# Monitoring regional lung ventilation to predict extubation failure

Vincent Janiak





# Sorbonne Université

École Doctorale Informatique, Télécommunications et Électronique
ED130
Laboratoire d'informatique de Paris 6

# Monitoring regional lung ventilation to predict extubation failure

Thèse de Doctorat

Par Vincent Janiak

Dirigée par Andrea Pinna, Martin Dres
Encadrée par Christophe Marsala, Umar Saleem

Présentée et soutenue publiquement le 6 décembre 2022

Devant un jury composé de :

| | | |
|---|---|---|
| Maria Rifqi | Panthéon-Assas Université | Rapporteuse |
| Damien Bachasson | Institute de Myiologie | Rapporteur |
| Jean-Gabriel Ganascia | Sorbonne Université | Examinateur |
| Alexandre Demoule | APHP | Examinateur |
| Andrea Pinna | Sorbonne Université | Directeur de thèse |
| Martin Dres | APHP | Co-Encadrant |
| Christophe Marsala | Sorbonne Université | Co-Encadrant |
| Hélène Malka-Mahieu | Bioserenity | Examinatrice |

# Acknowledgments

I express my deep gratitude to the members of the jury for having agreed to participate in the evaluation of my thesis work, Jean-Gabriel Ganascia, Alexandre Demoule, Hélène Malka-Mahieu and especially Maria Rifqi and Damien Bachasson for the time devoted to reviewing my thesis manuscript.

I would like to thank my thesis director Andrea Pinna and my co-supervisor Martin Dres, Christpohe Marsala and Umar Saleem for supporting me for my work and morally during those 3 years of thesis, for teaching me the right methods for research, and for keeps me in the right direction to finish this thesis.

I also would like to thank Isabelle Bloch and Alexandre Demoule who have been part of my committee for monitoring my thesis and for giving me valuable help from an external point of view.

My thanks also go toward Bioserenity and Marc Frouin who granted me the possibility to realize this thesis with them. More especially, I would like to thank all my colleagues at the Innovation team at Bioserenity, Sarah-Lise Bounaix, Charles Beye, Corentin Janiaud, Paul Durand-Estebe, Holger Graef, Nicolas Nghe, Paul Roujansky, Mathilde Herouan and Antoine Droneau. And a special thanks to Adrien Foueillasard for helping me at the start of my thesis to better understand the EIT as well as Gwendoline Tallec and Jonas Milstein, which greatly helped, in supporting me morally during those 3 years.

I would like to thank my colleague at the LIP6 laboratory for the help and support, more especially Thomas Garbay and Songlin Li, as we started our thesis the same years.

Finally, I would like to thank my family and my friends Eliot Pothet and Tom Wurtz, for their continuous support throughout my thesis.

# Abstract

Mechanical ventilation is a lifesaving treatment for critically ill patients. When the precipitating reason for which mechanical ventilation has been resolved, the challenge is to safely and promptly separate the patients from the ventilator. In some patients, the separation from the ventilator is poorly tolerated, a process referred to as extubation failure. Extubation failure occurs in 10% to 25% of the patients and is associated with increased morbidity and increased mortality. To prevent the occurrence of extubation failure, noninvasive respiratory supports are employed. Though, clinicians lack reliable monitoring tools to identify the patient that would most likely benefit from preventive measures.

To address this monitoring shortcoming, we want to use the Electrical Impedance Tomography. This monitoring device uses current injection and voltage measurement on electrodes place around the thorax. From this, images that show the conductivity variation are reconstructed. Applied to the thorax, we can monitor the distribution of the ventilation. This technology offers many advantages, it is non-invasive and can be used continuously at the bedside of the patient. With such technology, we want to monitor patients once they are extubated, and attempt to predict as soon as possible the extubation failure.

For this goal, first, we put in place a clinical study called EXIT, in which patients undergoing extubation were monitored with EIT for 48 hours. During the study that lasted 2 years, 37 patients were included, though 2 were not included in the database due to too much noise in the signal.

During the inclusions time, we setup our Electrical Impedance Tomography (EIT) framework, to pre-process the raw EIT data, reconstruct the images and finally, extracting the EIT features from the images. In total, we use 61 EIT features. Most of them are coming from the scientific literature. But after analyzing the first EXIT patients and discussing with clinicians, we introduce 4 new EIT metrics: lung_area, lung_shape, $Flow_{EIT}$, $RSBI_{EIT}$. The last two are based on already existing metrics, used by clinicians that we have adapted to be computed on EIT data. The addition of those features allows to significantly improve the extubation failure's prediction results of our models (+10.8% of sensitivity and +17.9% of specificity for the failure class).

For the learning task, we study 3 different dataset models composed of the EIT data coming from EXIT. The goal of this study is to evaluate the impact of the ventilation variation after extubation on the prediction learning models. Three different inference's algorithms are learned

on each dataset, they are: the Decision tree, the Random Forest and the Support Vector Machine.

Our results show that the EIT offers good prediction capabilities. We are not only able to predict the extubation failure, but we are able to predict them hours before the clinician's re-intubation decision. The sensitivity for the failure class increases throughout the weaning observation. Overall, our failure extubation prediction yields a sensitivity of 0.80 and a specificity of 0.73. This shows the usefulness of the EIT during this crucial period.

**Key words**: Electrical Impedance Tomography, Mechanical ventilation, Weaning, Failure prediction.

# Résumé

La ventilation mécanique est un traitement salvateur pour les patients de réanimation. Lorsque la raison précipitante pour laquelle la ventilation mécanique est résolue, le défi consiste à séparer rapidement et en toute sécurité les patients du ventilateur. Chez certains patients, la séparation du ventilateur est mal tolérée, un processus appelé échec d'extubation. L'échec de l'extubation survient chez 10 % à 25 % des patients et est associé à une morbidité et à une mortalité accrue. Pour prévenir l'apparition d'un échec d'extubation, des supports respiratoires non invasifs sont utilisés. Cependant, les cliniciens manquent d'outils de surveillance fiables pour identifier le patient le plus susceptible de bénéficier de mesures préventives.

Pour répondre à cette lacune de surveillance, nous souhaitons utiliser la tomographie par impédance électrique. Ce dispositif de surveillance utilise l'injection de courant et la mesure de tension sur des électrodes placées autour du thorax. À partir de là, des images montrant la variation de conductivité sont reconstruites. Appliqué au thorax, on peut suivre la répartition de la ventilation. Cette technologie offre de nombreux avantages, elle est non invasive et peut être utilisée en continu au chevet du patient. Avec une telle technologie, nous voulons suivre les patients une fois qu'ils sont extubés, et tenter de prédire le plus tôt possible l'échec de l'extubation.

Pour cet objectif, nous avons d'abord mis en place une étude clinique appelée EXIT, dans laquelle des patients subissant une extubation ont été suivis avec EIT pendant 48 heures. Pendant la durée de l’étude qui a duré 2 ans, 37 patients ont été inclus, bien que 2 ne soient pas inclus dans la base de données en raison d'un trop grand bruit dans le signal.

Pendant le temps des inclusions, nous avons configuré notre cadre de tomographie par impédance électrique (EIT), pour prétraiter les données brutes EIT, reconstruire les images et enfin, extraire les caractéristiques EIT des images. Au total, nous utilisons 61 fonctionnalités EIT. La plupart d'entre eux sont issus de la littérature scientifique. Mais après avoir analysé les premiers patients EXIT et discuté avec des cliniciens, nous introduisons 4 nouvelles métriques EIT : lung_area, lung_shape, $Flow_{EIT}$, $RSBI_{EIT}$. Les deux derniers sont basés sur des métriques déjà existantes, utilisées par les cliniciens, que nous avons adaptées pour être calculées sur des données EIT. L'ajout de ces fonctionnalités permet d'améliorer significativement les résultats de prédiction d'échec d'extubation de nos modèles (+10,8% de sensibilité et +17,9% de spécificité pour la classe d'échec).

Pour la tâche d'apprentissage, nous étudions 3 modèles de jeux de données différents composés des données EIT provenant d'EXIT. Le but de cette étude est d'évaluer l'impact de la variation de la ventilation après extubation sur les modèles d'apprentissage de prédiction. Trois algorithmes d'inférence différents sont appris sur chaque jeu de données, ce sont : l'arbre de décision, la forêt aléatoire et la machine à vecteurs de support.

Nos résultats montrent que l'EIT offre de bonnes capacités de prédiction. Nous sommes non seulement capables de prédire l'échec de l'extubation, mais nous sommes capables de les prédire des heures avant la décision de réintubation du clinicien. La sensibilité pour la classe d'échec augmente tout au long de l'observation du sevrage. Globalement, notre prédiction d'échec d'extubation donne une sensibilité de 0,80 et une spécificité de 0,73. Cela montre l'utilité de l'EIT pendant cette période cruciale.

**Mots clés** : Tomographie d'impédance électrique, Ventilation mécanique, Sevrage, Prédiction d'échec.

**Table of contents**

**Table of figures**

**Table of tables**

# Acronyms

ICU: Intensive Care Unit

MV: Mechanical Ventilation

PSV: Pressure Support Ventilation

Fc: Cardiac frequency

Fr: Respiratory frequency

BMI: Body Mass Index

COPD: Chronic Obstructive Pulmonary Disease

SOFA: Sequential Organ Failure Assessment

Tidal: Represent the lung at end-inspiration minus the lung at end-expiration

Vt: tidal volume

CV: Coefficient of variation

EIT: Electrical Impedance Tomography

EELV: End-Expiratory Lung Volume

EELZ: End-Expiratory Lung Impedance

$\Delta Z$: conductivity variation

CoV: Center of ventilation

GI: Global Inhomogeneity index

RR: Respiratory Rate

Std: Standard deviation

Nb: Number

# 1. Introduction

## 1.1. Background

Positive pressure ventilation has seen a global implementation in the hospital after 1951[1], [2]. During an outbreak of poliomyelitis that occurred in Copenhagen, Dr. Ibsen recognized that patients died from respiratory failure. The use of positive pressure ventilation was then tested. Medical students were recruited to manually insufflate air in the chest of patients with rubber bags. This produced impressive results with a decrease of mortality from 87% previously to 40%[3]. This method was then generalized to all patients in hospitals.

The goal of artificial ventilation was to replace the natural respiratory muscles pump. Nowadays, the classical form of invasive mechanical ventilation is positive pressure ventilation which consists in cyclic insufflation inside the thorax by an external ventilator. The insufflation is permitted by a flow between the ventilator and the lungs, the flow being driven by developing higher pressure outside the thorax. The technological development as well as increased knowledge let to optimized ventilation strategies that now aims to support the patients' ventilation and not replace it[1]. This intends to reduce the possibility to injure the lungs due to the mechanical ventilation.

To help clinicians to better adapt the ventilator assistance to the patient's need, they have at their disposal different monitoring devices. The main one is the ventilator itself, which can measure important metrics such as the airway pressure, respiratory rate and tidal volume. To examine the condition of the lungs, they are 3 methods that are usually available in hospital:

- X-ray: it is a great tool to identify the cause of the ventilation distress. This technology has been used for decades, and thus, the different ventilation problems are easily recognizable by clinicians. Though, it offers low spatial resolution. Plus, it has never been shown that performing X-ray may allow to better adapt the patient's ventilation treatment.
- CT-scan: this method offers high spatial resolution. This gives it a vast range of applications in the medical field to better understand the problem of a patient. For ICU applications, the CT-scan can be used to simply monitor the lung distribution with a good accuracy. Or it could be used to detect overdistended alveolar induces by a ventilator. However, in practice, the use of the CT-scan for ICU[1] patients is limited.

[1] Intensive Care Unit

Indeed, this limitation is due to the patient's transportation to the CT-scan, which can induce complications for the patient.

- Ultrasound: it is used by clinicians to detect ventilation pathology, such as pneumothorax or pulmonary edema. Though, like the X-ray, it has not been proven that this method can help clinicians to set the ventilator. Plus, the results are operator dependent, which can pose a problem to compare results.

In an ICU setting, the ultrasound is the best candidate from those three methods for lung monitoring for two main reasons. The ultrasound can be done at the patient's bedside, and it does not use ionizing radiation. Though as mentioned, the primary used of the ultrasound is to find the cause of the ventilation distress. Therefore, the main tool used by clinicians to guide them to better adapts the ventilation treatment, is the ventilator itself which give vital information about the patient's ventilation state. However, once the patient is extubated from the ventilator, there are no more efficient tools available to monitor the patient's ventilation.

Although, there is another monitoring device that was first published in 1984 by Mr. Barber and Mr. Brown [4], called Electrical Impedance Tomography (EIT). Many improvements to this method had to be realized for it to be commercialized. But it then emerged in the market in 2011, by the company Drager which made it available through his EIT devices called "Pulmovista 500". This method allows to monitor the distribution of the ventilation through impedance measurement. It is non-invasive and non-ionizing, so it can be used continuously at the patient's bedside. It has very good temporal resolution as the frame rate can go up to 50 frames/s. Though, it does not have great spatial resolution like the CT-scan. Most research using the EIT applied to ICU patients, has been done to better guide the clinician during the ventilator treatment.

Though, there are very few studies that focus on guiding the clinician once the patient is extubated. Indeed, when the precipitating reason for which mechanical ventilation has been resolved, the challenge is to safely and promptly separate the patients from the ventilator. In some patients, the separation from the ventilator is poorly tolerated, a process referred to as extubation failure. Extubation failure occurs in 10% to 25% of the patients and is associated with increased morbidity and increased mortality. To prevent the occurrence of extubation failure, noninvasive respiratory supports are employed. Though, clinicians lack reliable monitoring tools to identify the patient that most likely would benefit from preventive measures. Plus, once extubated, clinicians lose all information about the ventilation of the patient that was recorded through the ventilator. They now have to rely on other means to detect if the patient

is at risk of extubation failure. But they lack the ability to monitor in real-time the ventilation state of the patient.

Our goal would be to use the monitoring capability of the EIT after extubation. With those EIT data, we want to build a prediction model that could predict the extubation outcome as soon as possible. Such prediction model could allow clinicians to identify the patient at risk before they experience clinical signs of acute respiratory failure and to administer treatment to prevent extubation failure.

their distress and administer preemptive treatment to avoid it.

This is a CIFRE (Conventions Industrielles de Formation par la Recherche) thesis, this work is based on a partnership between academic laboratories and an industrial partner:

- Academic laboratory:
    - LIP6 is a computer science laboratory part of Sorbonne Université which focuses on modeling and solving fundamental application-driven problems, as well as implementing and validating solutions through academic and industrial partnerships. The research challenges are addressed among four transversal axes: Artificial intelligence and data science; Architecture, systems and networks; Safety, security and reliability and Theory and mathematics of computing. Two teams participate to this work: SYEL (Systèmes Electroniques).

    - INSERM UMRS_1158, Neurophysiologie respiratoire expérimentale et Clinique is a research unit located in the hospital of la Pitié Salpétrière (AP-HP) and work in the department of Respiratory and Critical Care. Their focus is on the relationships between the nervous system and the respiratory system.

- Bioserenity: the core of this company is on remote healthcare. More precisely, they are developing smart textiles embedding medical sensors that would allow monitoring of patients from home.

## 1.2. Thesis organization

The introduction chapter briefly introduces the problematic and provides the scientific rationale for conducting this thesis.

Then in the state of the art, we dive more into the mechanical ventilation treatment, with the risks of the mechanical ventilation, the process to extubate patients and the weaning phase following. We also present some of the main prediction models for the extubation outcome already existing. Those methods entail using data from the ventilator, lung ultrasound or blood analysis. There is only one method that uses the EIT data for such purpose.

In the chapter 3, we detail our clinical study called EXIT, with the protocol that we have implemented, as well the recruited patients.

In the next chapter, we define all the EIT features used in this study. With, in addition, 4 new features that we introduced and tested in order to improve the prediction results.

The chapter 5 relates about the EIT framework built to study and test different prediction models. In total, 3 different estimators were tested with 3 different dataset training models. The results obtained from each model are presented.

In chapter 6 we have tuned the training algorithms and optimized the EIT features number in order to find the best set-up to improve the performances. The results obtained from each optimization are then shown and compared with the firsts ones.

At last, chapter 7 presents the conclusion of this thesis work and bring about future development.

# 2. Context and state of the art

*This chapter aims at describing the current epidemiology, definitions and process referred to as weaning from mechanical ventilation. To that end, the chapter will describe clinical challenge of mechanical ventilation, from its initiation until the readiness to undergo safe and prompt separation. Then, it will review the interests and limits of indices developed to help clinicians to predict the outcome of the patients after extubation. Eventually, the electrical impedance tomography technology will be discussed before to raise the hypothesis of the present project.*

## 2.1. Mechanical ventilation

### 2.1.1. Benefits and risks of mechanical ventilation

Worldwide, millions of patients are admitted to intensive care units (ICU) and require invasive mechanical ventilation (MV). Admission to the ICU is a heavy burden at the collective level in terms of healthcare costs and resources and also at the individual level in terms of functional prognosis and survival. In view of the current Covid-19 pandemic, ICU bed availability has become a subject of crucial importance since it may affect triage decisions and, moreover, the use and relevance of MV has increased at a scale not previously imaginable. MV is usually considered to be a supportive therapy and is often lifesaving. MV increase oxygenation, decreases the work of breathing, helps to reopen or to keep open collapsed alveoli, and improves respiratory mechanics.

However, MV per se can injure the lungs, process referred to as ventilator-induced lung injury [5]. Moreover, MV can also lead to other complications such as ventilator-induced diaphragm dysfunction [6] and brain dysfunction (delirium) [7]. The main reason for which MV is initiated is acute respiratory failure that mainly originates from the community-acquired pneumonia, viruses (such influenza or SARS-CoV-2), exacerbation of chronic obstructive pulmonary disease or acute pulmonary edema. MV is only a supportive treatment and cannot be viewed in any way as a curative treatment. Therefore, in parallel to MV initiation, it is fundamental to address the primary reason that precipitated the need of MV. After the resolution of the acute disease, the priority of clinicians is to separate patients from the ventilator, a process referred to as weaning from mechanical ventilation. It is sometimes a long and complex process when faced with various difficulties, but it is simple for a majority of patients. This simplicity, however, does not imply that it is conducted in an optimal way in such cases. Indeed, discontinuation of ventilatory support and extubation should occur as expeditiously as possible to minimize the iatrogenic consequences of intubation and invasive ventilation, including

infectious complications, complications of bed rest and, importantly, ventilator-induced diaphragm dysfunction [6]. Therefore, the weaning process aims at identifying patients early to separate them from invasive ventilation. (cf. Figure 1).

