Computer Methods and Programs in Biomedicine 238 (2023) 107613

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Optimal machine learning methods for prediction of high-flow nasal cannula outcomes using image features from electrical impedance tomography

Lin Yang[a,1], Zhe Li[b,1], Meng Dai[c,1], Feng Fu[c], Knut Möller[d], Yuan Gao[b,1,*], Zhanqi Zhao[d,1,*]

[a] Department of Aerospace Medicine, Fourth Military Medical University, Xi'an, China
[b] Department of Critical Care Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
[c] Department of Biomedical Engineering, Fourth Military Medical University, Xi'an, China
[d] Institute of Technical Medicine, Furtwangen University, Villingen-Schwenningen, Germany

ARTICLE INFO

Article history:
Received 21 October 2022
Revised 26 April 2023
Accepted 15 May 2023

Keywords:
High-flow nasal cannula
Electrical impedance tomography
Image features
Early prediction
Machine learning

ABSTRACT

Background: High-flow nasal cannula (HNFC) is able to provide ventilation support for patients with hypoxic respiratory failure. Early prediction of HFNC outcome is warranted, since failure of HFNC might delay intubation and increase mortality rate. Existing methods require a relatively long period to identify the failure (approximately 12 h) and electrical impedance tomography (EIT) may help identify the patient's respiratory drive during HFNC.

Objectives: This study aimed to investigate a proper machine-learning model to predict HFNC outcomes promptly by EIT image features.

Methods: The Z-score standardization method was adopted to normalize the samples from 43 patients who underwent HFNC and six EIT features were selected as model input variables through the random forest feature selection method. Machine-learning methods including discriminant, ensembles, k-nearest neighbour (KNN), artificial neural network (ANN), support vector machine (SVM), AdaBoost, xgboost, logistic, random forest, bernoulli bayes, gaussian bayes and gradient-boosted decision trees (GBDT) were used to build prediction models with the original data and balanced data proceeded by the synthetic minority oversampling technique.

Results: Prior to data balancing, an extremely low specificity (less than 33.33%) as well as a high accuracy in the validation data set were observed in all the methods. After data balancing, the specificity of KNN, xgboost, random forest, GBDT, bernoulli bayes and AdaBoost significantly reduced ($p<0.05$) while the area under curve did not improve considerably ($p>0.05$); and the accuracy and recall decreased significantly ($p<0.05$).

Conclusions: The xgboost method showed better overall performance for balanced EIT image features, which may be considered as the ideal machine learning method for early prediction of HFNC outcomes.

1. Introduction

At present, high-flow nasal cannula (HFNC) is widely used to prevent or minimize the duration of invasive mechanical ventilation [1], and recent studies have shown that it improves respiratory drive as well as lung mechanics, enhances CO_2 removal, and reduces 90-day mortality [2,3]. However, not all patients can benefit from HFNC. For instance, Roca's research on patients with severe hypoxemia showed that the failure rate of HFNC treatment was as high as 28% [4]. The mortality rate of patients in ICU with tracheal intubation after more than 48h of HFNC could be higher than that of patients with re-intubation within 48h of HFNC, whereas the success rate of re-extubation could be lower [5]. HFNC may delay the tracheal intubation of patients in some cases, leading to the deterioration of prognosis.

The best-known parameters, such as respiratory rate-oxygenation (ROX) ratio, respiratory rate (RR), oxygenation index, peripheral capillary hemoglobin oxygen saturation (SpO_2), acute physiology score II, the severity of hypoxemia and C-reactive

* Corresponding authors.
E-mail addresses: daimeng@fmmu.edu.cn (M. Dai), rj_gaoyuan@163.com (Y. Gao), zhanqi.zhao@hs-furtwangen.de (Z. Zhao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cmpb.2023.107613

protein level, may help predict HFNC failure. However, these parameters can not directly reflect lung ventilation status and thus a longer period (approximately 12 h) is required [6].

Electrical impedance tomography (EIT) is a novel non-invasive, radiation-free, bedside method for monitoring of ventilation changes related to different lung conditions, which include lung regional recruitment and overdistension during PEEP titration in patients with acute respiratory distress syndrome [7–10]. Recent study has indicated that EIT can help to identify the overdistension caused by HFNC [11]. Besides, EIT can also observe pendelluft and diaphragm activities and monitor the related lung injuries, which may facilitate the identification of patients' respiratory drive [12–14].

In our previous study, we used EIT to observe changes in spatial and temporal ventilation distributions in 46 patients with acute respiratory failure (ARF) during the first hour of HFNC application [15]. The present study aims to propose a proper machine-learning model to predict HFNC within 48h outcomes using image features captured by various EIT indexes, so as to indicate immediate mechanical ventilation treatment in case of invalid HFNC, which is critical for patients to choose sequential oxygen therapy after ventilator weaning, especially for those difficult-to-wean patients.

2. Materials and methods

As shown in Fig. 1, the study aims to adopt the following process: Firstly, collect patient EIT data in clinic continuously; Secondly, analyze and extract patient EIT data features, extract patient basic clinical records (including name, age, height, weight, etc.), and integrate label information based on clinical observation information (including Apache II, complications, initial FiO_2, PaO_2, HFNC intervention outcomes, etc.); Thirdly, after data preprocessing, feature selection method is used to select features from EIT features as model features; Fourthly, integrate EIT data features with label to form sample clusters, then balance the samples based on the feature of HFNC result; Fifthly, Divide the samples into training and testing groups; Sixthly, construct the 48 h HFNC treatment prediction model (hereinafter referred to as the prediction model) by machine learning; Seventhly, testing and comparing the effectiveness of predicted models.

2.1. Data measurement

The data from our previous study were analyzed retrospectively [16]. The study was approved by the ethics committees of Renji Hospital, School of Medicine, Shanghai JiaoTong University (KY2021-057-B). Written informed consent was obtained from all patients or their legal representatives prior to the study.

Patients were treated with HFNC after ICU admission from 2021.05.27 to 2021.06.20 and continuously screened by EIT during underwent HFNC for a period of 1h. Only patients with acute respiratory failure ARF (respiratory rate >25 breaths/min, PaO_2/FiO_2 <300mmHg) were included. Exclusion criteria included age <18 years, pregnancy, and lactation period, weaning from the ventilator, intubation required, tracheotomy, bronchoscopy, absence of commitment to pursue full life support, and any contraindication to the use of EIT. We have to drop 3 HFNC successes samples because in the 2021 paper, the data analysis was conducted for each period for 5 min. However, in the present study, when we analyzed the entire 1 h EIT data, data in 3 patients were with insufficient quality in some parts. The patients had spontaneously breathing in the supine position during the process, at the end of which 11 patients failed HFNC and 32 patients succeeded. The detailed patient information is available from our previous study. An EIT electrode belt with 16 electrodes was placed around the thorax at the 4∼5th intercostal space, and one reference electrode was placed on the abdomen (PulmoVista 500, Dräger Medical, Lübeck, Germany). The frequency and the amplitude of the currents were selected automatically according to the background noise of the measurement environment. The EIT images were reconstructed with the software of the manufacturer (EIT Data Review Tool, Dräger Medical, Lübeck, Germany).

