cancer. *Ann Thorac Surg*. 1985;39(3):260–5.
- Subotic D. Lung function assessment before anatomical lung resections—is everything so clear?—a narrative review. *AME Med J*. 2022;7:27.
- Varela G, Brunelli A, Rocco G, et al. Predicted versus observed FEV1 in the immediate postoperative period after pulmonary lobectomy. *Eur J Cardiothorac Surg*. 2006;30(4):644–8.
- Subotic DR, Stevic R, Gajic M, Vesovic R. Diaphragm motion and lung function prediction in patients operated for lung cancer—a pilot study on 27 patients. *J Cardiothorac Surg*. 2013;8:213.
- Kendall F, Oliveira J, Peleteiro B, Pinho P, Bastos PT. Inspiratory muscle training is effective to reduce postoperative pulmonary complications and length of hospital stay: a systematic review and meta-analysis. *Disabil Rehabil*. 2018;40(8):864–82.
- Cavayas YA, Eljaiek R, Rodrigue É, Lamarche Y, Girard M, Wang HT, Lev- esque S, Denault AY. Preoperative diaphragm function is associated with postoperative pulmonary complications after cardiac surgery. *Crit Care Med*. 2019;47(12):e966–74. <https://doi.org/10.1097/CCM.00000000000004027>. PMID: 31609771.
- Huh S, Cho WH, Kim D, et al. Clinical impact of preoperative diaphragm dysfunction on early outcomes and ventilation function in lung transplant: a single-center retrospective study. *J Intensive Care*. 2023;10:23. <https://doi.org/10.1186/s40560-022-00614-7>.
- Kwon OB, Han S, Lee HY, Kang HS, Kim SK, Kim JS, Park CK, Lee SH, Kim SJ, Kim JW, Yeo CD. Prediction of postoperative lung function in lung cancer patients using machine learning models. *Tuberc Respir Dis (Seoul)*. 2023;86(3):203–15.
- Bolourani S, Wang P, Patel VM, Manetta F, Lee PC. Predicting respiratory failure after pulmonary lobectomy using machine learning techniques. *Surgery*. 2020;168(4):743–52. <https://doi.org/10.1016/j.surg.2020.05.032>. Epub 2020 Jul 15. Erratum in: *Surgery*. 2021 Apr;169(4):1001. PMID: 32680748.
- Salati M, Migliorelli L, Moccia S, Andolfi M, Roncon A, Guiducci GM, Xiumè F, Tiberi M, Frontoni E, Refai M. A machine learning approach for postoperative outcome prediction: surgical data science application in a thoracic surgery setting. *World J Surg*. 2021;45(5):1585–94. <https://doi.org/10.1007/s00268-020-05948-7>. Epub 2021 Feb 16 PMID: 33594578.
- Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, Stockley RA. Global strategy

- for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187(4):347–65.
16. American Thoracic Society/European Respiratory Society. ATS/ERS statement on respiratory muscle testing. *Am J Respir Crit Care Med*. 2002;166(4):518–624.
 17. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. *Advances in Neural Information Processing Systems 30* [Internet]. Curran Associates, Inc.; 2017. p. 4765–4774. Available from: <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>
 18. Samek W, Wiegand T, Müller K-R. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models [Internet]. arXiv [cs.AI]. 2017 [cited 2023 Aug 22]. Available from: <http://arxiv.org/abs/1708.08296>
 19. Scikit-learn [Internet]. scikit-learn. [cited 2023 Nov 15]. Available from: <https://scikit-learn.org/stable>
 20. NumPy [Internet]. numpy. [cited 2023 Nov 15]. Available from: <https://numpy.org>
 21. Pandas [Internet]. pandas. [cited 2023 Nov 15]. Available from <https://pandas.pydata.org>
 22. Sklearn.Model_selection.KFold [Internet]. scikit-learn. [cited 2023 Aug 22]. Available from: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html
 23. Sklearn.Metrics.Mean_absolute_error [Internet]. scikit-learn. [cited 2023 Aug 22]. Available from: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_absolute_error.html
 24. Miyoshi S, Yoshimasu T, Hirai T, Hirai I, Maebeya S, Bessho T, Naito Y. Exercise capacity of thoracotomy patients in the early postoperative period. *Chest*. 2000;118(2):384–90.
 25. Ercegovic M, Subotic D, Zucic V, et al. Postoperative complications do not influence the pattern of early lung function recovery after lung resection for lung cancer in patients at risk. *J Cardiothorac Surg*. 2014;9:92.
 26. Batchelor TJ, Rasburn NJ, Abdelnour-Berchtold E, Brunelli A, Cerfolio RJ, Gonzalez M, Ljungqvist O, Petersen RH, Popescu WM, Slinger PD, Naidu B. Guidelines for enhanced recovery after lung surgery: recommendations of the Enhanced Recovery After Surgery (ERAS[®]) Society and the European Society of Thoracic Surgeons (ESTS). *Eur J Cardiothorac Surg*. 2019;55(1):91–115.
 27. Varela G, Brunelli A, Rocco G, Novoa N, Refai M, Jiménez MF, Salati M, Gatani T. Measured FEV1 in the first postoperative day, and not ppoFEV1, is the best predictor of cardio-respiratory morbidity after lung resection. *Eur J Cardiothorac Surg*. 2007;31(3):518–21.
 28. Oswald NK, Halle-Smith J, Mehdi R, Nightingale P, Naidu B, Turner AM. Predicting postoperative lung function following lung cancer resection: a systematic review and meta-analysis. *EClinicalMedicine*. 2019;10(15):7–13.
 29. Mans CM, Reeve JC, Elkins MR. Postoperative outcomes following preoperative inspiratory muscle training in patients undergoing cardiothoracic or upper abdominal surgery: a systematic review and meta analysis. *Clin Rehabil*. 2015;29(5):426–38.
 30. de Oliveira VC, Martha BA, Macagnan FE. Effect of inspiratory muscle training associated or not to physical rehabilitation in preoperative anatomic pulmonary resection: a systematic review and meta-analysis. *Support Care Cancer*. 2022;30(2):1079–92.
 31. Laghi FA Jr, Saad M, Shaikh H. Ultrasound and non-ultrasound imaging techniques in the assessment of diaphragmatic dysfunction. *BMC Pulm Med*. 2021;21(1):85.
 32. Takazakura R, Takahashi M, Nitta N, Murata K. Diaphragmatic motion in the sitting and supine positions: healthy subject study using a vertically open magnetic resonance system. *J Magn Reson Imaging*. 2004;19(5):605–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