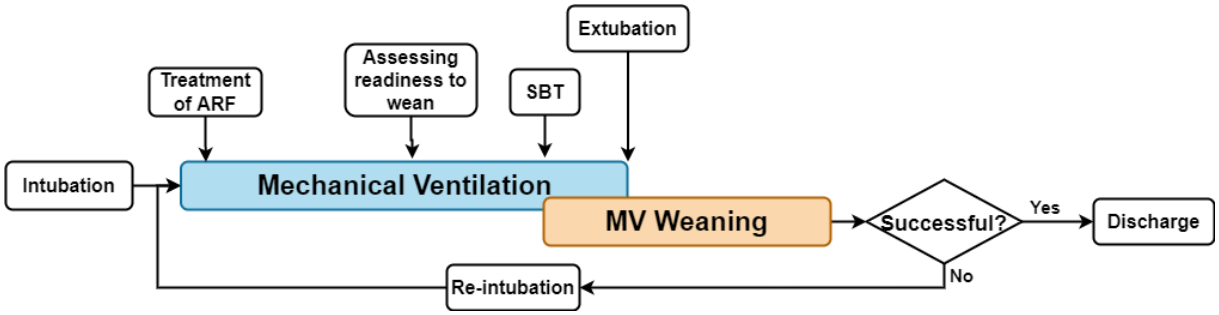


Figure 1 - Different stages a patient goes through once intubated.

### 2.1.2. Weaning from mechanical ventilation

Weaning failure is defined as the inability to separate a patient from the ventilator but actually weaning from mechanical ventilation implies two major milestones. First, the transition from artificial breathing delivered by the ventilator toward natural breathing that is driven by the contraction of the respiratory muscles. Generally, the ability of patients to tolerate the return toward natural breathing is ascertained by a so-called spontaneous breathing trial (SBT) during which the assistance of the ventilator is set to zero to challenge the cardiac and respiratory muscles functions. The second milestone is the ability to control the upper airways (glottis) without the endotracheal tube that connects the patient to the ventilator. Following this, weaning can be viewed as a continuum multi-steps process.

It starts by screening readiness criteria that basically indicate to the clinicians that the patient may be ready to breathe without the assistance of the ventilator. If those criteria are present, it is current practice to “challenge” the respiratory and cardiac system by performing a SBT. If the SBT is successful which means that patients tolerate a period of 30 to 60 minutes with ventilator assistance set to zero or low level of pressure, extubation can be considered. The majority of patients are safely weaned from the ventilator after a first attempt, some are even extubated without any SBT [8], [9]. This group of patients is referred to as “simple weaning”. The other groups are referred to as difficult or prolonged weaning (taking <1 week or >1 week) and are explained by specific causes of weaning difficulties, including positive fluid balance and cardiogenic pulmonary edema and respiratory muscle dysfunction. The third group (prolonged weaning) represents patients entering the general definition of ventilator-dependent patients, who are usually tracheotomies and may benefit from specialized weaning centers.

It is still common belief that the separation from the ventilator should always be gradual, which is not justified [10]. In addition, the assessment of the degree of the reversal of the causative process is often based on subjective grounds, with a frequent natural tendency for clinicians to keep their patients on the 'safe' side, i.e., considering them as not being ready for separation. A landmark trial describing a systematic approach aimed at finding simple criteria for performing a SBT has shown interest in significantly shortening the duration of invasive ventilation and reducing associated complications. In this study, Ely et al. (1996) have shown the interest of a systematic approach to "patient weanability" by the daily search for predefined criteria that, if necessary, led to the completion of a SBT. The criteria were variables that were easily obtained at the patient's bedside and that showed the level of oxygenation, hemodynamic stability, the presence of a cough and the measurement of the respiratory rate/tidal volume ratio. Using this approach significantly reduced the duration of invasive ventilation without any additional complications such as reintubation [11]. More recently, a strategy that paired spontaneous awakening, based on the interruption of sedatives, with SBTs improved extubation rates, reduced ICU length of stay and decreased mortality by 32%. Therefore, bundles have been proposed to facilitate separation from invasive ventilation [12]. It seems therefore important to facilitate the weaning process and to allow extubation as soon as it is possible when the patients are ready. The advantages of early discontinuation must also be balanced against the detrimental consequences of extubation failure. Patients who fail the extubation (extubation failure) and require reintubation have a high mortality [13], which to some degree may be precipitated by the extubation failure itself. Extubation failure defined by the need of reintubation or the development of post extubation acute respiratory failure is associated with prolonged duration of mechanical ventilation and increased mortality. In studies, extubation is considered as being successful if the patient is not re-intubated after 48 hours [22] [23]. Though, some studies extend this period to 72 hours [16]. The timing of reintubation is likely to influence the outcome: delayed reintubation is associated with a higher mortality rate [13]. Therefore, early identification of patients at high risk of extubation failure is of great importance. Accordingly, studies have identified risk factors of extubation failure [17]–[19]. In presence of one of these risk factors (age>65 years, chronic respiratory disease, chronic heart disease), the rate of extubation failure is around 15% and can reach up to 30% [20]. By contrast, patients without risk factors have lower rate of extubation failure, between 5 to 10% [19]. In clinical practice, recognition of clinical worsening after extubation can be delayed because key elements of respiratory monitoring (e.g., tidal volume) are no longer available and respiratory rate alone is not a good indicator of inspiratory effort. Since decades, physicians struggle to develop monitoring tools that could predict the outcome of extubation, such tools would permit to

carefully monitor a patient with high risk of extubation failure and to personalize a preventive strategy with multifaced approach including non-invasive ventilation, high flow oxygen cannula, chest physiotherapy, mobilization and armchair.

### 2.1.3. Extubation Predictors

Extubation failure is provoked by a constellation of causes, some of them potentially linked. When the capacity of the system (respiratory muscle strength, respiratory drive) is reduced, the load (lung disease, fluid overload, cardiac dysfunction, abnormality of the chest wall) is increased, leading to a respiratory load/capacity imbalances. To prevent the risk of extubation failure, clinicians have searched for monitoring tools and clinical indicators that could predict the extubation outcome. A number of indexes such as vital capacity, maximal inspiratory pressure, arterial carbon dioxide tension has been studied as predictors but their predictive performance is poor [21]. In 1991, came a landmark study introducing the Rapid Shallow Breathing Index (RSBI) [22]. This index reflects the efficiency of the respiratory system to cope with unassisted breathing. It is calculated by dividing the tidal volume ($V_T$), by the respiratory rate (RR). A value superior to 105 bpm/min indicates that the patient has a high risk to fails the extubation. This index serves as the gold standard for extubation prediction. Since then, several new metrics have been tested, here is presented a non-exhaustive list:

- **Tidal volume**: the measure of tidal volume has been tested on patients, however, the difference observed between the patient succeeding and failing the extubation is not always significant [23], [24]. However, studies have shown that the coefficient of variation (CV) is more efficient to separate success from extubation failure [25], [26]. The Figure 2 shows the difference in performance between the average Vt and CV Vt;

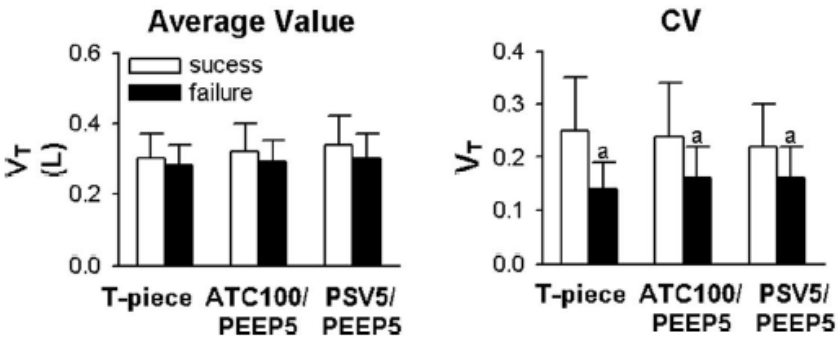


Figure 2 - Comparison of the performance between Average Vt and Coefficient of variation Vt under different ventilator settings [26].

- **cough strength:** for a patient to succeed the extubation, he needs to be able to have a sufficient cough strength to expel mucus from the breathing airway. To that end, test

has been done, such as measuring the peak ventilation flow of a cough, to study the prediction on the extubation depending on the cough strength [27], [28]. Though, from the study Gobert et al, the difference in cough peak flow (CPF) does not seem to yield significant difference between the two populations [29].

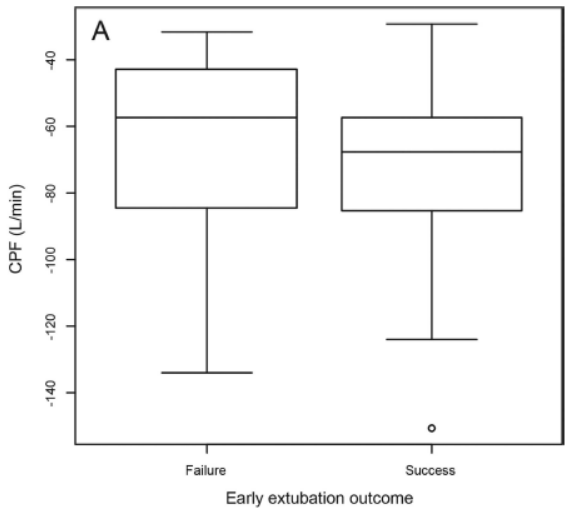


Figure 3 - Cough peak flow results for the failure and success extubation groups [27].

- **blood analysis:** signs of alveolar hypoventilation (hypercapnia notably) can be observed in the blood gas analysis. A lot of markers have then been tested [23], [24], [30], but there are no consensus on which is the best one yet. The ratio between PaO2 (oxygen partial pressure) and the FiO2 (fraction inspired in oxygen) seems to be providing interesting prediction capability depending on the study though;

- **ultrasound:** diaphragm ultrasound has been increasingly used in the ICU and several studies have been done to investigate its use and relevance to predict weaning outcome. When the diaphragm contracts, it shortens, thickens, displaces, and stiffens which provide interesting estimates of diaphragm function, a key determinant of weaning the outcome. For instance, a thickening fraction of the diaphragm below than 29% (cf. Figure 4) has been associated with the diaphragm dysfunction in mechanically ventilated patients [31] and cut-off values for predicting successful weaning (success of the spontaneous breathing trial and/or extubation success) range between 25 and 35% [32]. Another helpful indication for ultrasound is to image the lungs to evaluate the loss of lung aeration during spontaneous breathing (cf. Figure 5). Excessive lung aeration loss at the time of the spontaneous breathing trial is associated with weaning failure. There are many causes for lung aeration losses, including alterations in lung compliance/resistance, ventilation/perfusion mismatch, atelectasis and pulmonary edema. Respiratory muscle weakness has also been suggested to play a role in lung

aeration loss during weaning, but has never been investigated. Lung ultrasound is an interesting tool in this context since it can show the occurrence of B lines that are mainly generated by weaning-induced pulmonary edema [33], [34].

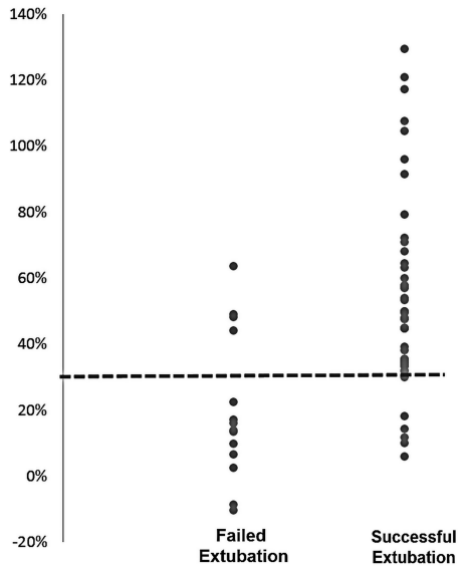


Figure 4 – Percentage change of the diaphragm thickness between end-expiration and end-inspiration [35]

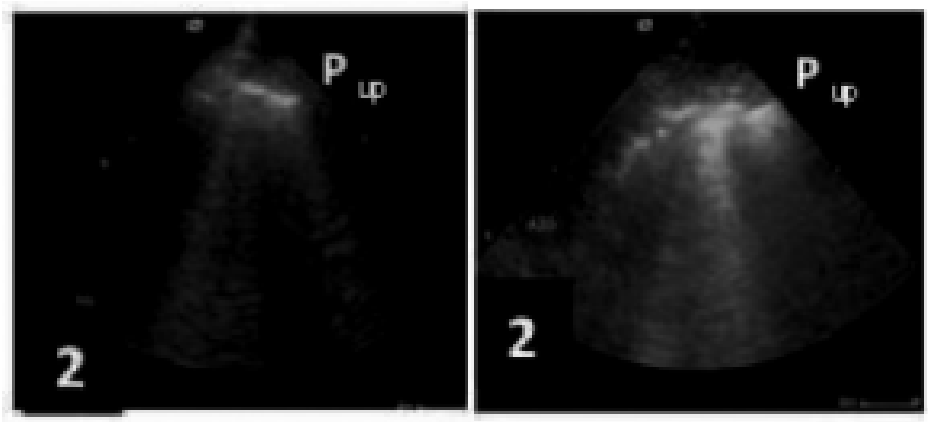


Figure 5 – Ultrasound of the left lung on the upper posterior thoracic region. The image on the left shows a good aeration, while it is degraded in the right image [36].

Some of the prediction results for the extubation are displayed in the Table 1.

Table 1 – Protocol setup and prediction results on the extubation using different methods.

| **Authors** | **Measurement period** | **Measures** | **Prediction method** | **Nb of patients** Success-Failure | **Inclusion parameters** | **Sensitivity** | **Specificity** |
|---|---|---|---|---|---|---|---|
| Segal et al. [37] | During the SBT | Measures from the ventilator | Rapid Shallow Breathing Index | 100 86% - 14% | - Patient's age > 18 - MV duration > 48h | 0.89 | 0.89 |
| Perren et al. [38] | Prior to the SBT | Asking patient's opinion | Patient were asked if he felt confident in his extubation outcome | 211 78% - 22% | - Patient's age > 18 - MV duration > 10h | 0.72 | 0.33 |
| | After the SBT | | | | | 0.86 | 0.64 |
| Saugel et al. [24] | Prior to the extubation | Blood analysis | Serum anion gap | 61 89% - 11% | - Patient's age > 18 | 0.7 | 0.86 |
| | Prior to the extubation | | Corrected serum anion gap | | | 0.88 | 0.71 |
| DiNino et al. [35] | During the SBT | Lung ultrasound | Diaphragm thickness | 63 78% - 22% | - Patient's age > 18 | 0.88 | 0.71 |
| Gobert et al. [27] | After the SBT | Measures from the ventilator | Cough peak flow | 92 76% - 24% | - Patient's age > 18 - MV duration > 24h | 0.7 | 0.64 |
| | After the SBT | | Tidal volume | | | 0.64 | 0.86 |
| Llamas-Álvarez et al. [16] | Cohort of study with different methods | Lung ultrasound | Diaphragmatic excursion | 1071 Depends on the study | Depends on the study | 0.75 | 0.75 |
| Hsieh et al. [39] | During the SBT | Review medical file | Neural network | 3602 95% - 5% | - Patient's age > 18 - MV duration > 24h | 0.82 | / |

More complex prediction models using multiple variables have been investigated. Baptistella et al. [40] tested 28 variables such as blood gas (arterial pH, hematocrit, arterial oxygenation tension, …), hospital meta data (days of MV, days of sedation, fluid balance, …) or ventilator data (RSBI, tidal volume, lung compliance,…). Features were then chosen for the extubation's prediction model based on their area under the curve (AUC) scores. For computing their development cohort, they recruited consecutive patients with the following criteria: older than 18 years old and mechanical ventilation longer than 24 hours. Eventually, 112 patients were enrolled, from them 8% had extubation failure. After reviewing the prediction performance of all 28 individual features, 7 Features were kept for the prediction model. Those features were the RSBI, the lung compliance, the duration of MV, the maximal inspiratory pressure, Glasgow coma scale, hematocrit and serum creatinine. Those features were then combined into one metric called ExPreS (Extubation Prediction Score). From the ExpreS score, three groups of extubation outcome probability were determined (low, intermediate and high of success). The

result for each group is detailed in Table 2. After creating this model, the authors tested their model in a validation cohort in which 232 patients were recruited. They were able to reduce the original failures rate of 8.2% from the original database to 2.4% in the validation cohort.

Table 2 - Extubation results from the extubation prediction score ExPreS [40].

| | Sensitivity | Specificity | Success rate (%) | Success probability |
|---|---|---|---|---|
| ExPreS ≤ 44 points | 0.96 | 0.33 | 57.10 | Low |
| ExPreS from 45 to 58 points | 0.950 - 0.667 | 0.333 - 0.772 | 88.30 | Intermediate |
| ExPreS ≥ 59 points | 0.89 | 0.75 | 98.70 | High |

An artificial network has also been tested on ICU data by Hsieh et al [39]. They retrieved medical data from patients going through the ICU between 2009 and 2011 in a Taiwanese hospital. All patients older than 18 years old that went through at least 24 hours of mechanical ventilation could be included in the study, upon the patient's consent. Their database is composed of 3602 patients with 161 patients receiving Non-Invasive Ventilation (NIV) after the extubation. They monitored the patients for 72h after the extubation, after which the patient was considered an extubation success, even if they received NIV. In total, 37 features were included in the prediction model. These 37 features gathered blood gas analysis, meta data such as the patient's age or past medical history or ventilator data such as maximum inspiratory pressure, tidal volume or respiratory rate. For the prediction model, they used a Multilayer Perceptron neural network [41] (cf. Figure 6), which is a common neural network algorithm. They used 3 layers, the first one includes of one input layer with 37 dimensions composed of the 37 medical features, the hidden one has 19 dimensions and the output has 2 (success or failure). This model provides good prediction results with a F1 score of 0.867, a precision of 0.939 and a recall (also called sensitivity) of 0.822.