2.2. Feature calculation

As shown in Table 1, several EIT-based indexes were calculated at three time points (before HFNC, T1; 30 min after HFNC commencement, T2; 1h after HFNC termination, T3) as the features of EIT images, which included global inhomogeneity (GI), the center of ventilation (CoV), regional ventilation delay (RVD), rapid shallow breathing index (RSBI), tidal impedance variation (TIV), end-expiratory lung impedance (EELI), minute volume (MV), and inspiration time (T_{insp}). To investigate the changes from the baseline and the line representing HFNC treatment, the differences of the EIT indexes between the time points were calculated and normalized to the values at T1.

EELI, end-expiratory lung impedance; TIV, tidal impedance variation; RVD, regional ventilation delay; CoV, center of ventilation; SD, standard deviation; GI, global inhomogeneity index; RSBI, rapid shallow breathing index; Tinsp, inspiration time; MV, minute volume.

Inspiratory time (T_{insp}) refers to the time interval from the beginning of inspiration to the beginning of exhalation. GI indicates the degree of the heterogeneity of ventilation, which is calculated from the difference between the impedance change of each pixel in the image of the target area and the average impedance change of the whole lung [17,18]. CoV depicts the ventilation distribution influenced by gravity or various lung diseases(relative impedance weighted with a location in the anteroposterior coordinate) [19]. Regional ventilation distribution was obtained by dividing the tidal image into four horizontal, anterior-posterior segments of equal height (regions of interest, ROI), which was denoted as ROIs 1-4 [20]. Changes of EELI compared to T1 (ΔEELI) were also calculated for individual ROIs. RVD index characterizes the regional ventilation delay as pixel impedance rising time compared to the global impedance curve [21], which may be used to assess tidal recruitment/derecruitment.

In addition to these conventional EIT-based indexes, we proposed two parameters relevant to spontaneously breathing patients, namely RSBI and MV.

RSBI is defined as the ratio of the respiratory rate to TV. Since the change in TV can be estimated by the measured impedance, $RSBI_{EIT}$ was calculated as the ratio of the respiratory rate to tidal impedance variation in arbitrary units. where i and N denote the pixel i in the lung area N. Similarly, MV was estimated as the multiplication of the respiratory rate and tidal impedance variation in arbitrary units (MV_{EIT}).

$$MV \approx MV_{EIT} = RR \times \sum_{i=1}^{N} TV_i$$

Inspiration time over expiration time (I:E) was calculated based on the global impedance-time curves.

2.3. Data grouping and preprocessing

A total of 46 features were obtained from EIT images at the three time points and used as predictor variables. The patients were grouped as follows: 30 cases in the training group (22 cases of HFNC success and 8 HFNC failure) and 13 cases in the test group (10 cases of HFNC success and 3 HFNC failure), as shown in Table 1. The outcome was defined as HFNC result with the two groups (success and failure), which were converted to 1 and 0 (Table 2).