Using a single parameter or a bunch of variables, studies have tried to develop the best predictor of extubation failure. However, it must be underlined that that the predictive performance of any predictor highly depends on the pretest probability of the occurrence of the event [42]. In addition, the timing of the test is crucial. The predictive performance is completely different if the test is performed before the SBT or after extubation.

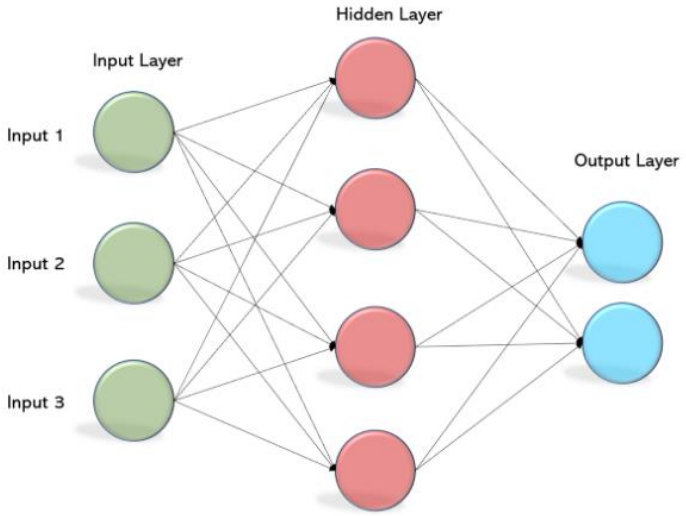


Figure 6 - Architecture example of a Multilayer Perceptron neural network.

On the one hand, the SBT actually challenges the cardiac function and the muscular pump to mimic the breathing of a patient without the ventilator. On the other hand, the extubation means to breathe without the endotracheal tube. Therefore, it is important to ask to a predictor only what it can actually do. In other words, a predictor related to load/capacity respiratory balance could perfectly predict a successful SBT but poorly extubation if extubation failure is related to aspiration and weak control of the upper airways. As a consequence, rather trying to predict the extubation outcome, a pragmatic approach would be to carefully monitor the patient's status in order to detect as early as possible a sub-clinical deterioration. This approach requires continuous and reliable monitoring tools.

## 2.2. Electrical Impedance Tomography

### 2.2.1. Introduction to the EIT

The first use of EIT on humans was published in 1984, where it was shown that it was possible to observe the distribution of conductivity variation in vivo [4]. Since then researchers have tried this imaging method for different applications in the medical field, such as monitoring the brain activity [43], breast cancer detection [44] or even estimating the urinary volume [45]. Though it seems that application to monitor the ventilation through the EIT are the most commonly find in research.

The EIT apply on the thorax offers image with great contrast of the ventilation in the image (cf. Figure 7). Indeed, the EIT apply to the thorax, measure the movement of air, which is not conductive, in a tissues environment which is conductive to some degree. Thus, during an inspiration, we observe a decrease of conductivity induce by the air filling the lung.

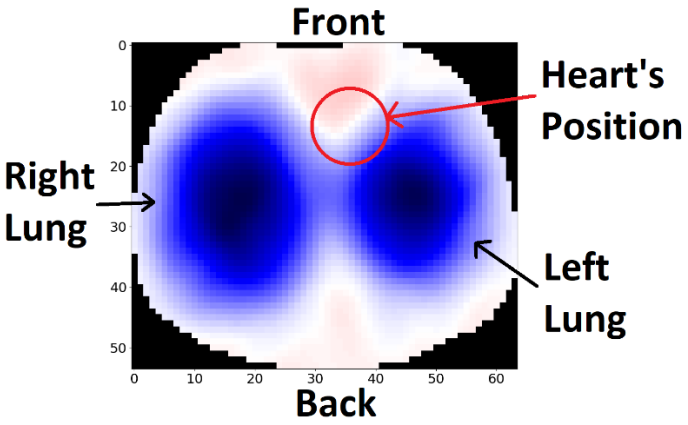


Figure 7 - EIT images from a healthy subject. The blue colors represent negative conductivity variation induce by the ventilation.

EIT measurements work by placing electrodes around the subjects' thorax (using an electrode belt or individual electrodes depending on the system). Then, a current is injected between a pair of electrodes while measuring the voltage on the following electrodes (see Figure 8). These measurements are repeated until current has been injected into all pairs of electrodes. The voltages are then processed by a reconstruction algorithm which makes it possible to map the conductivity variations. There are so-called "absolute EIT" algorithms which allow the conductivity values to be displayed directly. However, this technique is not very successful, because it is highly sensitive to errors due to the movement of the electrodes or variation in contact impedance [46]. Therefore, in a large majority, the 'time-difference EIT' is used. This method displays the variations of conductivity by realizing the difference of the measurements between two instants.

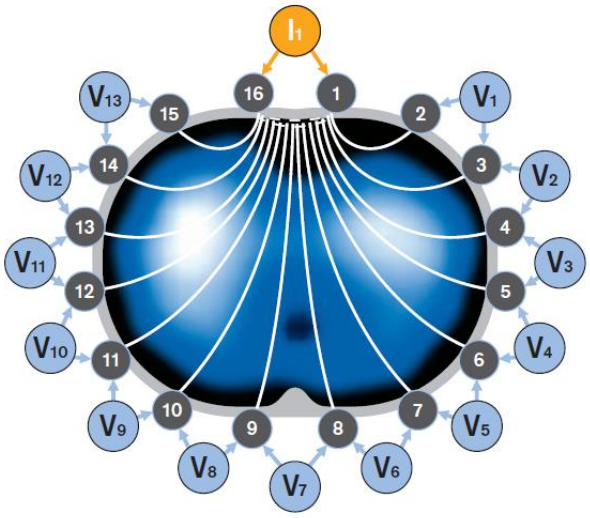


Figure 8 - Current injection and voltage measurement pattern [47]

Several EIT devices are commercially available (cf. Table 3). There are two types of devices, depending on whether they have the CE mark. They are either intended in a hospital

environment for monitoring and test on patients; or for research purpose like the Swisstom Pioneer Set where you can adjust the current injection parameter.

Table 3 – List of available EIT devices marked CE on the market [48]

| Manufacturer | EIT System | Electrodes | | Image Reconstruction algorithm | Measurement and Data Acquisition |
|---|---|---|---|---|---|
| | | Number | Configuration | | |
| Swisstom AG | BB² | 32 | Electrode belt | GREIT | Adjustable skip |
| Timpel SA | Enlight | 32 | Electrode stripes | Newton-Raphson | 3-electrode skip |
| CareFusion | Goe-MF II | 16 | Individual electrodes | Sheffield back-projection | Adjacent |
| Dräger Medical | PulmoVista 500 | 16 | Electrode belt | Newton-Raphson | Adjacent |
| Maltron Inc | Mark 1 | 16 | Individual electrodes | Sheffield back-projection | Adjacent |

The EIT also found other applications to it than the medical field. A variant to this method, ERT (Electrical Resistivity Tomography) is used in geology [49], [50]. Electrodes are placed on the surface of the ground. The measurement allows to map the underground by displaying the different components in the ground that have different resistivity.

The EIT also found an application in the robotic field [51], [52]. Researchers created artificial skin from conductive materials. They then apply EIT measurement to it to detect the position where the skin is touched.

### 2.2.2. EIT reconstruction process

Finding the different conductivity inside a system from voltage measurements realized on the boundary is an ill-posed problem. Meaning that for a set of voltage measurements, they are not a unique solution to the problem. Thus, different reconstruction algorithm could arrive to different solutions and recreate different EIT images showing different conductivity variations.

Historically, the first reconstruction algorithm was the ‘Sheffield backprojection’ algorithm [4]. This method consists of adding the equipotential of each current injection couple, which makes it possible to highlight the zones which have had the most variation in conductivity (see Figure 9).

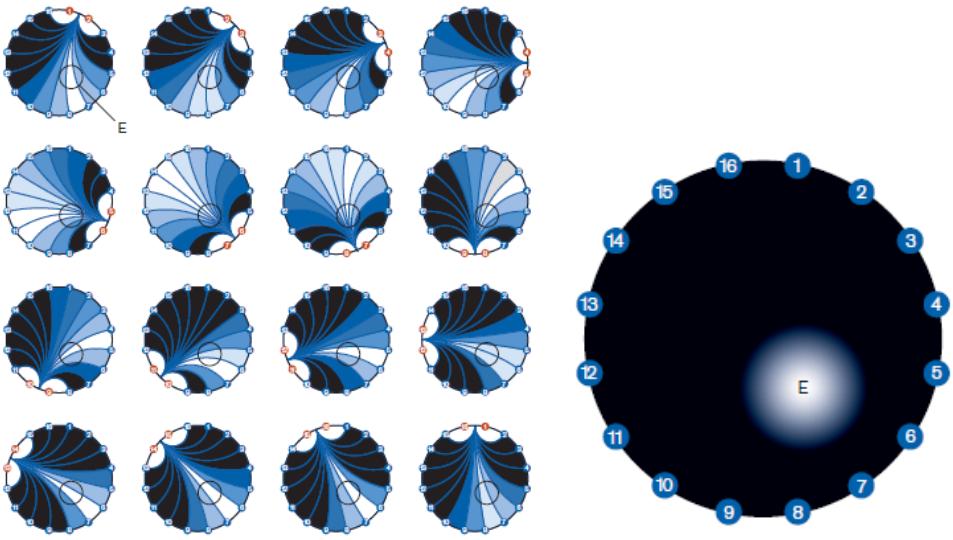


Figure 9 - The superposition of each of the equipotential on the image on the left makes it possible to obtain the result which is observable on the image on the right [47].

Yet, most of the reconstruction algorithms are based on sensitivity models (see Figure 10). This method consists of comparing the real boundary voltage measure, from estimation of boundary voltages calculated thanks to the forward solution. In the case of the EIT, the forward solution calculated the boundary voltage from the system conductivity. So, algorithms iterate over themselves to minimize the difference between the real boundary voltages and the estimates ones, in order to find the right conductivity. From this original method, many algorithms emerged throughout the years, with different regularization methods or other improvements devoted to the image quality [53]–[55].

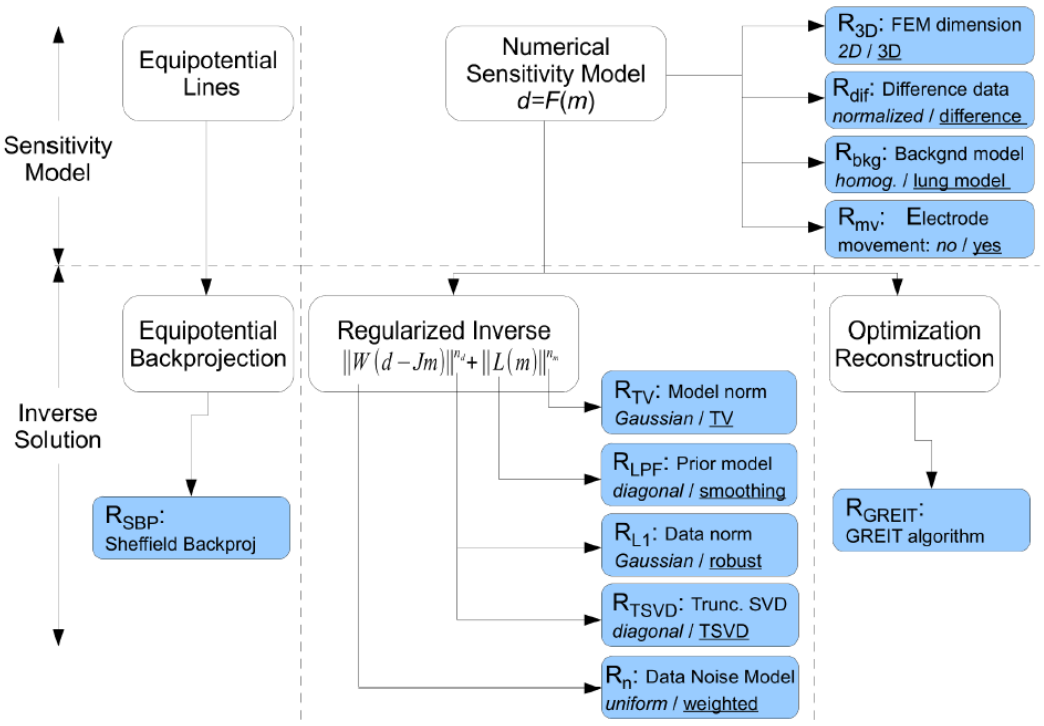


Figure 10 - Description of the different sorts of EIT reconstruction algorithms [56]

In 2009, a consensus of experts chose to opt for a new approach and created the GREIT algorithm which improves the image by optimizing the reconstruction matrix R [57].

To configure R, GREIT creates a simulation set with different conductivity variations (later called target). From this, two EIT images are created: the desired EIT image and the real image (cf. Figure 11). It is possible to see that the desired image results in a larger circle, because GREIT considers the blurring effect inherent in EIT. Indeed, it was found with other reconstruction algorithms that the conductivity variation in the center could be blurred or not detect at all. This problem is due to the ill-conditions of the EIT reconstruction [58]. It can be solved by increasing the regularization parameters during the reconstruction.

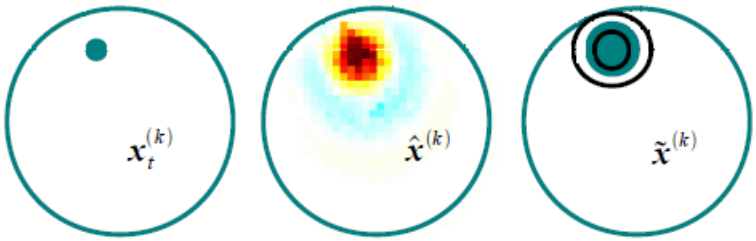


Figure 11 - On the left, the simulation of a variation in conductivity; in the middle, the real image; on the right, the desired image [57].

From the real image and the desired one, GREIT adjust certain pixels weight in the image, to have an image reconstruction closer to the desired image. The creation of the reconstruction matrix is done by following the objectives of six performance parameters which the expert consensus has deemed to be the most important. Its parameters are as follows (ordered from most important to least important):

- **amplitude response:** knowing that the amplitude of the targets is the same throughout the test set, the amplitude on the EIT image must be the same between each test;
- **position error:** the center of gravity of the object in the real image must be at the same position as the target;
- **resolution:** the size of the actual image must be equal to the size of the desired image;
- **shape deformation**: the reconstructed image must be a circular shape just like the target;
- **circle effect:** the W matrix tries to limit the circle effect that can be created around objects reconstructed in EIT;
- **noise amplification**: It must be ensured that the noise on the signals is not amplified.

Once the training is complete, the reconstruction matrix R can be used to reconstruct the EIT images $\hat{x}$ with a simple multiplication with the voltage vector y (Cf. Equation 1).

$$\hat{x} = Ry$$ Equation 1

The whole process is described in the Figure 12. With an EIT system using 16 electrodes, this results in voltages vector of 208 measures. This number comes from the fact that, for a certain electrode pair injection, 13 measures are realized (cf. Figure 8). The 13 measures are then repeated for 16 different electrode pair injection. Once the 208 measures are done, we have all the measures corresponding to an EIT frame. For the time-difference EIT, a reference has to be subtracted. The reference chosen depends on the EIT application. For pulmonary monitoring, the reference chosen is usually a frame corresponding to an end-expiration. Now that the voltage dataset was subtracted to the reference, the image can be reconstructed. For each frame, the voltage vector is multiplied by the reconstruction matrix. This results in an image represented as a vector, which then have to be rearranged to be displayed as an image.

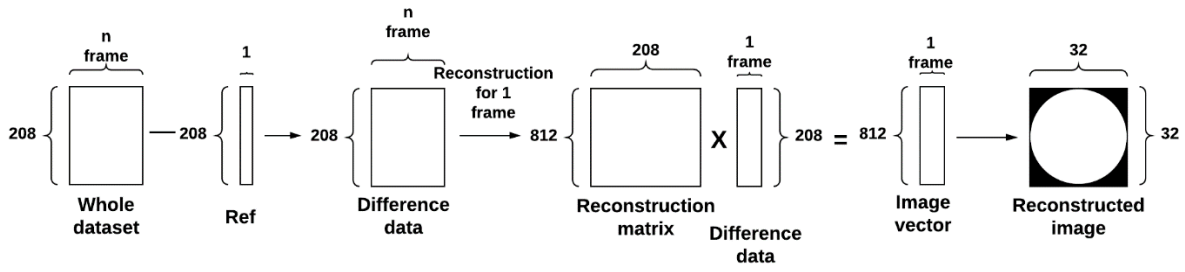


Figure 12 - Whole EIT reconstruction process.

Among the linear-one step algorithms, it has been shown that GREIT is the best one [56]. This sort of algorithm has the advantage to simplify the EIT image computation by the multiplication of the reconstruction matrix with the voltage vector.

## 2.3. EIT applications in the ICU

Although the use of EIT is not used in routine in the ICU, it represents a promising approach to continuously monitor the distribution of the ventilation inside the lungs. It is non-invasive, it provides continuous visualization of the ventilation and it is not ionized. It may be useful to detect heterogenous ventilation which is harmful or to detect the progressive loss of lung aeration during a spontaneous breathing trial or later after extubation. EIT could also be useful to personalize the ventilator settings.

It was found that through the EIT measurement, it is possible to estimate the variation of the ventilation volume [59], [60]. Indeed, this estimation is rendered possible through the calibration with the real ventilation volume measured by the ventilator or another device like a spirometer. This allows to translate EIT metrics into ones uses in the ICU so that the doctors

can better grasp their meaning. The estimation of the volume variation through the EIT is calculated thanks to the calibration coefficient a (see Equation 2). The calibration coefficient corresponds to the slope coefficient of the line that best fit the linear relationship between the ventilation volume and the conductivity variation ΔZ (cf Figure 13), which is calculated by summing every pixel in the EIT image.

However, it is important to note that the EIT cannot measure the actual ventilation volume but only the variation. This limitation is due to that fact that we realized time-different EIT, and thus always compare the measure to a reference from which we calculate the variation.

$$V_{EIT}(t) = a \times \Delta Z(t) \qquad \text{Equation 2}$$

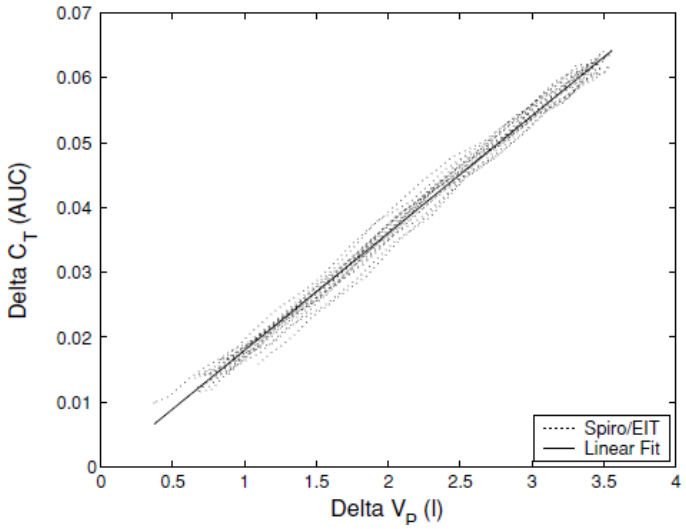


Figure 13 - Conductivity variation $\Delta Z$ (called here Delta C) depending on the variation of volume [61].

Study have been realized to check the validity of the EIT against the CT-scan, which has the highest resolution compared to other monitoring devices [62]. Tests were realized on porcine mechanically ventilated models. The authors of the study compared the estimation of volume through the two different methods, by calculating the tidal volume and the variation End-Expiratory Lung Volume, EELV. EELV is the volume present in the lungs at the end of expiration and is also called functional residual capacity in spontaneously breathing subjects. The EIT counterpart is the variation End-Expiratory Lung Impedance, ΔEELZ, the calculation process is the same except that we used ΔZ instead of the ventilation volume. Using the Pearson correlation score, they found a good correlation between the different metrics, 0.84 for the tidal volume against ΔZ and 0.94 for ΔEELV against ΔEELZ. Thus, proving the reliability of this method against a gold standard.

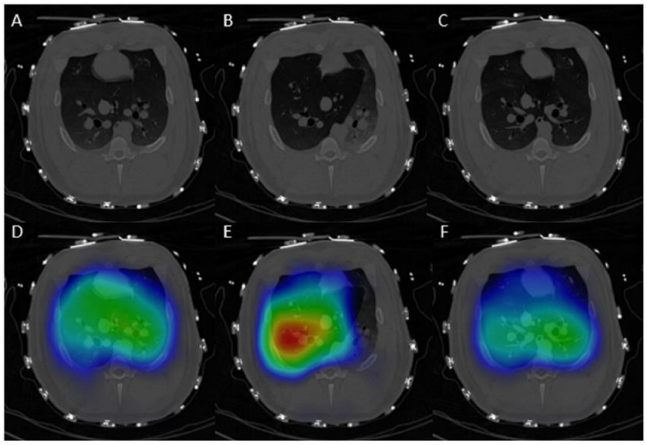


Figure 14 - Measurement of CT-scan and EIT realize on a porcine [62]. On top, the CT scan images. At the bottom, the EIT images are added.

Another application of EIT is for setting the right level of Positive End-Expiratory Pressure (PEEP). The PEEP allows to keep the lungs opened at the end of expiration. The optimal level of PEEP is a critical setting when delivering ventilation because there is a tradeoff between higher PEEP (and increased EELV but a risk of hyperinflation) and lower PEEP (and decreased EELV but a risk of atelectasis). Before EIT, doctors relied on different baseline rules but had to rely on their experience to find the right settings for an individual patient [63]. With the EIT, it is now possible observe the area that is over distending or collapsing depending on the level of pressure [64], [65]. Thus, allowing to find the optimal ratio between collapse and over distension.

The EIT also have different metrics to quantify the distribution of the ventilation in the lungs. Such as the center of ventilation (CoV) [66], [67] which indicates the distribution of the tidal volume between the anterior-posterior parts. This is realized by calculating the sum of pixels of each line $\Delta Z_{line}$ and realizing a weighted average of each of them (cf. Equation 3).

$$CoV = \frac{number\ of\ line \times \Delta Z_{line}}{count\ of\ all\ line} \qquad \text{Equation 3}$$

The calculation process can be simplified by looking at the Figure 15.

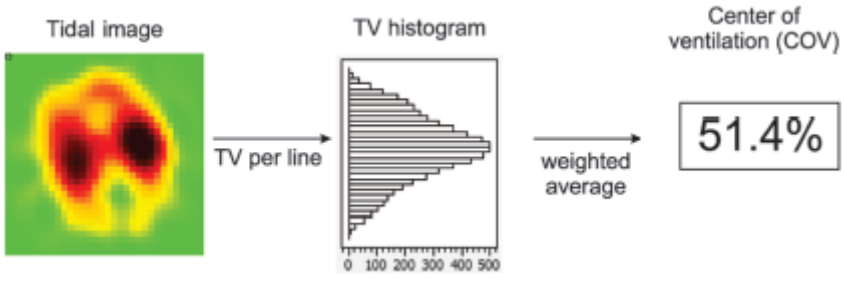


Figure 15 - Process to calculate the EIT features CoV.

One other application of EIT has been tested by Longhini and his team which is especially interesting for our study. They tested the capability of the EIT to predict extubation failure [68]. Measurements were realized before and after the 30 minutes SBT, as well as right after the extubation and 30 minutes later. Three EIT metrics were tested: estimation of the tidal volume by EIT measurement $\Delta Vt_{EIT}$ (after calibration with the ventilator volume), ΔEELZ and the global inhomogeneity index (GI), which is a metric that reflect the homogeneity of the volume distribution in the lungs [69], [70] and is calculated as follows:

$$GI = \frac{\sum_{x,y \in lung} |\Delta Z_{xy} - Median(\Delta Z_{lung})|}{\sum_{x,y \in lung \Delta} Z_{xy}} \quad \text{Equation 4}$$

With $\Delta Z_{x,y}$ being the value of the pixel located in the position x, y; and $\Delta Z_{lung}$ being the sum of all pixels belonging to the lung in the EIT image.

To test the extubation capability of those EIT metrics, a threshold was used to separate both class (success and failure). The Youden index method was used to calculate the threshold [71]. They included 61 patients whom 36.1% of them had extubation failure.

With the success defined as the negative class and the extubation failure as the positive class, they found the results displayed in the Table 4. The best predictors come from ΔEELZ after the SBT and the GI, 30 mins after extubation.

Table 4 - Results on the prediction of the extubation outcome using EIT metrics [68]

| | ΔVtEIT | | ΔEELZ | | GI | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| Before SBT | 0.68 | 0.66 | 1.00 | 0.17 | 0.68 | 0.68 |
| After SBT | 0.77 | 0.63 | 0.82 | 0.49 | 1.00 | 0.32 |
| Right after extubation | 0.55 | 0.78 | 0.46 | 0.66 | 0.64 | 0.73 |
| 30mins after extubation | 0.73 | 0.59 | 0.64 | 0.63 | 0.86 | 0.51 |

Longhini and his team demonstrated the ability of the EIT to predict extubation failure. Though, their study is limited to the 3 EIT metrics they tested. In addition to that, there weaning observation was limited to 30 mins.

## 2.4. Conclusion

Over this overview of the existing literature on weaning and prediction of extubation failure, we introduced the importance to rightfully predict the prediction outcome. Prediction models

have been developed and tested by researchers. Some provide good results. Like the models from Baptistella et al [40] where they have been able to show the model capability by reducing the extubation failure rate on a validation set. Though, most of the proposed models focus the observation during the spontaneous breathing trial or before the extubation. This limitation pushed us to develop our clinical study. Our hypothesis is that continuous monitoring of regional volume distribution by EIT after extubation may detect patients at risk of subsequent extubation failure. The aim is to give a more extending understanding of the capability of the EIT to predict extubation outcome. To do this, we included a large number of EIT features and we observe the patients for 48 hours in order to describe how the ventilation patterns change after extubation, as well as developing more complex prediction models to predict the extubation failure.

# 3. Clinical study EXIT

*This chapter describes the clinical study that was put in place to study the EIT performance to monitor patients after extubation (NCT04180410). The clinical study was designed at the beginning of the thesis by my thesis supervisors and myself. This chapter describes the clinical study that was conducted to ascertain the EIT performance to monitor patients after extubation. Dr Martin Dres, in collaboration with Dr Vincent Bonny and Dr Vincent Joussellin, were responsible of the inclusion and setup. M. Andrea Pinna and Vincent Janiak oversaw the testing and analyzing the prediction model, in collaboration with the Pr Christophe Marsala. The inclusion of the patients occurred at the intensive care unit of the hospital La Pité Salpêtrière in Paris and lasted between the February 2020 to March 2022.*

## 3.1. EXIT scope

For this study, the main goal is to monitor patients during the weaning phase with EIT. Such study has never been realized before, so we do not know how the ventilation will change according to the EIT monitoring.

From those EIT data, we want to propose a prediction model that could predict the extubation outcome. We want to include as many as possible EIT features to have a whole view of the patient's ventilation from the EIT perspective.

## 3.2. Study population

We included ICU patients older than 18 years, who were mechanically ventilated for at least 48 hours through an orotracheal tube, with at least one risk factor of extubation failure (age > 65 years old, chronic heart or pulmonary disease). They had to succeed to a spontaneous breathing trial. Patients undergoing extubation during weekends were not considered for inclusion. Non-inclusion criteria were the following: pregnant women, patients under extra-corporal assistance (ECMO), patients without social insurance.

## 3.3. Weaning observation

Patients were included in the study once the physician prescribed the extubation, and if the patient met the inclusion criteria. After enrollment, a silicon 16-electrode EIT belt of proper size was placed around the patient's chest between the 4th and 6th intercostal spaces and connected to the EIT device (PulmoVista 500; Draeger Medical GmbH, Lübeck, Germany). EIT was connected to a ventilator (Infinity V500; Drager Medical GmbH, Lübeck, Germany)

through a RS232 interface to calibrate impedance variation with volume while patients were still intubated.

After 48 hours, if the patient was not re-intubated, she/he was considered as an extubation success. Post-extubation acute respiratory failure was predefined by the presence of one or more of the following criteria (if persistent over 5 minutes): SpO2 < 90% with an oxygen support ≥ 5 L/min, a respiratory rate ≥ 35/min, a pH < 7,35 with a $pCO_2$ > 45 mmHg. Patients were allowed to receive preventive NIV.

### 3.4. Data acquisition

For this clinical study, we decided to realize 7 visits with different measures that are realized. The first one occurred right before the extubation (H0), then 2 hours after the extubation (H2), then 6 hours (H6), 12 hours (H12), 24 hours (H24), 36 hours (H36) and at last 48 hours (H48). Between those visits, no measures were realized on patients.

At each visit, EIT measurement was realized by the Drager Pulmovista 500 [1] set to 30 frames/s. Two EIT recording of 5 minutes each was realized at each visit. The 16 electrodes belt was positioned between the 4th and 6th intercostal space. A temporary mark is placed on the thorax to make sure the belt does not move significantly between takes. Each EIT measures last for 5 minutes which produced 9000 frames.

In addition to the EIT measurement, several other variables were collected (see Table 5). The other measures realized at each visit are intended for other goals than the one defined for this thesis. The results from those are therefore not discussed in this report.

Table 5 - Data available for the first 5 visits.

| H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|
| Meta data | Usual ICU data | Usual ICU data | Usual ICU data | Usual ICU data |
| Arterial blood gas | Heart ultrasound | Dyspnea sign | EIT | Blood analysis |
| Usual ICU data | Lung ultrasound | EIT | | Usual ICU data |
| Heart ultrasound | Dyspnea sign | | | Heart ultrasound |
| Lung ultrasound | EIT | | | Lung ultrasound |
| EIT | | | | EIT |

### 3.1. Weaning outcome

During the 2 years inclusion periods, 37 patients were included. The first one occurred on February 28th in 2020, the last one was included on Marsh 17th in 2022 (cf. all inclusions date

in Table 30 in the Annex). From the 37 patients, 2 could not be analyzed, which left us with 35 patients in our database, with 26 success and 9 failures, thus a failure rate of 26%. Table 6 describes the anthropometric measures, as well as the medical condition of patients from both classes.

Table 6 - EXIT meta data. The first 7 lines display the following results: median [1st quartile; 3rd quartile].

| | **Success** | **Failure** |
|---|---|---|
| Age [years] | 61 [49; 66] | 59 [40; 69] |
| Weight [kg] | 72.5 [70; 94.9] | 75 \|69.3; 78] |
| Height [m] | 1.73 [1.68; 1.78] | 1.7 [1.68; 1.76] |
| Body Mass Index (BMI) [kg/m²] | 27 [23.8: 29.8] | 24.9 22.3; 28.3] |
| Mechanical ventilation duration [days] | 8 [5; 13] | 8 [7; 12] |
| Delay between extubation and failure [hours] | / | 8.3 [1.1; 14] |
| Sequential Organ Failure Assessment score (SOFA) | 8 [4; 11.8] | 6 [3.3; 8] |
| Patients with chronic respiratory disease [%] | 36 | 33 |
| Patients with chronic heart disease [%] | 36 | 11 |
| Patients with diabetes [%] | 22 | 22 |
| Patients with neuromuscular disease [%] | 5 | 11 |
| Number of males | 17 | 7 |
| Number of females | 5 | 3 |

After extubation, prophylactic NIV was used in 12 patients (34%) and high flow nasal cannula in 8 patients (23%). Extubation failure occurred in 9 (26%) patients within the 48 hours after extubation. The main cause of extubation failure was the presence of ineffective cough (n=6, 14%). Weaning-induced pulmonary edema was documented in one patient. One patient presented a hemorrhagic shock and was re-intubated to perform upper gastrointestinal endoscopy, while another patient was re-intubated immediately after extubation because of a laryngeal edema. Two patients were not re-intubated: the first had a sudden hypoxic cardiac arrest that occurred 74 hours after extubation, while the other developed an acute respiratory failure for which high-flow nasal oxygen therapy was intensified.

The Table 7 shows all EIT measures realized on the patients. Unfortunately, not all measures were realized or can be analyzable. Indeed, 3 success patients (P01, P15 and P22) exited the ICU before the 48h observation, as they were deemed treated. As due for the measure which cannot be analyzed, an example can be seen below in the Figure 16. In such cases, we are not sure what happens during the recording. The assumption is that the patient moved too much and displaced the electrode belt. One last case that need to be addressed is the patient P28. This patient is considered an extubation success even though he was re-intubated, as his failure was not due to ventilation reasons.

Table 7 - Each EIT measurements realized on success's patient (on the left) and on failure's patient (on the right).

| Patient | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
|---|---|---|---|---|---|---|---|
| P01 | | | | | | | |
| P04 | | | | | | | |
| P07 | | | | | | | |
| P08 | | | | | | | |
| P10 | | | | | | | |
| P11 | | | | | | | |
| P12 | | | | | | | |
| P15 | | | | | | | |
| P16 | | | | | | | |
| P17 | | | | | | | |
| P18 | | | | | | | |
| P19 | | | | | | | |
| P20 | | | | | | | |
| P21 | | | | | | | |
| P22 | | | | | | | |
| P23 | | | | | | | |
| P24 | | | | | | | |
| P25 | | | | | | | |
| P26 | | | | | | | |
| P28 | | | | | | | |
| P30 | | | | | | | |
| P31 | | | | | | | |
| P32 | | | | | | | |
| P33 | | | | | | | |
| P34 | | | | | | | |
| P35 | | | | | | | |
| P37 | | | | | | | |
| **Total** | **26** | **25** | **24** | **14** | **18** | **6** | **13** |

| Patient | H0 | H2 | H6 | H12 | H24 | H36 | H48 | Re-intubation delay [hours] |
|---|---|---|---|---|---|---|---|---|
| P02 | | | | | | | | 15 |
| P03 | | | | | | | | 35 |
| P05 | | | | | | | | 45 |
| P06 | | | | | | | | 24 |
| P09 | | | | | | | | 0.5 |
| P13 | | | | | | | | 0 |
| P14 | | | | | | | | 8 |
| P27 | | | | | | | | 1 |
| P29 | | | | | | | | 1.5 |
| P36 | | | | | | | | 24 |
| **Total** | **9** | **6** | **5** | **2** | **3** | **0** | **0** | |

| | |
|---|---|
| | Good measure |
| | ICU discharge |
| | Measure not analyzable |
| | Measure not realize |
| | Re-intubated patient |
| | Ventilation distress |
| | Study exit |

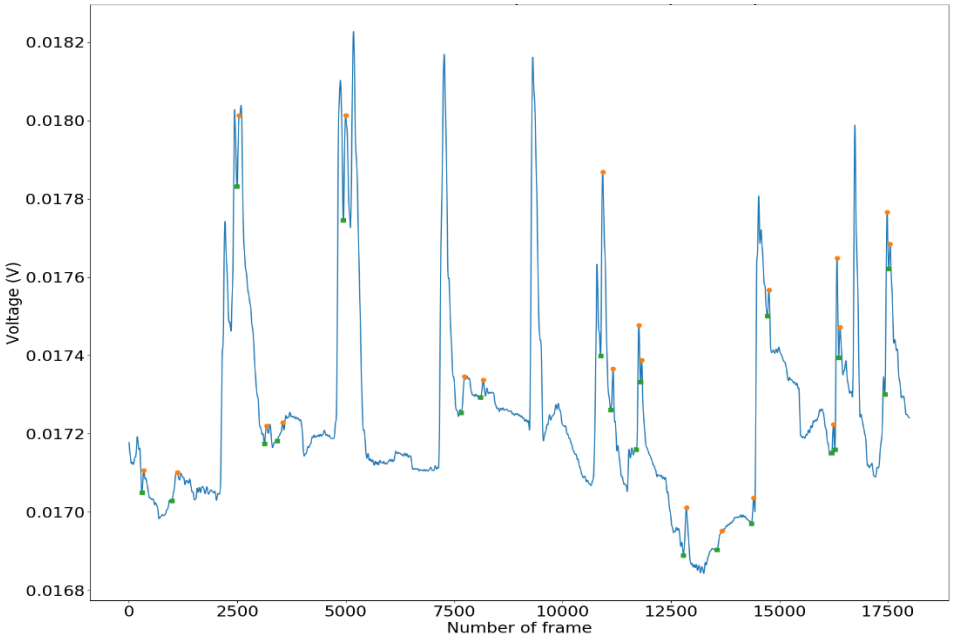


Figure 16 - Example of a take that cannot be analyzed (P35-H12).