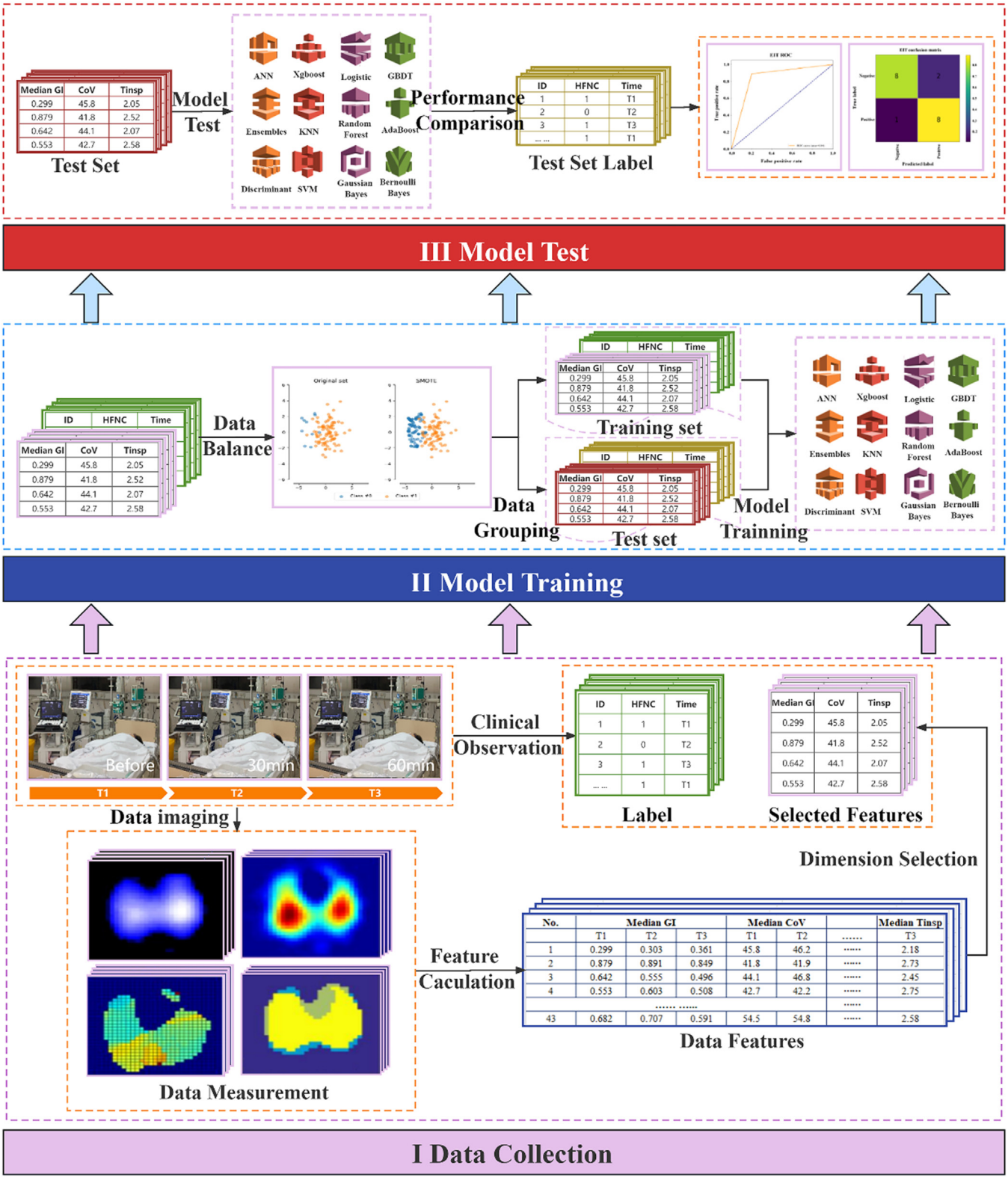

Fig. 1. Schematic diagram of the proposed methodology. The graph can be divided into three parts: data collection, model training, and model test. In data collection section, there are three functional modules as follows: Data measurement: Collect patient EIT data in clinic continuously, which receives and converts the EIT data into image information. Data features: Analyze and extract patient EIT data features from EIT images, can extract patient basic clinical records (including name, age, height, weight, etc.), and integrate label information based on clinical observation information (including Apache II, complications, initial FiO_2, PaO_2, HFNC intervention outcomes, etc.). Dimension selection: Select critical dimensions from image feature data and form sample feature data. These critical features are selected from data features by random forest feature selection method and are coupled with the sample label selected from electronic medical record. The model training section includes three parts as data balancing, data grouping and model training. Data balancing: Integrate EIT data features with label to form sample clusters. The sample is balanced by SMOTE method, then is divided into training and test set in data grouping section. Model training is used to train and optimize the model, which applies the balanced samples to build model and optimize the model through iterative training. In model test section, the test data is applied to the model which have been trained for performance test. Compare with test set label and generates the confusion matrix, AUC curve and other testing tables by the performance comparison function.

Table 1
Summary the EIT-based features.

No.	Name	Features	Parameters
1	EELI	EELI corresponds to the end-expiratory lung volume, the changes of which compared to T1 (ΔEELI) are calculated for individual ROIs, such as EELI_T2-T1_ROI1 indicates the alteration between T2 and T1 in ROI1.	EELI_T2-T1_ROI1, EELI_T2-T1_ROI2, EELI_T2-T1_ROI3, EELI_T2-T1_ROI4
2	TIV	TIV is measured by the difference between the impedance at the end of inhalation and exhalation, can reflect the ventilation status within the ROI.	TIV_T1_ROI1, TIV_T1_ROI2, TIV_T1_ROI3, TIV_T1_ROI4, TIV_T2_ROI1, TIV_T2_ROI2, TIV_T2_ROI3, TIV_T2_ROI4, TIV_T3_ROI1, TIV_T3_ROI2, TIV_T3_ROI3, TIV_T3_ROI4
3	RVD	RVD can be used to assess tidal recruitment/ derecruitment and represents the time heterogeneity of quantifiable regional ventilation, which can be measured by pixel impedance rising time compared to the global impedance curve.	RVD_T1_ROI1, RVD_T1_ROI2, RVD_T1_ROI3, RVD_T1_ROI4, RVD_T2_ROI1, RVD_T2_ROI2, RVD_T2_ROI3, RVD_T2_ROI4, RVD_T3_ROI1, RVD_T3_ROI2, RVD_T3_ROI3, RVD_T3_ROI4
4	CoV	CoV depicts the ventilation distribution influenced by gravity or various lung diseases (relative impedance value weighted with a location in the anteroposterior coordinate).	CoV_T1, CoV_T2, CoV_T3
5	TIV_SD	TIV_SD, the standard deviation of TIV, estimates the ventilation distribution heterogeneity.	TIV_T1_SD, TIV_T2_SD, TIV_T3_SD
6	GI	GI indicates the degree of the heterogeneity of ventilation and is calculated from the difference between the impedance change of each pixel in the image of the target area and the average impedance change of the whole lung.	GI_T1, GI_T2, GI_T3
7	RSBI	RSBI is defined as the ratio of the respiratory rate to TV. Since the change in TV can be estimated by the measured impedance, can be calculated as the ratio of the respiratory rate to tidal impedance variation in arbitrary units.	RSBI_T1, RSBI_T2, RSBI_T3
8	T_{insp}	T_{insp} indicates the time interval from the beginning of inspiration to the beginning of exhalation.	T_{insp}_T1, T_{insp}_T2, T_{insp}_T3
9	MV	MV can be estimated as the multiplication of the respiratory rate and tidal impedance variation in arbitrary units.	MV_T1, MV_T2, MV_T3

Table 2
Sample grouping.

group	total	training group	test group
HFNC success (1)	32	22	10
HFNC failure (0)	11	8	3

Because the ratio of data dimension of the training sample size was greater than 0.1 (approximately 3:1), the dimension of the training set should be reduced according to Johnson Lindenstrauss principle. In this study, the random forest feature selection method was used to rank the original features by computing the random forest scores of importance based on the correlation. First, a decision tree was built and out of bag (OOB) data were selected to calculate the error of out of bag data (err_{OOB1}). Next, the noise interference was added randomly to the characteristics of all samples of OOB data and the error of OOB (err_{OOB2}) was calculated again. Assuming that there were M trees in the forest, the importance of feature X is as follows:

$$p_X = \frac{\sum_{i=1}^{M} (err_{i,OOB2} - err_{i,OOB1})}{M}$$

where $err_{i,OOB1}$ and $err_{i,OOB2}$ denote the error of out of bag data and the error of out of bag data with added noise interference of *i*th tree in the forest, respectively.

Additionally, the Z-score method was employed to standardize the predictor variables to avoid an enormous changing range in the model weights [19], which might lead to the instability of numerical calculation; also, standardization could help accelerate algorithm convergence. Considering the imbalanced classes between the HFNC success and failure (approximately 3:1) samples, the synthetic minority oversampling technique (SMOTE) was adopted to process the predictor variables to balance the two groups.

2.4. Model derivation and validation

The following machine learning methods were the widely used to construct predictive models in the field of medical applications: discriminant, ensembles, k-nearest neighbour (KNN), artificial neural network (ANN), support verctor machine (SVM), AdaBoost, xgboost, logistic, random forest, bernoulli bayes, gaussian bayes and gradient-boosted decision trees (GBDT). Accordingly, these methods were involved in this study as well. Details of the model settings were summarized in the Appendix 1. Kolmogorov Smirnov test was used to test the normality of samples. To explore the influence of data balancing operations on the performance of machine learning methods, accuracy, recall, precision, specificity and area under curve (AUC) in the SMOTE-balanced data sets were compared with those in the imbalanced data sets by employing the Wilcoxon rank sum test in the statistical analysis software SPSS 22 (IBM Software, Armonk, NY, USA). A p value$<$0.05 was considered statistically significant.

3. Results

3.1. Selection of important predictors for the HFNC prediction model

As shown in Fig. 2, the importance of predictor variables was ranked by sorting the top 20 items in a descending order. In this study, the top six items were selected as the important predictor variables for the HFNC prediction model, which were EELI_T2-T1_ROI4, TIV_T1_ROI1, EELI_T2-T1_ROI2, TIV_T2_ROI2, RVD_T1_ROI3 and TIV_T1_ROI2, respectively.