## 3.2. Conclusion

This chapter presented the clinical trial. From the measurement protocol, the inclusion parameter and the weaning observation process. The 35 included patients were presented. Table 8 sums up the study and echoes the Table 1 to compare with the previous studies on the subject. Next chapter will review the whole EIT framework from which the EIT metric are calculated and used for prediction.

Table 8 – Protocol setup for our clinical study EXIT.

| **Measurement set** | **Measures** | **Nb of patients** Success-Failure | **Inclusion parameters** |
|---|---|---|---|
| 7 measures after extubation | Electrical Impedance Tomography | 35<br>75% - 25% | - Patient's age > 65<br>- MV duration > 48h |

# 4. EIT features extracted from the images

*Using the Drager Pulmovista 500, we had access to the EIT reconstruction from their software. In addition to that, their software allows to calculate only a couple of EIT metrics (ΔZ, CoV, GI and lung compliance if ventilator data available). Thus, due to the great limitation that are imposed when using their software, we decided to realize our own EIT framework. This enabled us to better optimize the process to fit our need, including adding more EIT features. The whole framework is summarized by the* Figure 17.

*This chapter describes every EIT features that are calculated which are then used in the prediction model. First, the usual features found in the literature are described. Then new EIT features are introduced. Those were developed based on observation realized on the EIT images from both groups (success vs failure). As well as recreating ICU metrics with the EIT signal. All in all, 151 EIT takes were analyzable from 35 patients. Each takes lasted for 5 minutes which yields 9000 EIT frames with the Drager Pulmovista 500. The size of each file is around 50 Mo, which made for a database of 7.5 Go.*

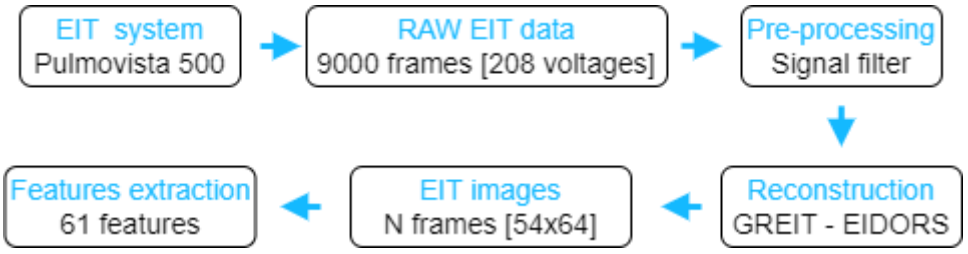


Figure 17 - EIT framework: from the EIT system to the EIT images from 5 minutes takes.

## 4.1. Preprocessing and reconstruction of the EIT data

### 4.1.1. Preprocessing

Patients who were extubated may cough due to accumulation of mucus in the upper airways. Plus, they can be confused due to spending days in artificial coma. Those two factors induce noise in the EIT voltage measurement. As a means to observe if the take is good enough to analyze, we are observing the variation of Vmean. It is the average of the 208 voltages (corresponding to all the voltages of a frame using a 16 electrodes EIT system) and represent the air volume variation in the lungs. An example of the Vmean variation can be seen in Figure 18. This signal is also used to detect the tidal cycles by finding the end-expiration and the end-inspiration events in the signal. These events are located by finding the local maximum and minimum in the Vmean signal.

The tidal cycle detection can be badly impacted by the noise generated by cough or by movements and can generate a wrong detection.

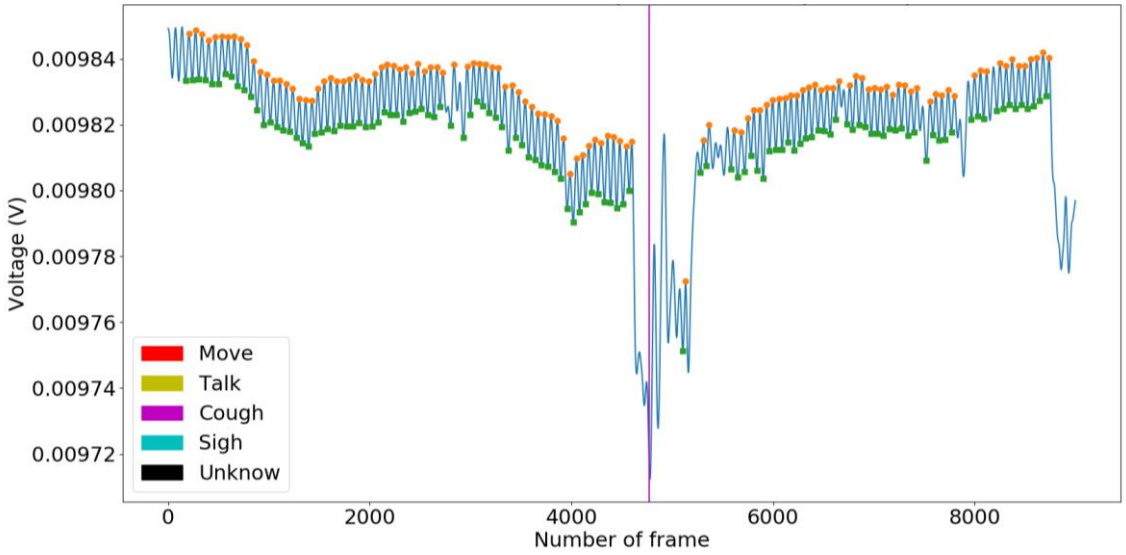


Figure 18 – Unfiltered variation of Vmean during a 5 minutes EIT measures for the patient P06 at H0

In addition to the signal spikes generated by cough or movement, a drift can be seen in most measurements (cf. Figure 19). This does not come from the patient's ventilation. Indeed, the drift is equivalent to 6 times the patient's tidal volume. So, the patient could not have progressively exhaled such volume over the 5 minutes period. Therefore, this drift must come from measurement errors. We are unclear as to why the drift occurs. The drift could come from different sources such as: small movement of the electrode belt, changes in contact impedance between the electrode and the skin (drying of the electrode or sweating).

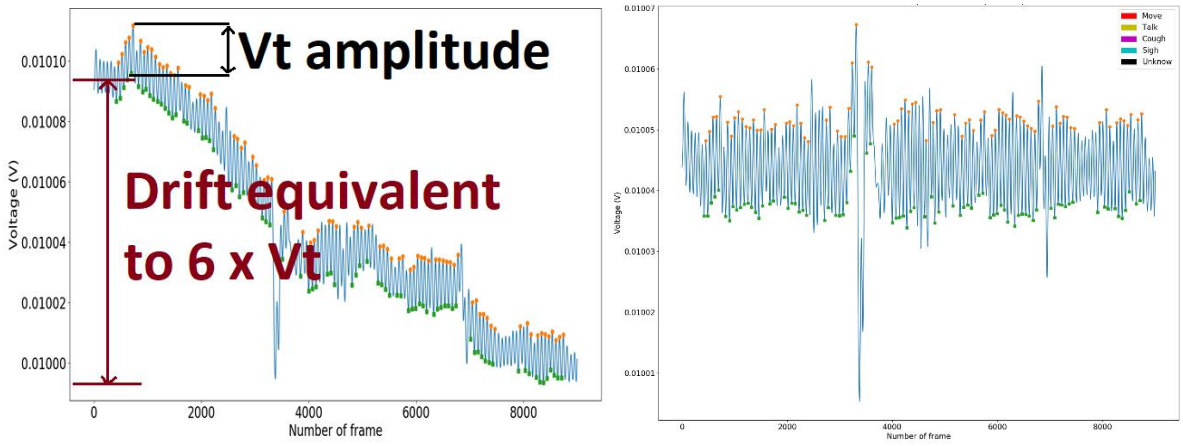


Figure 19 – On the left, unfiltered variation of Vmean during a 5 min EIT measures for the patient P06 at H2. On the right, the same signal but filtered.

To reduce the chance to have a bad take that has too much noise or drift, 2 EIT measures are realized at each measurement set (H0, H2, …). The one with the less disturbance is then chosen. However, the spikes and drift present in certain measures still need to be addressed. To achieve

that, we are applying a band pass filter with cutoff frequencies equal to 0.05Hz and 0.7Hz, that is applied on each of the 206 voltages channel measurements.

As mentioned above, the end-inspiratory and end-expiratory frames are found by searching for the local maximum and minimum in the signal. Though due to the noise, some of the tidal images found are not relevant (cf. Figure 20). To address this issue and only detect real tidal variation, thresholds were put in place to select the tidal image. The tidal amplitudes that were either higher than $Threshold_{high}$ or lower than the $Threshold_{low}$ were not kept (cf. Equations 5 and 6). The thresholds were set as to remove most false tidal detection, even if it meant removing true tidal cycle. From this, we found empirically the following values for the coefficient: $Threshold_{high} = 1.0$; $Threshold_{low} = 0.6$.

$$Threshold_{low} = mean\big(Vmean(tidal_{amp})\big) - 0.6 * std\big(Vmean(tidal_{amp})\big) \quad \text{Equation 5}$$

$$Threshold_{high} = mean\big(Vmean(tidal_{amp})\big) + 1 * std\big(Vmean(tidal_{amp})\big) \quad \text{Equation 6}$$

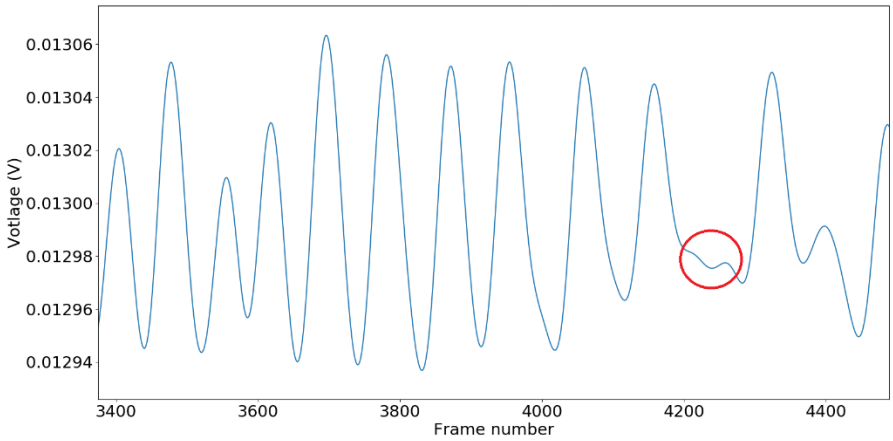


Figure 20 - Example of noise variation that could be mistakenly recognized as a tidal variation

With this process to remove any false tidal detection, we end up having in average of 74 tidal cycles detected between all patients throughout the 48 hours (Cf. Table 9).

Table 9 – Number of tidal cycles kept at each measurement set.

| | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
|---|---|---|---|---|---|---|---|
| Min | 27 | 48 | 39 | 41 | 47 | 60 | 43 |
| Max | 156 | 114 | 133 | 91 | 110 | 135 | 92 |
| Mean | 71 | 75 | 73 | 68 | 71 | 87 | 75 |
| Std | 28 | 19 | 22 | 18 | 17 | 26 | 15 |

This number can vary greatly between patients (Cf. Table 31 in Annex) as the standard deviation results show. But this should not pose any significant problem. Indeed, the patient's ventilation should be relatively stable during the 5 minutes measurement. Thus, the different tidal images do not really change between the start and the end (cf. Figure 21). Plus, at the end, we average the results on each tidal image, as we need a one-dimension results for each feature to feed our chosen estimators that are discussed in the next chapter.

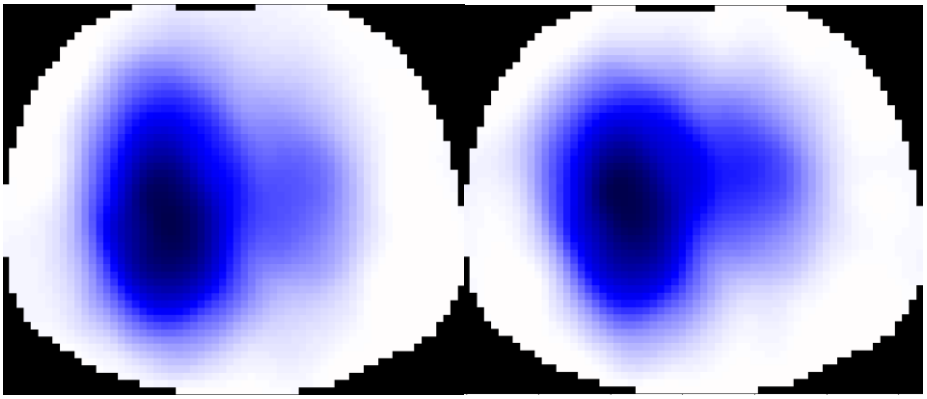

Figure 21 – First and last tidal images for the patient P18 take H2.

Once this step is done, we have removed any false tidal detection from our analysis, though the perturbation is still present in the overall signal. To further clean it, we removed any parts that contain false tidal detection and then we reassemble the signal.

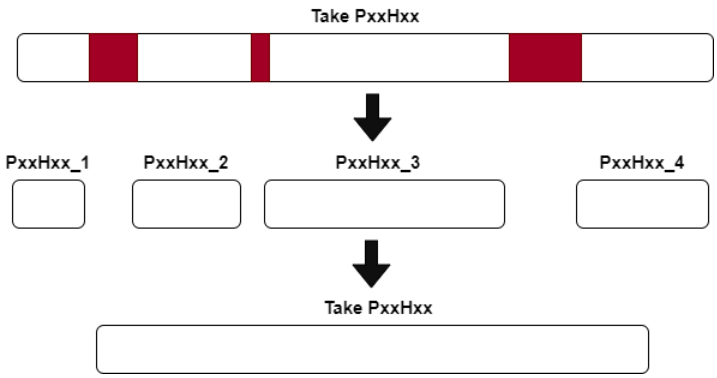


Figure 22 – Reunification of the different EIT parts after removing the false tidal detection due to noise. The false tidal parts are depicted in red.

With this last filtering of the EIT signal, we are left with an average between the measurement set of 3455 frames. This corresponds to using only 38% of the whole 5 minutes frames. The 62% other frames correspond for the most parts as noise that was detected from the voltage signal.

Table 10 - Number of EIT frames kept at each measurement set after removing the part containing noise.

| | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
|---|---|---|---|---|---|---|---|
| Min | 919 | 923 | 1267 | 2153 | 1690 | 3446 | 2019 |
| Max | 6622 | 6608 | 6621 | 4886 | 5432 | 5070 | 5751 |
| Mean | 3244 | 3469 | 3402 | 3175 | 3175 | 4252 | 3465 |
| Std | 1310 | 1225 | 1268 | 866 | 931 | 636 | 1218 |

### 4.1.2. Image reconstruction

The algorithm GREIT is used for the image's reconstruction. To calculate the GREIT reconstruction matrix R, the EIT open-source toolkit EIDORS [72] is used in Matlab. GREIT have many parameters that need to be set for the training of R. Like the target size, which define the size of the target simulation (cf. Figure 11), the distribution of the target or the number of simulations. Thürk and his teams published two articles on the best methods to find the right settings [73], [74]. They recommend an empiric method where different values for each parameter are tested. Then they tested each reconstruction matrix by reconstructing the image and calculating certain EIT features to check against gold standard. They used the same analyze framework proposed by Grychtol et al [56] that was previously explained in the part EIT reconstruction process part in the state of the art. They also used the same data as theirs. Though, as we wanted to have a reconstruction that was optimized using our setup, we used our own data. EIT recordings were realized on 2 healthy subjects using the Drager Pulmovista 500. Two takes of 5 minutes were realized for each subject. Subjects had to ventilate through a spirometer in order to measures the true ventilation volume. In total, over 1200 reconstruction matrix were tested, examples of different reconstruction matrices are shown in the Figure 23.

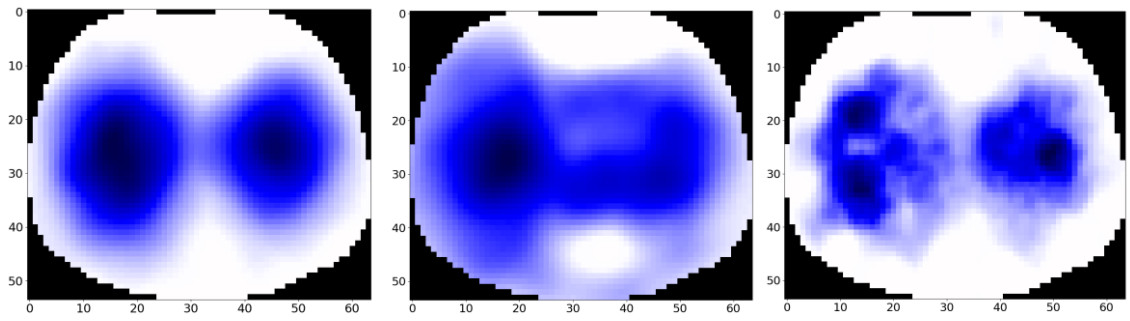

Figure 23 – Reconstruction example of an EIT tidal image using different GREIT reconstruction matrix. The image on the left is from our reconstruction matrix.

Estimation of the tidal volume was calculated for each reconstruction matrix and evaluated against the true volume. The one that yielded the lowest difference in tidal volume was then chosen to reconstruct all images from the clinical study. The best mix of GREIT parameters under the EIDORS implementation is shown in Table 11.

Table 11 – Best GREIT parameters chosen for this study.

| | |
|---|---|
| Image size [pixels] | 64 x 54 |
| Distribution of simulated conductivity variation | Random |
| Size of conductivity variation as proportion of mesh radius [%] | 0.08% |
| Number of simulations | 500 |
| Noise level [Ø] | 0.3 |

The reconstruction matrix is then exported in a Python IDE where the whole framework takes place, which includes the preprocessing, the image reconstruction and the calculation of the EIT features. The image reconstructed from the reconstruction matrix have a size of 54x64 pixels. They are two sorts of EIT image reconstruction: the tidal images and the dynamic images. As mentioned in the state of the art, the tidal images are the difference between the end-inspiration and the end-expiration, thus showing the lung without the residual volume at the end-expiration. The dynamic images are the difference between each EIT frames and a reference. The reference chosen is the end-expiratory frames with the lowest Vmean. This simplifies the interpretation of the EIT images, as it induces that the negative conductivity variations are due to air movement, and positive variations are noise in our case (as the focus is on the ventilation).

## 4.2. Usual EIT features

The only similar study on the subject comes from the prediction models developed by Longhini et al [68]. Indeed, they tested the prediction capability of 3 EIT features. Though, their observation windows after the extubation lasted for 30 minutes. Our study extends this duration to 48 hours, as we want to monitor the whole weaning process. Not knowing initially what kind of ventilation profile we observe in the study, we wanted to have a comprehensive set of EIT features that could capture the information on the EIT images from different perspectives. Therefore, most of the usual EIT features found in the literature were included in our framework. They are three different perspectives from which the EIT images are interpreted: the ventilation distribution, the volume and the frequency (cf. Figure 24).

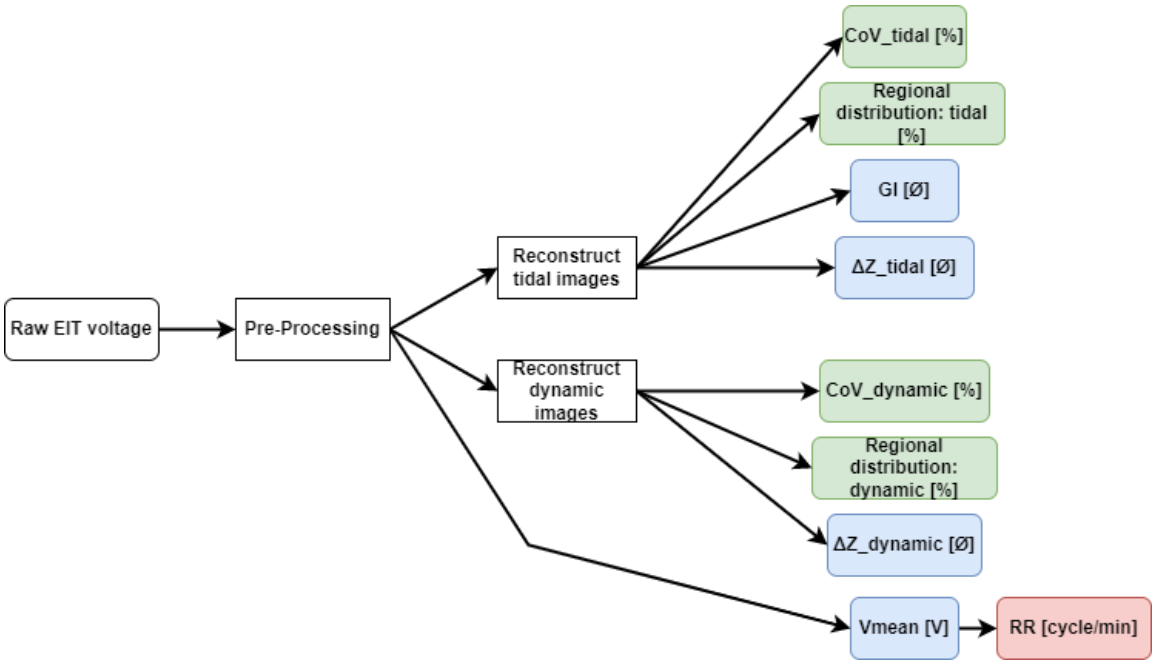


Figure 24 - All usual EIT features calculated. In green, the distribution features; in blue, the features related to the ventilation volume; in red, the temporal feature.

The implementation of the different usual EIT features is reviewed in this part. Except for CoV, Vmean and ΔZ that were explained in the state of the art and does not need further precision. This step is necessary to remove any variation that is not due to ventilation, thus corresponds to noise for our study. Now, unto the explanation of the remaining features.

### 4.2.1. Regional distribution

The regional distribution is a feature that divides the EIT images in four quadrants and calculate the percentage of involvement of each of them. Then, the percentage of involvement of each quadrant is calculated with the following formula:

$$Reg\ distri_{\ q} = \frac{\Delta Z_q}{\Delta Z_{all\ image}} \times 100 \qquad \text{Equation 7}$$

With q, being one of the four quadrants; $\Delta Z_q$ the sum of each pixel value from the quadrant q and $\Delta Z_{all\ image}$ the sum of each pixel from the whole image. The whole process is described in the Figure 25.

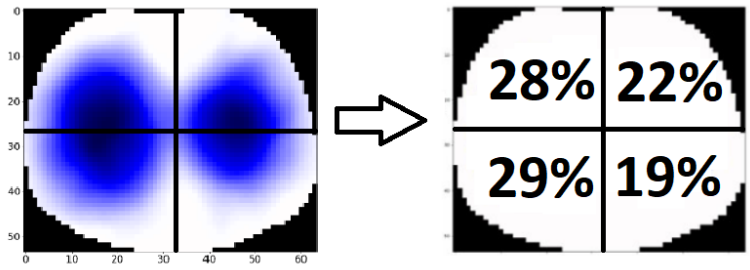


Figure 25 – Example of the calculation of the regional distribution on a healthy subject. On the left is the tidal image with the 4 quadrants; on the right is the percentage of the participation of each quadrant.

### 4.2.2. Global Inhomogeneity Index

For the Global Inhomogeneity index (GI), the equation for the calculation was shown in the state of the art (cf. Equation 4). Though, we left out the explanation about how the lungs are identified in the image. During the state of the art, we came across diverse method to identify the lungs in the EIT image [75]. The most used method in the literature is the uses of standard deviation (std) method. This entail calculating the standard deviation of each pixel during the entire measurement and applying a threshold depending on the maximum value, to remove the background noise of the EIT images. This method works well for identifying the ventilation in the thorax. However, upon a patient who does not have any ventilation in one of his lungs, the standard deviation method will not detect this lung. This would result in a bad estimation of the GI. Indeed, the GI calculates the homogeneity of the volume distribution over the whole lung and not just the ventilated area. We can observe an illustration of this problem from the patient 5 that had atelectasis on the left lung (cf. Figure 26). Only the ventilated area was detected. Therefore, we needed a method capable of identifying the whole lung region, even if no ventilation occurs.

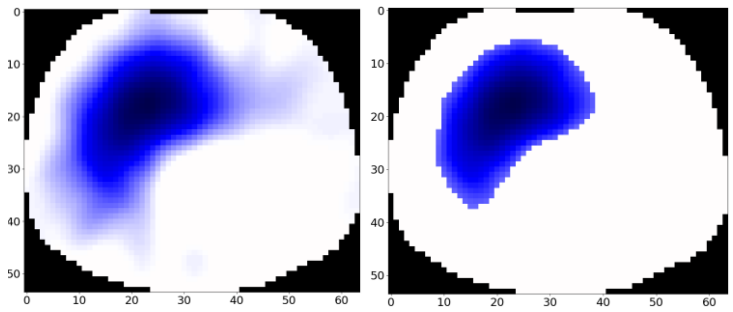

Figure 26 - Uses of the Std method to find the ROI in the EIT images. On the left, the original tidal image; on the right the ROI detected.

To suits our need, we designed our own method to identify the lung. We use the EIT data from measures on 2 healthy subjects. Two measures of 5 minutes were realized while the subject was seated and asked to ventilate in a normal fashion. The standard deviation method was then

applied to those measures with a threshold of 30% of the maximum value. We use the threshold recommended from study conducted by Frerichs et al [76]. The results for the 4 measures in shown in the Figure 27. With those, we combined them to form the final binary mask that will be applied to all EXIT patients. To combine them, we opted to keep every pixel that is present in each of the 4 EIT images. This gives the overall pixel mask that can be seen in the Figure 28.

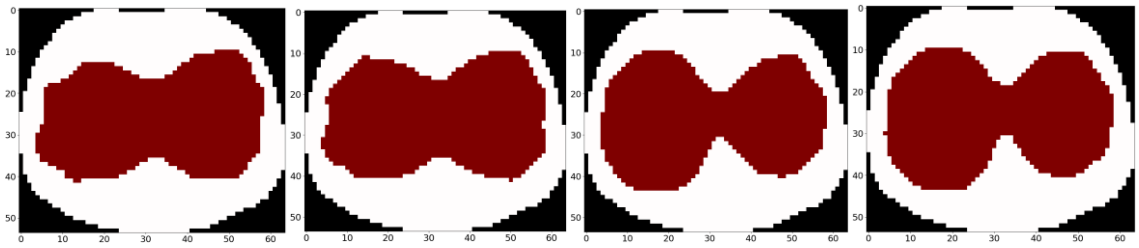

Figure 27 - Lung detection after the 30% threshold applied. The two images on the left are from the 2 measures from the subject 1, and the other 2 from the subject 2.

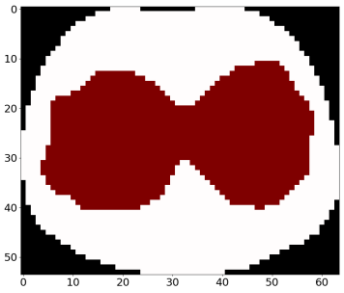

Figure 28 - Overall pixel mask to detect the lungs in the EIT images.

This new method allowed to have a GI result more aligned with other studies. Indeed, with the Std method, we had a GI equal to 0.28 with the example from P05h24 (cf Figure 26). Such result seems inaccurate, as previews study showed that for healthy subjects, GI ranged from 0.4 to 0.5 [69]. As for patient having ventilation distress, GI range from 0.5 to 1.5. With our ROI method, the GI is equal to 0.65 which is more in line with what we expect.

This our ROI method was designed to fix the problem that occurs when the lungs are not properly ventilated. With it, we are able to calculate the GI in accordance with other studies. Though, while this method works, it is somewhat rude in the making. Only 2 healthy subjects were used to construct the ROI. Which does not allow for a good generalization of the model. Plus, other parameters could have been considered to realize more custom ROI depending on the patient's profile. Like considering the patient's morphology or his body mass index.

### 4.2.3. Respiratory rate

For the respiratory rate, we use the Vmean signal. The method is to count the number of inspiration points (cf. Figure 29) over the 5 minutes measures. The count is then divided by 5 to express the results as breaths/minutes.

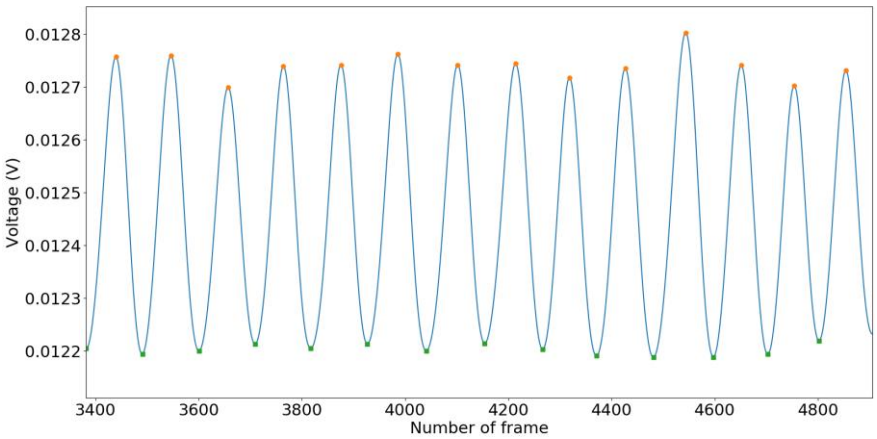


Figure 29 - Detection of the inspiration point (orange) and expiration point (green).

### 4.2.4. Other features

EIT features such as the compliance or calibrating the EIT data to express the results as a volume variation could not be realized. Indeed, the ventilator data was only available at H0. Therefore, those metrics would only have been calculated for only one measurement set.

In addition to all EIT features included, we are also included another version of each of them. It has been shown the coefficient of variation of the tidal volume is a better extubation predictor that the actual tidal volume [25], [26]. From this spirit, we calculated the coefficient of variation of every EIT features.

To finish, we can note that some features are calculated on the tidal image. While for some, we also calculated the same feature on the dynamic images (cf. Figure 24). Which corresponds to each EIT frames measures during the 5 minute period. We then only used the average of the features from the 5 minutes as we are limited from our chosen prediction's estimator. Indeed, they do not allow for vector input.

## 4.3. Added features

### 4.3.1. Features based on existing ICU metrics

We wanted to add EIT features that could be more recognize by clinicians. To do that, we had to select ICU metrics that could be estimated from EIT measurements. We opted to estimate the RSBI and the ventilation flow. The variation of both metrics could be estimated through the variation of ΔZ. Both metrics are interesting as they mix the temporal aspect of the ventilation with the volume. Thus, bringing a new optic to the EIT analysis.

The $RSBI_{EIT}$ is calculated by dividing ΔZ by the respiratory rate RR. From this global formula, we decline it into 2 features, $RSBI_{EIT}$(overall) and $RSBI_{EIT}$(cycle) that are calculated from the Equations 8 and 9. $RSBI_{EIT}$(overall) calculate the overall RSBI, with ΔZt which is calculated by summing the pixel from an average tidal image from the 5 minutes take. $RSBI_{EIT}$(cycle) calculates another version from only one tidal cycle, with Ti being the inspiration time.

$$RSBI_{EIT}(overall) = \frac{RR}{\Delta Zt(5min)} \quad \text{Equation 8}$$

$$RSBI_{EIT}(cycle) = \frac{Ti}{\Delta Zt} \quad \text{Equation 9}$$

The second added feature is the ventilation flow. It is calculated from the tidal images (inspiration frame minus the expiration frames) as well as the expiration images (expiration frame minus the inspiration frames). Those two-ventilation flows are calculated from the equation below. With Te, being the expiration time and ΔZte, the conductivity variation of the expiration image.

$$flow_{EIT}(inspiration) = \frac{\Delta Zt}{Ti} \quad \text{Equation 10}$$

$$flow_{EIT}(expiration) = \frac{\Delta Zte}{Te} \quad \text{Equation 11}$$

### 4.3.2. Features based on the lung's appearance on EIT images

With the first patients included, we reviewed the EIT data to check if we could improve our analysis framework. From it, we could observe that the ventilation distribution was different in

most cases between the two classes (cf. Figure 30). Indeed, for patients succeeding the extubation, we can observe EIT images that look like lungs. However, for patients failing it, the ventilation looks particularly odd, and the lungs are not clearly visible.

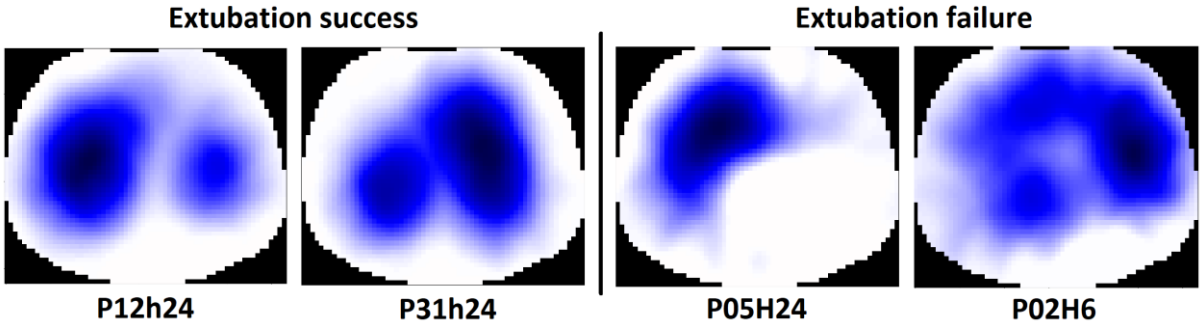


Figure 30 - Tidal images of different patients.

To try to capture this aspect, we created 2 EIT features, lung area, which capture the ventilation surface measures on the EIT images; and lung shape, which captures the ventilation appearance. Both EIT features are calculated on the tidal images and on the left and right sides separately to monitor both lung ventilation. To measures the features, we first binarized the tidal EIT images after applying a 20% threshold of the maximum pixel value to remove any background noise (cf. Figure 31).

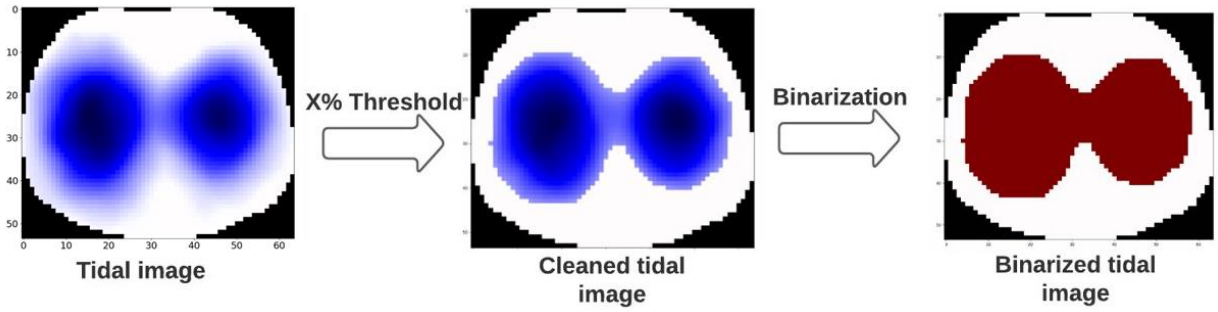


Figure 31 - Binarization process of the tidal images.

From the binarized images, the lung area is calculated by counting the number of pixels presents in 4 areas: right, left, front and back (cf. Figure 32). While lung shape is calculated by counting the height and width of both lungs. In addition, to have a feature combining both, we calculated the compacity of the ventilation (cf Equation 12). This metric is also applied on both sides.