3.2. Performance comparison among machine learning methods for the analysis of imbalanced data against balanced data

Fig. 3 shows accuracy, recall, precision, specificity and AUC of the twelve machine learning methods for the analysis of the imbalanced data sets. An extremely low specificity (less than 33.33%) as well as a high accuracy in the validation data set were observed.

Fig. 4 shows accuracy, recall, precision, specificity and AUC of the twelve machine learning methods for analysis of the SMOTE-balanced data sets. After the data balancing process, the specificity increased from approximately 33.33% to nearly 100.00%.

For all the machine learning methods, accuracy and recall significantly reduced after the data balancing process compared with that before balancing ($p<0.05$), and AUC did not have significantly improved before and after balancing ($p>0.05$). The specificity of KNN, xgboost, Random Forest, GBDT, Bernoulli Bayes and

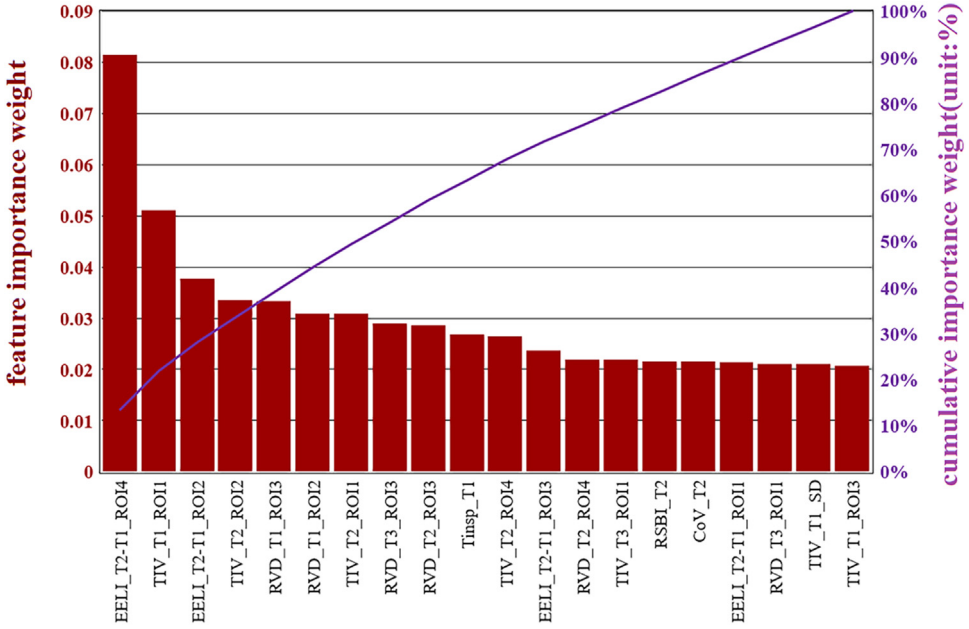

Fig. 2. Important Predictors for the HFNC Prediction Model (Top 20 items).

AdaBoost all significantly improved after the data balancing process ($p<0.05$).

PPV represents the proportion of actual patients who have successfully intervened by HFNC in all judged as success by the prediction method; NPV represents the proportion of actual patients who have failed intervened by HFNC in all judged as failure by the prediction method, which reflect the credibility of the "success" and "failure" in the predicted results. As shown in Table 3, the NPV was all less than 50% in the prediction results of the model trained with imbalanced data. KNN, SVM, and logistic models, the results of "intervention failure in actual was predicted as successful mistakenly " and "predicting the result of intervention as failure correctly" were all 0, resulting in a denominator of 0 in the NPV calculation formula. This reflects the low credibility of the model in predicting patients with potential intervention failure. In application, following the model's prediction may results "blindly" may lead to erroneous judgments of "HFNC being too optimistic"; After data balancing, NPV has achieved an overall improvement, with XGBoost, Random Forest, and Bernoulli Bayes models all exceeding 80% (Table 4). However, the "misdiagnosis" of successful HFNC treatment but failed HFNC treatment in actual is more critical in clinic. Therefore, it is more recommended to use models with higher recall and NPV as Random Forest and GBDT in clinic.

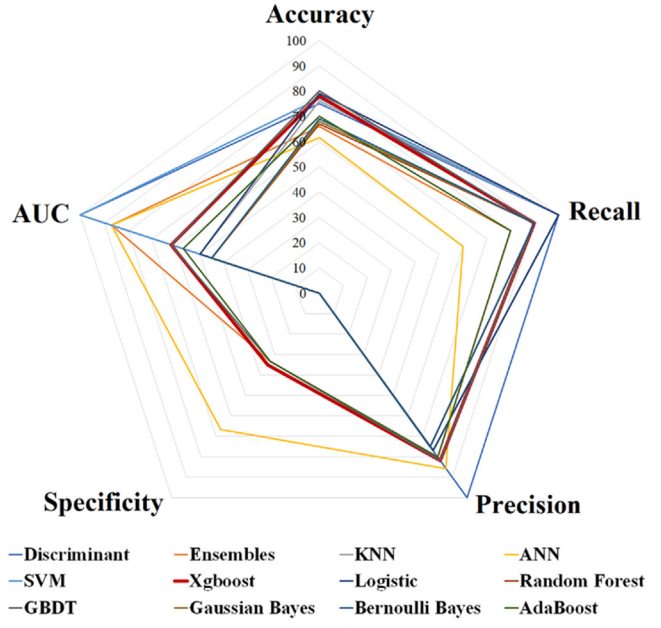

Fig. 3. The performance of discriminant, ensembles, KNN, ANN, SVM, AdaBoost, xgboost, GBDT, logistic, random forest, Gaussian Bayes and Bernoulli Bayes for analysis of the imbalanced data. Xgboost was highlightedfor its best performance.

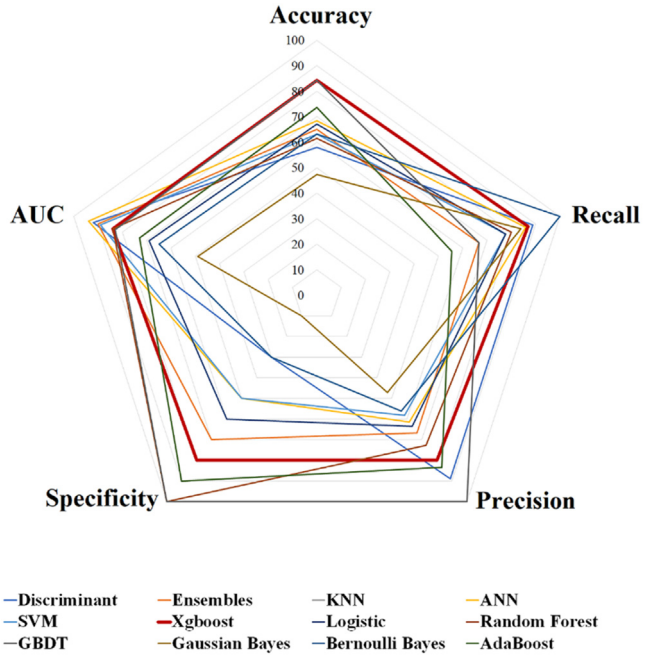

Fig. 4. The performance of discriminant, ensembles, KNN, ANN, SVM, AdaBoost, xgboost, GBDT, logistic, random forest, Gaussian Bayes and Bernoulli Bayse for the analysis of the balanced data. Xgboost was highlighted for its best performance.