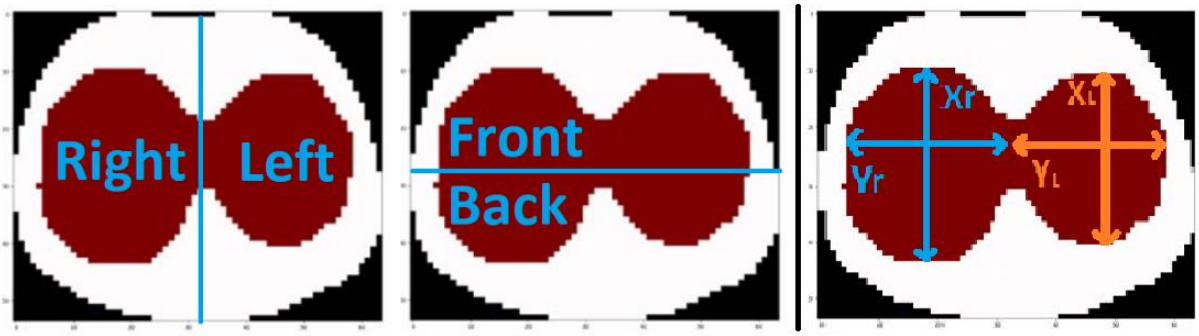


Figure 32 – On the left, the lung area measures on the lung; on the right, the lung shape measures. $X_r$ and $Y_r$ are the height and width of the right lung. $X_L$ and $Y_L$ are the height and width of the left lung.

$$Lung\ shape_{compasity} = \frac{Y}{X} \qquad \text{Equation 12}$$

## 4.4. Conclusion

This chapter presented the EIT framework. With the EIT voltages filtering to remove any noise present in the signal. Then, we presented the different features that are calculated from the reconstructed EIT images. In total, we compute 61 EIT features. From of all those, they will have to be sorted to examine which features are best to separate the success from the extubation failure. Which will give us an insight as to what kind of ventilation profile a failing patient has, with the EIT perspective. Now, unto the next chapter where the different prediction models are details as well as the training method.

# 5. Framework to construct and test the prediction models.

*We had 3 main concerns at the beginning:*

1. *We are dealing with a database with few examples.*
2. *The ratio success patients versus failure one is clearly not balanced in favor for the success class.*
3. *The number of patients per measurement set H is not constant.*

*As we will discuss in this chapter, those concerns greatly weighted in the conception of our framework, as we had to work around them to find the best solution.*

*For the prediction, we decided to focus on the prediction of the failure patient. The extubation outcome for the failure is then set as the positive class and the success are set as the negative one. Plus, we decided to realize unique new prediction model for each measurement set. This allows to update the prediction with the new data and allows to disregard the third concern with the unbalanced number of patients between measurement sets.*

*In this chapter, we are first going to present how the EIT data is used. Then how the model training is organized. To conclude, we show the first prediction results without any optimization.*

## 5.1. Learning model based on the EIT features

For this study, we primarily focus on detecting patients failing the extubation. Though, from our database, there are no more patients in the failing group after the measurement set H24. Thus, the data gather at H36 and H48 were not used to build prediction models. So, the prediction model was created only for the measurement set H0, H2, H6, H12 and H24.

As briefly mentioned in the introduction of this chapter, the dataset that we are working with includes 3 main challenges. For this part, the most constraining challenges from the 3 is the last one, which correspond to the variability of the number of patients per measurement set. Having fewer patients in our database through time is inherent to this kind of study, as it is bound to have patients failing the extubation at different hours. Though, other factors contributed to this variation (cf. Table 7):

- Whole EIT measures could not be taken into account due to the presence of too much noise in the signal (cf. Figure 16 for example). Those represents 20 measures in total from both classes combined.
- Certain measures were not realized as they occurred during the night, which includes a more limited medical crew present in the ICU. In total, 11 measures were not realized due to this limitation.

- Finally, 2 patients were discharged from the ICU as their ventilation condition was now stable and healthy.

This inconsistency of patients in the database limited us in our possibility to organize it. To overcome any problem that could occur from this inconsistency, we decided to realize new prediction models at each new measurement set (Hx). Rendering each new prediction model independent from the other. Plus, this way we are able to re-evaluate our prediction at each new set, and hopefully improving our prediction results. A drawback from this approach, though, is that we are not taking into account the temporal aspect, that would be possible when more data will be available.

During the conception of the models, we also had to decide how we consider the failure patients. There are two possibilities:

1. patients are declared as extubation failure only on the measurement set (H) preceding the failure;
2. patients are set as failure from the start observation period whatever measurement set they are failing.

Table 12 – Description of the failure database using the first option.

| Patient | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
|---|---|---|---|---|---|---|---|
| P02 | | | 1 | | | | |
| P03 | | | | | 1 | | |
| P05 | | | | | 1 | | |
| P06 | | | | 1 | | | |
| P09 | 1 | | | | | | |
| P13 | | | | | | | |
| P14 | | | 1 | | | | |
| P27 | 1 | | | | | | |
| P29 | 1 | | | | | | |
| P36 | | | | | 1 | | |
| **Total** | **3** | **0** | **2** | **1** | **3** | **0** | **0** |

| | |
|---|---|
| | Good measure |
| | ICU discharge |
| | Measure not analyzable |
| | Measure not realize |
| | Re-intubated patient |
| | Ventilation distress |
| | Study exit |

The first option is closer to the ground reality as from a clinical point of view, a patient is not failing the extubation until he is not re-intubated. Though, it is possible that experienced clinicians could see the failure to come. Albeit this option being closer to reality, we decided to go for the second option, as the first one has a major drawback for the prediction model.

By defining the patient has failure just moments before it actually happens (first option), we are left with very few failure patients in the database at each measurement set. This can be seen in the Table 12 which presents the number of patients that would be counted as failure to use the first option. Such implementation would therefore render the construction of the prediction models not easy. However, we would not even be able to have a prediction model at H2 as there is no failure. Therefore, as stated we use the knowledge that we have of the real outcome, to set the patients as failure from the start. This method offers the advantage that the model could recognize the ventilation distress way before it actually happens.

Now unto the learning model, we have 61 features that were extracted from the EIT images. As they have never been previous experiment where EIT measures were realized to observe the weaning phase, we did not know what kind of evolution we should expect.

Therefore, we decided to test different hypotheses for sorting the EIT features, to then feed the prediction model. In total, we tested 3 hypothesis which are described as follows:

- **Hypothesis 1 - Single H:** at each epoch i, the measurement set H(i) is used to learn a prediction model. Only the EIT features from that epoch i are used for the training (cf. Figure 33). This is the most basic model. Though if it proves to be accurate, it will make it easier to use the EIT to predict the extubation.

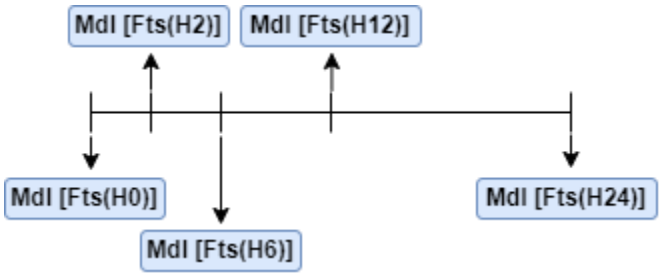


Figure 33 - Illustration of the Single H model

- **Hypothesis 2 - Variation H:** at each epoch i, the measurement set H(i) a prediction model is learned, and each dataset is composed of the difference of each feature computed between the measurement set H(i) and H(i-1) (cf. Figure 34). With this model, we wanted to ascertain whether calculating the ventilation behavior variation between two sets carries a more accurate prediction.

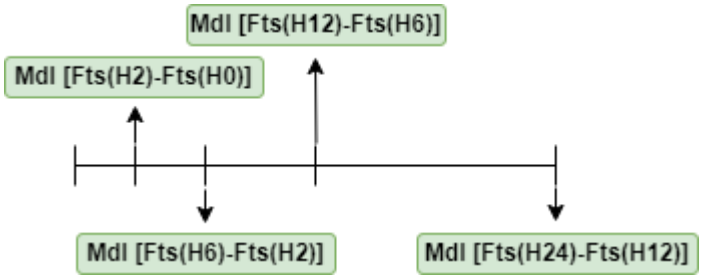


Figure 34 - Illustration of the Variation H model.

- **Hypothesis 3 - Global H:** between measures from two epochs on patients, we do not have the same measurement condition. Indeed, the electrode belt could have moved slightly, the patient do not have the same position and the contact electrode/skin changed. Thus, each new EIT measures are independent even for the same patient. Thus, we want to test a more global prediction model that would treat each new measure independently. Therefore, for each new measurement set H(i), we then integrate the new examples to a whole database that include each past measures for all patients.

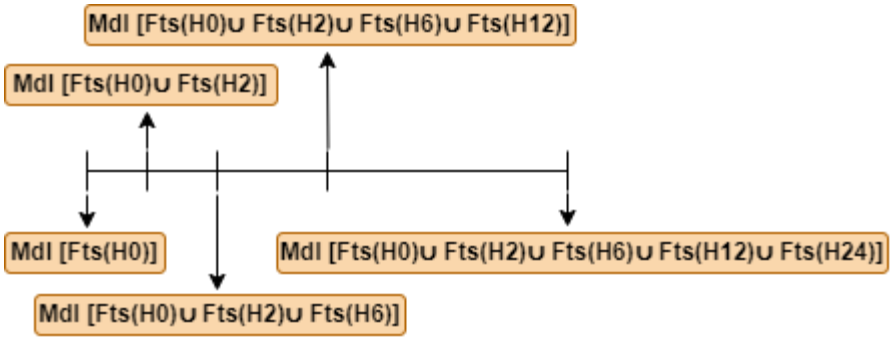


Figure 35 - Illustration of the Global H model.

## 5.2. Cross-validation

The limited number of patients present in our database, as well as the unbalanced distribution of the two classes (success and failure) is another challenge that we need to address for the learning dataset.

With those concerns in mind, we searched for a method that could handle such database. First, we need to organize the database into 3 set: training, testing and validation set. The training set is the set that is used to train the model. The testing is the one used to test the prediction capability of the trained model. Data present in this set can also be used during the training during the cross-validation step. At last, the validation set is used to compute the prediction performance of the model. The difference with the testing set is that data present in this set was

left untouched during the training set. Thus, this is a test on an independent set to study the generalization of the model.

In order to not overfit the model to the testing set, most solution divide and optimize their model by training on subsets of data randomly pick from the training set. This method works greats on usual datasets. Though our lack of patients and even more so, our lack of enough failure patients, made it difficult. By using those methods, we could only manage to realize 2 or 3 folds of subsets. After which, we would not have any more sufficient patients of the failure classes to make any more folds. With such method, our prediction model would be based on very few examples, making it hard to have a generalized model.

For those reasons, we decided to realize our own cross-validation-like algorithm. Our main concern was to have a training set balanced between the success and failure patients so that we would not introduce any bias into the model.

Our method consists in separating the dataset into two sets: the training set (70% of the data) and testing set (30%). The limit values of examples to take from each group (failure or not) is set to 70% of the examples of the failure class (cf. Equation 13).

$$Nb\ of\ examples = 0.7 \times Nb\ of\ failure(H_i) \qquad \text{Equation 13}$$

We then randomly pick patients on both the success and failure group, until both groups reach the number of examples fixed earlier. This ensures that the training set is balanced. The corresponding test set is composed of all the examples not picked.

$$Training\ set = \begin{cases} success\ patients\ (Nb\ of\ examples) \\ failure\ patients\ (Nb\ of\ examples) \end{cases}$$

The Figure 36 shows an example of this process with 12 patients. This process is repeated for each measurement set, thus the limit of patient change depending on the number of patients failing the extubation.

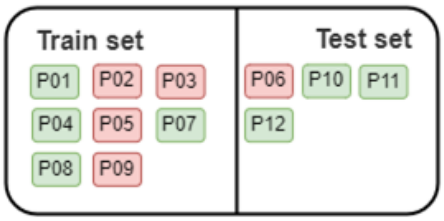


Figure 36 - Example of dataset separation with 12 patients.

In addition to those rules, we decided to pick the last included patients to be part of the validation set. The patient P33 to P37 are then chosen for this application. From those, the patient P35 is not included as the EIT data from him are not analyzable. Which left 4 patients, with one extubation failure among them. We then have 25% of failure patients in the validation set, same as on the whole database.

When this cross-validation rule is applied to our dataset with the different prediction, this results in the dataset separation that is details in the Table 13, Table 14 and Table 15. From those, it is observed that for two last measurement sets of each model, they are very few examples in the training sets, as it is limited by the number of extubation failures. However, this would have been a difficult problem to avoid giving the nature of the protocol. As first, they are around 10 to 25% chance to fails the extubation. Second, the risk of failing the extubation decrease throughout the hours after the extubation. Thus, there are only a few patients that failed the extubation at the H24 mark. A solution to this problem would have been to either recruits more patients in the hope to have more patients reaching that stage. Or recruiting patients with even higher risk of failures, though they would still need to pass the SBT to be extubated.

Table 13 - Number of patients per measurement set (Hi) for the Single H model.

| Single H | | H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|---|---|
| Total | Success | 24 | 23 | 22 | 14 | 17 |
| | Failure | 8 | 5 | 5 | 2 | 2 |
| Training | Success | 6 | 4 | 4 | 1 | 2 |
| | Failure | 6 | 4 | 4 | 1 | 1 |
| Testing | Success | 18 | 19 | 19 | 13 | 15 |
| | Failure | 2 | 2 | 2 | 1 | 1 |

Table 14 - Number of patients per measurement set for the Variation H model.

| Variation H | | H2-H0 | H6-H2 | H12-H6 | H24-H12 |
|---|---|---|---|---|---|
| Total | Success | 23 | 22 | 13 | 10 |
| | Failure | 5 | 5 | 2 | 1 |
| Training | Success | 4 | 4 | 1 | 1 |
| | Failure | 4 | 4 | 1 | 1 |
| Testing | Success | 19 | 19 | 12 | 9 |
| | Failure | 2 | 2 | 1 | 0 |

Table 15 - Number of patients per measurement set for the Global H model. The orange number are the added data from the past measurement set.

| Global H | | H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|---|---|
| Total | Success | 24 | 23+24 | 22+47 | 14+69 | 17+73 |
| | Failure | 8 | 5+8 | 5+13 | 2+18 | 2+20 |
| Training | Success | 6 | 4 +24 | 4+47 | 1+69 | 2+73 |
| | Failure | 6 | 4+8 | 4+13 | 1+18 | 1+20 |
| Testing | Success | 18 | 19 | 19 | 13 | 15 |
| | Failure | 2 | 2 | 2 | 1 | 1 |

This whole cross-validation process is repeated 100 times to test most possible mix of patients. For each new cross-validation sets, a new prediction model is created based on this set. The final prediction result is then based on the results of those 100 prediction models. An example of how many times a patient is chosen for the training set is shown in the Figure 37. The distribution is not perfect as we do not have lots of iterations. Though, we can see that all patients are relatively all use the same number of times. Except, of course, for the difference between the success and failure patients, which was expected as they are way more success patients.

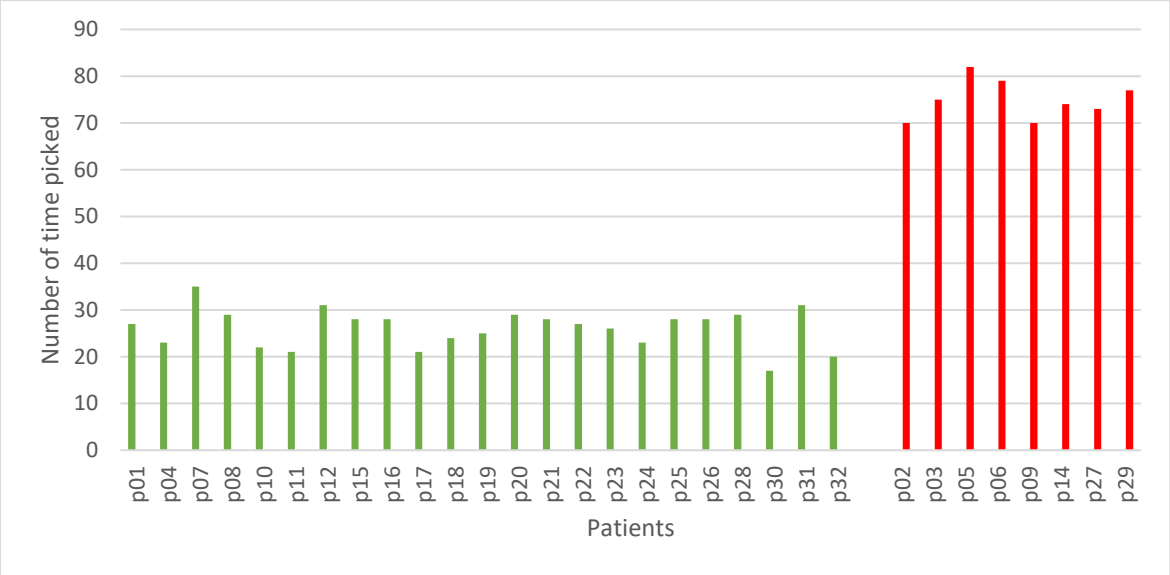


Figure 37 – Example of the number of times each patient is used in the training set with the 100 iterations (data at H0). In green, patients succeeding the extubation; in red, patients failing it.

## 5.3. Estimators

In this study, we are treating a supervised learning classifications case, as we knew the extubation outcome which can fall in only two categories, success or failure. With that in mind, we chose to use 3 estimators for the prediction, which are the decision tree, the random forest and the Support Vector Machine (SVM). Those estimators were chosen for two main reasons:

1. they are applicable on small databases;
2. the prediction results are easier to understand, in contrast to more complex algorithm like neural networks.