Table 3
The overall performance of prediction models for the imbalanced data set.

No.	Model	Accuracy	Recall	Specificity	PPV	NPV
1	discriminant	76.92	100.00	0.00	76.92	NaN
2	ensembles	69.23	80.00	33.33	80.00	33.33
3	KNN	76.92	100.00	00.00	76.92	NaN
4	ANN	61.54	60.00	66.67	85.71	33.33
5	SVM	76.92	100.00	0.00	76.92	NaN
6	XGBoost	76.92	90.00	33.33	81.82	50.00
7	Logistic	76.92	100.00	0.00	76.92	NaN
8	Random Forest	69.23	90.00	0.00	75.00	0.00
9	GBDT	61.54	70.00	33.33	77.78	25.00
10	Gaussian Bayes	69.23	90.00	0.00	75.00	0.00
11	Bernoulli Bayes	69.23	90.00	0.00	75.00	0.00
12	AdaBoost	69.23	80.00	33.33	80.00	33.33

3.3. Performance comparison among the machine learning models

Comparisons of AUCs were also conducted among different machine learning methods. The results indicated that, for the imbalanced data set (Fig. 3), the performance of discriminant, ANN and SVM was better than that of the other methods ($p<0.05$). Ensembles (modeling method: Boost, iterations:300, learning method: discriminant) and ANN (feedforwardnet:3, trainFcn: OneStep Secant Algorithm) had the best validation accuracy of 84.62% and recall of 100%. However, the specificities of all the machine learning methods except ANN (feedforwardnet:3, trainFcn: Levenberg-Marquardt) were less than 35%. For the SMOTE-balanced data set, the specificities of most machine learning methods were higher than 60%.

In general, the specificity performance of the machine learning methods improved greatly after the data balancing process. As shown in Fig. 4, after SMOTE balancing, the specificities of KNN, random forest and GBDT reached the maximum of 100.00%. The overall performance of xgboost method improved considerably, whose accuracy, specificity and AUC were much better than those for the imbalanced data set, as shown in Fig. 5.

4. Discussion

In the present study, we demonstrated the feasibility to predict HFNC outcomes with EIT measurements and machine learning methods. The feature parameters were selected based on EIT image analysis, and the machine learning methods were applied to establish the prediction model based on the feature parameters. The most important six predictor variables for the HFNC prediction model included EELI_T2-T1_ROI4, TIV_T1_ROI1, EELI_T2-T1_ROI2, TIV_T2_ROI2, RVD_T1_ROI3 and RVD_T1_ROI2. The highest sensitivity and specificity were attained with xgboost method.

The prediction of HFNC failure at an early phase is of clinical importance. Roca *et al.* introduced the ROX index to predict HFNC failure, which has been widely used in clinical practice [6]. The ROX index is caculated by indirect indexes representing oxygenation and respiratory drive, which begins to have optimal performance 12h after HFNC initiation. Colleges made efforts to improve the efficiency of the ROX index at an early phase. The values of the modified ROX index with heart rate (ROX/HR*100) above 6.80 were significantly associated with a lower risk of intubation in ARF patients at 10 h after HFNC [20]. Another retrospective study suggested that ROX index values greater than 5.98 and FiO_2 less than 0.59 at 8 h after HFNC treatment was associated with a lower risk of intubation in ARF patients [21]. However, the lack of early and accurate predictive instrument is so far the clinical challenge [22,23]. In our previous work, it was verified that the mROX index [24], which uses PaO_2 instead of SpO_2 to directly respond to oxygen status, could increase sensitivity slightly at 2h. The same finding was obtained by Chen *et al.* in a more recent study, in which the index of Vox, which used Vt instead of RR to indicate respiratory drive, was reported [25]. However, these parameters require invasive blood collection or non-invasive ventilation equipment. We attempted to explore the potential of prediction by EIT to monitor ventilation distributions 1h after HFNC initiation, yet we found that EIT-based spatial and temporal ventilation distribution indexes had slightly and insignificant difference between groups of HFNC success and failure. Notablely, in the present study, it is found that the most important six EIT predictor variables selected by feature weight for HFNC prediction were all at before (T1) and 30 min (T2) after HFNC initiation. These predictors showed high sensitivity and specificity in the machine learning algorithms. These results were consistent with clinical practice that respiratory and oxygen would change a few minutes after the adjustment of respiratory strategies [26,27]. Besides, previous studies have demonstrated that HFNC results in ventilation redistribution within a brief period [3,11]. Consequently, machine learning models are capable of capturing complex features based on multidimensional clinical EIT information during HFNC, and possess the potential to predict outcomes at 30min after HFNC initiation with considerable accuracy.

For the features selected by machine learning, △EELI related predictors EELI_T2-T1_ROI4 contributed the most proportion to the feature important weight (Fig. 2), followed by TIV related predictors TIV_T1_ROI1 and TIV_T2_ROI2, and RVD related predictors RVD_T1_ROI3 and RVD_T1_ROI2. △EELI is associated with changes in end-expiratory lung volume. Mauri *et al.* found that △EELI increased with an increasing flow rate during HFNC [28]. This may be the main pathophysiological benefits of HFNC therapy for ARF patients combined with ARDS, which is the main etiology for the patients in the original study population. △EELI was not comparable among patients treated by conventional methods because the impedance monitored by EIT was not normalized to volume. We presumed that the EELI_T2-T1_ROI4 highlighted the recruitment of collapsed alveolar in the most dependent regions. TIV

Table 4
The overall performance of prediction models for the balanced data set.

No.	Model	Accuracy	Recall	Specificity	PPV	NPV
1	discriminant	60.00	90.00	30.00	56.25	75.00
2	ensembles	70.00	70.00	70.00	70.00	70.00
3	KNN	60.00	70.00	50.00	58.33	62.50
4	ANN	70.00	80.00	60.00	66.67	75.00
5	SVM	65.00	80.00	50.00	61.54	71.43
6	XGBoost	84.21	88.89	80.00	80.00	88.89
7	Logistic	68.42	77.78	60.00	63.64	75.00
8	Random Forest	89.47	77.78	100.00	100.00	83.33
9	GBDT	84.21	66.67	100.00	100.00	76.92
10	Gaussian Bayes	47.37	88.89	10.00	47.06	50.00
11	Bernoulli Bayes	63.16	100.00	30.00	56.25	100.00
12	AdaBoost	73.68	55.56	90.00	83.33	69.23

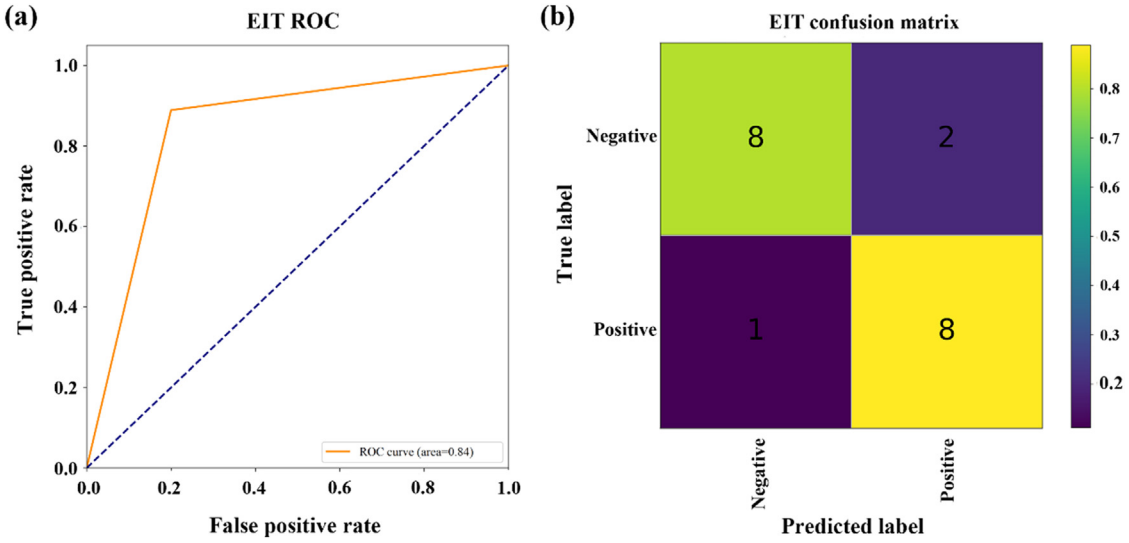

Fig. 5. The performance of xgboost method after SMOTE balancing: (a) ROC curve and (b) Confusion matrix.

at T1 and T2 were also important in predicting the outcome. TIV is related to tidal volume and could represent the respiratory drive during spontentous breath [29,30]. A previous study has indicated that EIT can help to identify the overdistension caused by HFNC [11]. The reduction of respiratory drive is a major benefit of HFNC therapy for ARF patients. When the respiratory drive of patients was too high, the high-flow rate might induce overdistension of not-dependent region [31]. We believed that TIV_T1_ROI1 and TIV_T2_ROI2 were the indicators recognized by machine learning for respiratory drive. In addition, RVD at T1 and T2 were the third relevant parameter. RVD was initially evaluated during low-flow maneuver [32] and could be unstable when the inspiration time is short [18]. Therefore RVD might not be suitable for evaluation of ventilatory status in ARF. Different from the previous study, the machine learning methods in this study demonstrated RVD_T1_ROI3 and RVD_T1_ROI2 as two high weighed features for HFNC outcome. The method of machine learning is helpful to facilitate the exploration of the internal relationship among parameters on the basis of prior knowledge. In practice, it is necessary to select appropriate modeling methods and parameters for the application scenario. For example, the overall performance of the XGBoost model is relatively balanced in test. However, wrong predicting of actual HFNC treatment failure will be more fatal in clinic. Therefore, it is generally more inclined to choose models with better specificity and NPV performance. The manuscript was focused on the feasibility of EIT for predicting HFNC outcomes, and further research will pay more attention to the implementation of the proposed approach.

In the current study, the data were quite imbalanced between the two classes (HFNC success vs. HFNC failure) with a ratio of approximately 3:1. The specificity performance of the machine learning methods was poor with the imbalanced data set. Meanwhile, the specificity performance of these prediction methods improved significantly after the data balancing process, and several methods (such as xgboost) had a considerable improvement, indicating that it was necessary to perform prediction with the balancing method. With data balancing techniques, severely imbalanced classes could be effectively avoided, which is crucial for an accurate forecast. However, accuracy and recall significantly reduced after the data balancing process. The reasons are as follows. First, sensitivity and specificity are two major statistical indicators to evaluate the classification model. Increasing sensitivity tends to reduce specificity and vice versa in general. A low specificity will indicate false positive patients in the clinic, which will delay the intubation time, affect the course of treatment and the prognosis, and even lead to the premature death of patients.On the other hand, a low sensitivity means too many false negative cases, which will waste medical resources and cause unnecessary panic and anxiety. Second, a small number of samples were used in the study, leading to sensitive test results, which demonstrates that the performance of machine learning methods largely depends on robustness and dimensional adaptability. For example, GBDT algorithm is suitable for low dimensional data modeling and has strong data compatibility, and in this study the algorithm showed quite favourable performance, especially specificity. Xgboost also showed excellent performance in the SMOTE-balanced data set in our study due to its high scalability and robustness as well as its advantage of over-fitting reduction.

There are several limitations in this study. The outcome variables in our study were selected by the investigators based on clinical practice, which might bring in subjective bias. In addition, owing to limited data availability, the population included in our study was not large enough. Only 43 samples were obtained because of measurement difficulty. Furthermore, with the development of data balancing techniques, a variety of methods are emerging and we only discussed the SMOTE-balanced method in this study. Limited by the number of samples, we selected only six predictors for the HFNC prediction model. In future studies, we will increase the sample size and select more different hospitals as the validation set to improve the generalizability of the study.

5. Conclusion

The xgboost method showed the best overall performance for balanced EIT image features in our test. It could also support parallelization technology and reduce the risk of over-fitting, which may be a potential machine learning method for early prediction of HFNC outcomes. In further research, we will increase the number of samples, integrate EIT and other related clinical indexes to establish an auxiliary decision-making tool that can be applied to early prediction of HNFC.

Statements of ethical approval

The study was approved by the ethics committees of Renji Hospital, School of Medicine, Shanghai JiaoTong University (KY2021-

057-B). Written informed consent was obtained from all patients or their legal representatives prior to the study.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Lin Yang: Visualization, Data curation, Formal analysis, Writing – review & editing. **Zhe Li:** Visualization, Data curation, Formal analysis, Writing – review & editing. **Meng Dai:** Data curation. **Feng Fu:** Visualization, Data curation, Formal analysis, Writing – review & editing. **Knut Möller:** Writing – review & editing. **Yuan Gao:** Data curation. **Zhanqi Zhao:** Data curation, Writing – review & editing.

Acknowledgment

The study was partially supported by National Natural Science Foundation of China (61901478, 52077216, 51837011, 2022YFC2404800 and 52277235), Natural Science Foundation of Shaanxi Province (2023-YBSF-130) and BMBF MOVE (FKZ 13FH628IX6).

Appendix 1
Parameter settings for the tested methods.

Method	Hyperparemter	Value
KNN	weights:strategy of neighbor weight evaluation	uniform: weight of points in each neighborhood is the same
	algorithm for calculating nearest neighbor	auto:select the most appropriate algorithm based on the samples
	evaluation strategy of nearest neighbor distance	euclidean_distance
	n_estimators	5(initialize value)
SVM	kernel function	RBF
	coefficient of kernel function	1/6
	residual convergence condition	1e-4
	maximum_iterations	unlimited
Logistic	selection of regularization term	L2 regularization
	optimization algorithm for selection parameters	Based on liblinear library, coordinate axis descent method is used as iteratively optimize loss function
	residual convergence condition	1e-4
	regularization degree	1.0
	maximum_iterations	100
Random Forest	max_depth:maximum depth of the tree	Each leaf node has only one category
	n_estimators:number of trees in random forest	8(initialize value)
	criterion: attribute calculation strategy	gini
	min_samples_split	2
	min_samples_leaf	1
	min_weight_fraction_leaf	0.0
	max_leaf_nodes	unlimited
	min_impurity_decrease	0.0
GBDT	measurement method of loss function	deviance
	learning_rate	1e-1
	n_estimators	100
	subsample	boost on full data set by decision tree
	criterion:strategy of sample set segmentation	friedman_mse
	min_samples_split	2
	min_samples_leaf	1
	min_weight_fraction_leaf	0.0
	max_depth	3
AdaBoost	base_estimator	decision tree
	n_estimators: number of base classifier cycles	50
	learning_rate	1e-1
	criteria of model promotion	SAMME.R
GaussianNB	priors: value of priori probability	calculate priori probability according to samples by maximum likelihood method
BernoulliNB	alpha: smoothing factor	1.0
	binarize: threshold of sample feature binarization	0.0
	fit_prior: learn the prior probability of the class or not	true
	class_prior	generate class prior probability from samples
ANN	hidden_layer_num	1
	feedforwardnet	3(initialize value)
	activation	relu
	solver	lbfgs, optimizer based on Quasi-Newton method
	alpha:parameter of regularization item	1e-4
	learning_rate	The initial learning_ rate is 1e-2. When the training loss cannot be reduced for two consecutive times or the verification score stops rising at least tol, divides the current learning rate by 5
	maximum_iterations	200
	tol	1e-4
	trainFcn	levenberg-marquardt
Xgboost	learning_rate	0.1
	n_estimator	10

(*continued on next page*)

Appendix 1 (*continued*)

Method	Hyperparemter	Value
	max_depth	6
	min_child_weight	1
	gama	0
	subsample	1
	colsample_btree	0.8
	scale_pos_weight	1
	max_delta_step	0
	alpha:weight of L1 regularization term	0
Ensemble	base_estimator	decision tree
	modeling method	boost
	n_estimators:number of training base estimators	10
	max_samples:percentage of samples trained for each base estimator	0.4
	max_features:the number of features extracted from the sample for training each base estimator	1.0
	bootstrap	true
Discriminant	solver	svd
	shrinkage	auto, automatic scaling by the Ledoit Wolf lemma
	priors	generate class prior probability from samples
	n_components	1
	tol	1.0e-4

References

[1] W. Alhazzani, M.H. Moller, Y.M. Arabi, M. Loeb, M.N. Gong, E. Fan, S. Oczkowski, M.M. Levy, L. Derde, A. Dzierba, et al., Surviving sepsis campaign: guidelines on the management of critically ill adults with Coronavirus Disease 2019 (COVID-19), Crit. Care Med. 48 (6) (2020) e440–e469.

[2] J. Braunlich, D. Beyer, D. Mai, S. Hammerschmidt, H.J. Seyfarth, H. Wirtz, Effects of nasal high flow on ventilation in volunteers, COPD and idiopathic pulmonary fibrosis patients, Respiration 85 (4) (2013) 319–325.

[3] T. Mauri, C. Turrini, N. Eronia, G. Grasselli, C.A. Volta, G. Bellani, A. Pesenti, Physiologic effects of high-flow nasal cannula in acute hypoxemic respiratory failure, Am. J. Respir. Crit. Care Med. 195 (9) (2017) 1207–1215.

[4] O. Roca, J. Messika, B. Caralt, M. Garcia-de-Acilu, B. Sztrymf, J.D. Ricard, J.R. Masclans, Predicting success of high-flow nasal cannula in pneumonia patients with hypoxemic respiratory failure: the utility of the ROX index, J. Crit. Care 35 (2016) 200–205.

[5] B.J. Kang, Y. Koh, C.M. Lim, J.W. Huh, S. Baek, M. Han, H.S. Seo, H.J. Suh, G.J. Seo, E.Y. Kim, et al., Failure of high-flow nasal cannula therapy may delay intubation and increase mortality, Intensive Care Med. 41 (4) (2015) 623–632.

[6] X. Zou, S. Li, M. Fang, M. Hu, Y. Bian, J. Ling, S. Yu, L. Jing, D. Li, J. Huang, Acute physiology and chronic health evaluation II score as a predictor of hospital mortality in patients of Coronavirus Disease 2019, Crit. Care Med. 48 (8) (2020) e657–e665.

[7] G. Franchineau, N. Brechot, G. Lebreton, G. Hekimian, A. Nieszkowska, J.L. Trouillet, P. Leprince, J. Chastre, C.E. Luyt, A. Combes, et al., Bedside contribution of electrical impedance tomography to setting positive end-expiratory pressure for extracorporeal membrane oxygenation-treated patients with severe acute respiratory distress syndrome, Am. J. Respir. Crit. Care Med. 196 (4) (2017) 447–457.

[8] S. Liu, L. Tan, K. Moller, I. Frerichs, T. Yu, L. Liu, Y. Huang, F. Guo, J. Xu, Y. Yang, et al., Identification of regional overdistension, recruitment and cyclic alveolar collapse with electrical impedance tomography in an experimental ARDS model, Crit. Care 20 (1) (2016) 119.

[9] F. Perier, S. Tuffet, T. Maraffi, G. Alcala, M. Victor, A.F. Haudebourg, K. Razazi, N. De Prost, M. Amato, G. Carteaux, et al., Electrical impedance tomography to titrate positive end-expiratory pressure in COVID-19 acute respiratory distress syndrome, Crit. Care 24 (1) (2020) 678.

[10] G. Scaramuzzo, S. Spadaro, F. Dalla Corte, A.D. Waldmann, S.H. Bohm, R. Ragazzi, E. Marangoni, G. Grasselli, A. Pesenti, C.A. Volta, et al., Personalized positive end-expiratory pressure in acute respiratory distress syndrome: comparison between optimal distribution of regional ventilation and positive transpulmonary pressure, Crit. Care Med. 48 (8) (2020) 1148–1156.

[11] R. Zhang, H. He, L. Yun, X. Zhou, X. Wang, Y. Chi, S. Yuan, Z. Zhao, Effect of postextubation high-flow nasal cannula therapy on lung recruitment and overdistension in high-risk patient, Crit. Care 24 (1) (2020) 82.

[12] H. He, Y. Chi, Y. Long, S. Yuan, I. Frerichs, K. Moller, F. Fu, Z. Zhao, Influence of overdistension/recruitment induced by high positive end-expiratory pressure on ventilation-perfusion matching assessed by electrical impedance tomography with saline bolus, Crit. Care 24 (1) (2020) 586.

[13] Q. Sun, L. Liu, C. Pan, Z. Zhao, J. Xu, A. Liu, H. Qiu, Effects of neurally adjusted ventilatory assist on air distribution and dead space in patients with acute exacerbation of chronic obstructive pulmonary disease, Crit. Care 21 (1) (2017) 126.

[14] T. Yoshida, R. Roldan, M.A. Beraldo, V. Torsani, S. Gomes, R.R. De Santis, E.L. Costa, M.R. Tucci, R.G. Lima, B.P. Kavanagh, et al., Spontaneous effort during mechanical ventilation: maximal injury with less positive end-expiratory pressure, Crit. Care Med. 44 (8) (2016) e678–e688.

[15] Z. Li, Z. Zhang, Q. Xia, D. Xu, S. Qin, M. Dai, F. Fu, Y. Gao, Z. Zhao, First attempt at using electrical impedance tomography to predict high flow nasal cannula therapy outcomes at an early phase, Front. Med. (Lausanne) 8 (2021) 737810.

[16] L. Yang, M. Dai, K. Moller, I. Frerichs, A. Adler, F. Fu, Z. Zhao, Lung regions identified with CT improve the value of global inhomogeneity index measured with electrical impedance tomography, Quant. Imaging Med. Surg. 11 (4) (2021) 1209–1219.

[17] I. Frerichs, G. Hahn, W. Golisch, M. Kurpitz, H. Burchardi, G. Hellige, Monitoring perioperative changes in distribution of pulmonary ventilation by functional electrical impedance tomography, Acta Anaesthesiol. Scand. 42 (6) (1998) 721–726.

[18] L. Yang, M. Dai, X. Cao, K. Moller, M. Dargvainis, I. Frerichs, T. Becher, F. Fu, Z. Zhao, Regional ventilation distribution in healthy lungs: can reference values be established for electrical impedance tomography parameters? Ann. Transl. Med. 9 (9) (2021) 789.

[19] T. Muders, H. Luepschen, J. Zinserling, S. Greschus, R. Fimmers, U. Guenther, M. Buchwald, D. Grigutsch, S. Leonhardt, C. Putensen, et al., Tidal recruitment assessed by electrical impedance tomography and computed tomography in a porcine model of lung injury, Crit. Care Med. 40 (3) (2012) 903–911.

[20] K.J. Goh, H.Z. Chai, T.H. Ong, D.W. Sewa, G.C. Phua, Q.L. Tan, Early prediction of high flow nasal cannula therapy outcomes using a modified ROX index incorporating heart rate, J. Intensive Care 8 (2020) 41.

[21] R.A. Ruiz, B.A. Jurado, F.C. Gueto, A.C. Yuste, I.D. Garcia, F.G. Delgado, J.A.G. Perez, M.L. Obispo, I.Q. Del Rio, F.R. Espinar, et al., Predictors of success of high-flow nasal cannula in the treatment of acute hypoxemic respiratory failure, Med. Intensiva (Engl. Ed.) 45 (2) (2021) 80–87.

[22] J.D. Ricard, O. Roca, V. Lemiale, Use of nasal high flow oxygen during acute respiratory failure, Intensive Care Med. 46 (12) (2020) 2238–2247 Dec.

[23] S. Oczkowski, et al., ERS clinical practice guidelines: high-flow nasal cannula in acute respiratory failure, Eur, Respir. J. 59 (4) (2022) 2101574 vol14 Apr, doi:10.1183/13993003.01574-2021.

[24] Z. Li, C. Chen, Z. Tan, Y. Yao, S. Xing, Y. Li, Y. Gao, Z. Zhao, Y. Deng, M. Zhu, Prediction of high-flow nasal cannula outcomes at the early phase using the modified respiratory rate oxygenation index, BMC Pulm. Med. 22 (1) (2022) 227.

[25] D. Chen, L. Heunks, C. Pan, J. Xie, H. Qiu, Y. Yang, L. Liu, A novel index to predict the failure of high-flow nasal cannula in patients with acute hypoxemic respiratory failure: a pilot study, Am. J. Respir. Crit. Care Med. 206 (7) (2022) 910–913.

[26] S. Rodriguez-Villar, B.M. Do Vale, H.M. Fletcher, The arterial blood gas algorithm: proposal of a systematic approach to analysis of acid-base disorders, Rev. Esp. Anestesiol. Reanim. (Engl. Ed.) 67 (1) (2020) 20–34.

[27] H.C. Gilbert, J.S. Vender, Arterial blood gas monitoring, Crit. Care Clin. 11 (1) (1995) 233–248.

[28] T. Mauri, L. Alban, C. Turrini, B. Cambiaghi, E. Carlesso, P. Taccone, N. Bottino, A. Lissoni, S. Spadaro, C.A. Volta, et al., Optimum support by high-flow nasal cannula in acute hypoxemic respiratory failure: effects of increasing flow rates, Intensive Care Med. 43 (10) (2017) 1453–1463.

[29] K. Vaporidi, E. Akoumianaki, I. Telias, E.C. Goligher, L. Brochard, D. Georgopoulos, Respiratory drive in critically ill patients. Pathophysiology and clinical implications, Am. J. Respir. Crit. Care Med. 201 (1) (2020) 20–32.

[30] R. Tonelli, R. Fantini, L. Tabbi, I. Castaniere, L. Pisani, M.R. Pellegrino, G. Della Casa, R. D'Amico, M. Girardis, S. Nava, et al., Early inspiratory effort assessment by esophageal manometry predicts noninvasive ventilation outcome in

De Novo respiratory failure. A pilot study, Am. J. Respir. Crit. Care Med. 202 (4) (2020) 558–567.

[31] D.L. Grieco, L.S. Menga, D. Eleuteri, M. Antonelli, Patient self-inflicted lung injury: implications for acute hypoxemic respiratory failure and ARDS patients on non-invasive support, Minerva Anestesiol. 85 (9) (2019) 1014–1023.

[32] Z. Zhao, K. Moller, D. Steinmann, I. Frerichs, J. Guttmann, Evaluation of an electrical impedance tomography-based Global Inhomogeneity Index for pulmonary ventilation distribution, Intensive Care Med. 35 (11) (2009) 1900–1906.