### 5.3.1. Decision tree

When using a decision tree, the algorithm builds up the prediction over a set of simple rules that were attributed to the features during the training [77]. Those rules correspond to thresholds that are applied on the features as well as the composition of the branch. This type of estimator renders the prediction easy interpretation to understand, as it just applied the rules that it learns during the training. Moreover, it can provide insightful information about which features are more relevant than other, by looking at the order of the used features and their repeated use in the tree. This kind of inference model can be interpreted by a physician in order to understand how the classification is made. To understand that aspect, we are presenting hereafter how this algorithm works. Before diving into the explanation, the Figure 38 shows the terminology used with this estimator.

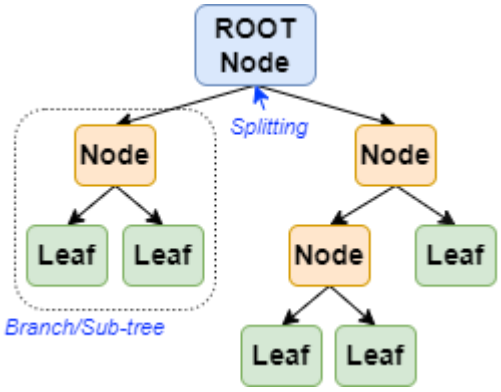


Figure 38 – Terminology related to Decision tree

The main challenge of this estimator is to find the best feature to use at each node to separate data. Indeed, the goal of the tree is to have leaves that are pure, meaning that only one class is present in it. To decide which features to use at each node, the algorithm uses a criterion. Two of the most widely used criteria are given here:

- shannon entropy: this criterion calculates the randomness of the information among the different classes. It is calculated by the **Erreur ! Source du renvoi introuvable.**, with $p_j$ the probability to belong to class j. To illustrate this, entropy equal to 0 would indicate that there is only one possible outcome. Whereas entropy reaches its

maximum for equiprobability of the classes. During the building of a tree, at each node, the feature that minimizes the entropy is selected to split the data;

$$Entropy = -\sum_j p_j . log_2 . p_j$$ Equation 14

- gini: the frequency of mislabeling data when it is randomly labeled. This is a similar idea as with the entropy but realized it in a different manner. As the entropy, the feature that minimized the Gini index value is then selected.

$$Gini = 1 - \sum_j p_j^2$$ Equation 15

One of the drawbacks of the Decision tree estimators is that it is prone to overfitting. Indeed, the algorithm creates new nodes until all leaves are pure. To solve this issue, pruning techniques can be used. This reduces the number of nodes or leaf and prevent overfitting to the training data.

### 5.3.2. Random forest

The random forest is an ensemble method [78], it corresponds to a method that uses multiple estimators to base their prediction. They are two kinds of ensemble method:

- bagging: this method uses a several estimators that all realized their prediction on data. The final prediction of an example is decided by voting between the estimators [79];
- boosting: it arranges the estimators in a sequential manner. This allows the next estimators to learn from errors produced by the previous one, which in turn improve the accuracy [80].

The random forest is a bagging method and uses decision tree as estimators. Though, one additional parameter to consider when using it, is the number of estimators to uses for the prediction.

This method returns more accurate prediction than a simple decision tree, due to the voting mechanisms that take place. Though, it is slower to compute due to the highest complexity.

### 5.3.3. Support Vector Machine (SVM)

The SVM attempts to find a hyperplane to separate the different classes. To do that, the algorithm calculated the vectors that link the point to the line. It then tries to maximize the

distance of all the different vectors, in order to have the best separation between the classes [81].

Data can have all sorts of distribution which is sometimes difficult to separate. For such problem, two parameters can be played into:

- the features to uses to represent the data in the space, this allows to have a different distribution which could render the separation easier;
- the kernel, which transforms the data into a representation space of higher dimensions where they could be linearly separated.

Common kernels are linear, sigmoid, RBF (Radial Basis Function) or polynomial kernels (cf. Equation below, X and Y refer to 2 different features).

$$Linear\ kernel: K(X,Y) = X^T Y \quad \text{Equation 16}$$

$$Sigmoid\ kernel: K(X,Y) = tanh\ (\gamma . X^T Y + r) \quad \text{Equation 17}$$

$$RBF\ kernel: K(X,Y) = exp(-\gamma \|X - Y\|^2) \quad \text{Equation 18}$$

$$Polynomial\ kernel: K(X,Y) = (\gamma . X^T Y + r)^d \quad , \alpha > 0 \quad \text{Equation 19}$$

## 5.4. Scoring methods

For the prediction model, we want to forecast the patient's extubation failing, because the percentage of comorbidity of those patients is too high. This outcome will help clinicians make the best decision in order to take the most suitable treatment. For example, they could use NIV as a first means treatment to prevent the re-intubation. Therefore, we defined the failures as the positive class and the success as negative class.

To score and compare the different prediction models, we are using the sensitivity and the specificity. The equation of both metrics can be seen below. With TP, the True Positive; TN, the True Negative; FP, the False Positive and FN, the False Negative. The sensitivity informs about how the model is able to recognize positive cases: it is valued as the probability of correct predictions of the positive class. The specificity informs about how the model is able to recognize negative cases: it is valued as the probability of correct prediction of the negative

class. Specificity and Sensitivity are used to evaluate separately the accuracy of the model when predicting positive and negative examples.

$$Sensitivity = \frac{TP}{TP+FN} \qquad \text{Equation 20}$$

$$Specificity = \frac{TN}{TN+FP} \qquad \text{Equation 21}$$

For the best learning model, we will also calculate the sensitivity and specificity for the success class (the success will be defined as the positive class and failure as the negative one).

It is important to note that the given results are evaluated for each inference model trained over the 100 iterations. Meaning that for each iteration, we count how many patients were predicted as TP, TN, FP and FN. All those counts are then added at the end to calculate the sensitivity and specificity of the model.

For the prediction results, we are also calculating the percentage of variation between the different version of the same model. This will better demonstrate the gains and losses. It is calculated through the following formula:

$$Var[\%] = \frac{Value_{final} - Value_{initial}}{Value_{initial}} \times 100 \qquad \text{Equation 22}$$

In addition to those two metrics, we also use the Cohen kappa coefficient for the model that generated the best prediction results. This coefficient value is used to measure the reliability between different raters. In other words, it measures the concordance of answers given between two instances. In our case, the two instances are the real extubation outcome versus our prediction model. It is calculated from the following formula:

$$K = \frac{p_o - p_e}{1 - p_e} \qquad \text{Equation 23}$$

With $p_o$ being the proportion of agreement between the raters and $p_e$ the probability of a random agreement. The Cohen kappa coefficient results are interpreted as:

- Inferior to 0: Poor agreement
- 0.01 – 0.20: Slight agreement
- 0.21 – 0.40: Fair agreement
- 0.41 – 0.60: Moderate agreement
- 0.61 – 0.80: Substantial agreement
- 0.81 – 1.00: Almost perfect agreement

## 5.5. Proposed training framework

With every different aspect of the prediction framework that is detailed in the previous part, we can show how they are connected. The Figure 39 shows the layout for the whole framework, which we resume hereafter.

This first start with all the EIT features calculated from the EXIT patients. We then have an array with 32 patients, each having 61 EIT features results. This array is then rearranged depending on which hypothesis we are testing, and it then feeds the cross-validation function, which separate it into the training and testing set. The training set is then used to train the prediction model using the 3 different estimators presented in section 5.3. Once the training is done, the generated model is used on the testing set to measure the performance of the model. The scoring counts the number of true positive, true negative, false positive and false negative. A scoring is also calculated for the validation set which is composed of Patient P33 to P37.

With the first iteration realized (cross validation + model training + scoring), a new iteration starts with a new dataset generated by a new random pick during cross-validation. For each hypothesis and each inference model (Decision Tree, Random Forest, SVM) tested, there are 100 iterations are realized. To score each whole prediction model (the results of the 100 iterations), we sum up the score of each iteration (Cf. Equation 24 for example on TP), and we calculate the sensitivity and specificity results based on those counts (Cf. Equation 25 and Equation 26).

$$TP_{whole} = TP_{iteration1} + TP_{iteration2} + \cdots + TP_{iteration100} \qquad \text{Equation 24}$$

$$Sensitivity = \frac{TP_{whole}}{TP_{whole} - FN_{whole}} \qquad \text{Equation 25}$$

$$Specificity = \frac{TN_{whole}}{TN_{whole} - FP_{whole}} \qquad \text{Equation 26}$$

It is important to take note that with each hypothesis or measurement set, we have a unique new model. This is due to the different composition of patients in the dataset. The way of arranging the EIT data through the 3 hypotheses, results in datasets with different numbers of patients included (Cf. Table 13 to Table 15). And each measurement set is composed of a different set of patients (Cf. Table 7). With that difficulty in mind, we can still compare the results between the different hypothesis, as we are interested in observing which one is more adequate to our studies. For the comparison between measurement set, the nature of this clinical study makes it impossible to have the same number of patients each time. As patients failing the extubation

can be re-intubated at any time. So, we still compare and observe the evolution between measurement set.

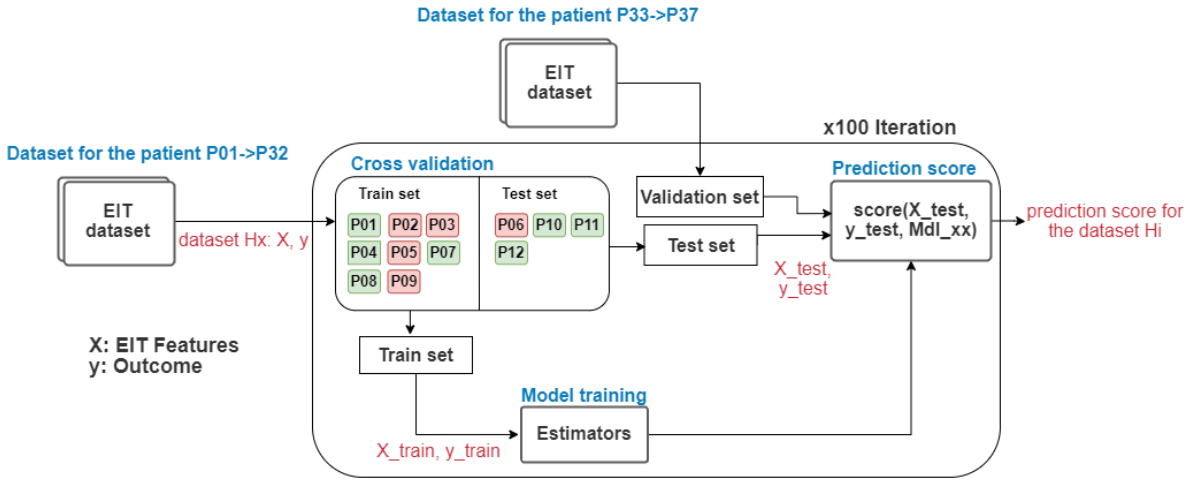


Figure 39 - Framework of the training and testing of the prediction models.

## 5.6. Results with the defaults settings

To have a first understanding of the different models, we use the 3 estimators with their default hyperparameter implemented in Sickit-learn package. This gives us a baseline from which we want to improve our results. The hyperparameter used for this test are described in Table 16.

Table 16 - Default hyperparameter used in the implementations of the Sickit-learn package.

| | | |
|---|---|---|
| **Decision tree** | Criterion | Gini |
| | Max depth | Until leaf is pure |
| **Random forest** | Nb of tree per forest | 100 |
| | Criterion | Gini |
| | Max depth | Until leaf is pure |
| **SVM** | Kernel | RBF |
| | Regularization parameter | 1 |

The sensitivity and specificity results for all models can be seen in Figure 40, Figure 41 and Figure 42. The results are also displayed as table in the Annex (cf. Table 32,

| P01->P32 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.47 | 0.61 | 0.45 | 0.69 | 0.31 | 0.8 |
| H2 | 0.53 | 0.61 | 0.59 | 0.72 | 0.66 | 0.63 |

| | | | | | | |
|---|---|---|---|---|---|---|
| H6 | 0.45 | 0.62 | 0.4 | 0.74 | 0.47 | 0.74 |
| H12 | 0.39 | 0.51 | 0.18 | 0.58 | 0.22 | 0.58 |
| H24 | 0.47 | 0.55 | 0.43 | 0.59 | 0.49 | 0.59 |
| **Mean** | **0.46** | **0.58** | **0.41** | **0.66** | **0.43** | **0.67** |

Table 33 and Table 34). From all the different hypotheses tested, the Global H model seems to perform best. The three of them starts relatively with the same sensitivity results. Then, the Global H sensitivity increases throughout the 48 hours. While the sensitivity of Single H and Variation H decrease.

The best estimator at this stage for the Global H model is the Decision Tree, with a sensitivity of 0.46; then the SVM, with 0.43; and the Random Forest with 0.41 (cf. Table 34).

However, most of the models do not yield good results, as the overall sensitivity average is around 50%. Though, in terms of specificity, the Global H model yields very good results with an average specificity between the estimators equal to 0.84.

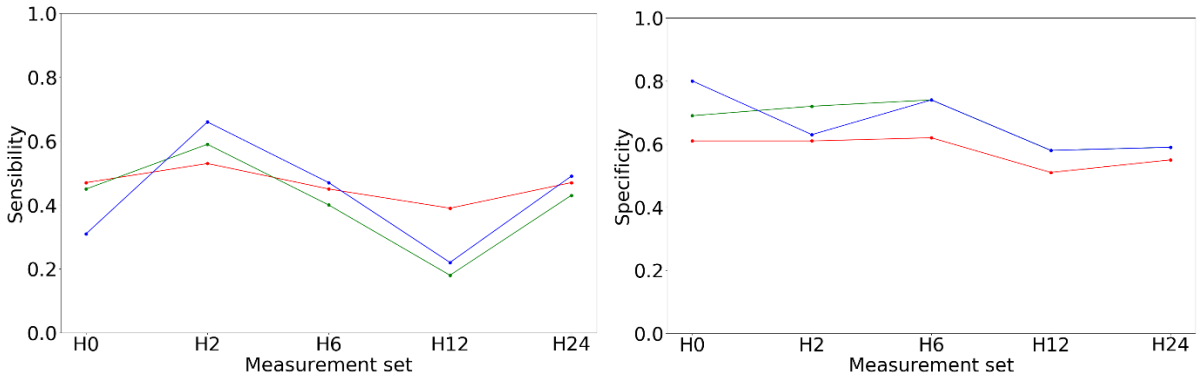


Figure 40 - Baseline results - Single H - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM.

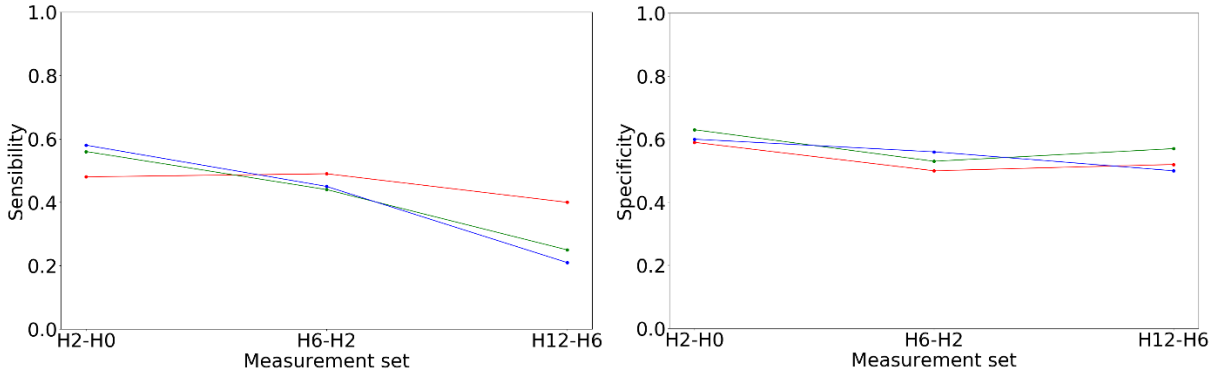


Figure 41 - Baseline results - Variation H - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM.

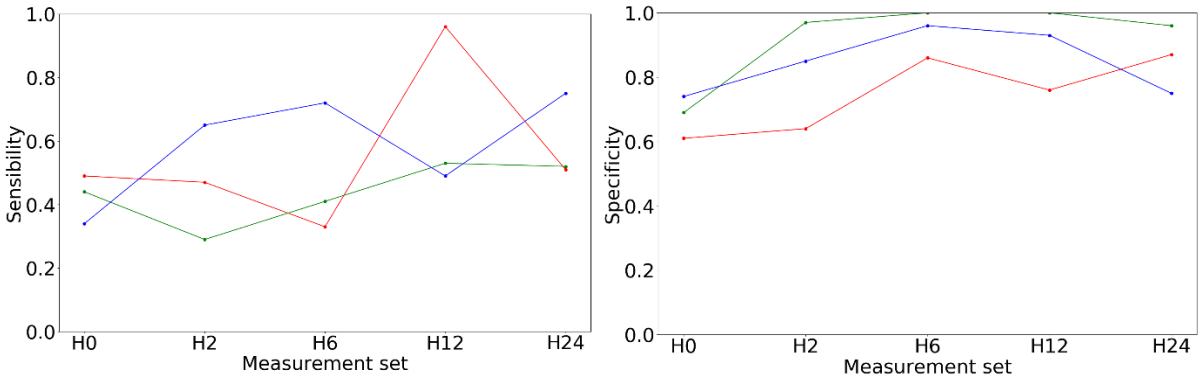


Figure 42 - Baseline results - **Global H** - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM.

If we dive further in the best hypothesis, Global H, we can observe from the Table 17, that more depth was needed after each measurement set. This can be explained by the addition of the added data in the Global H model. Thus, requiring more rules to better predict patients.

A table showing the most and least used EIT features for the Decision Tree and Random Forest can be seen in the Annex (Cf. Table 35). We can note that for both models, coefficient of variation features had the most used for the start of the weaning phase (H0 and H2). Then, the most important features are the $Flow_{EIT}$. The least uses features are for all measurement sets, features regarding the ventilation volume.

Table 17 - Depth of each the Trees for the 100 trained estimators. The Random Forest results show the depth's average of the Tree present in the Forest.

| | | H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|---|---|
| **Decision Tree** | Min | 1 | 3 | 4 | 6 | 6 |
| | Max | 3 | 6 | 7 | 6 | 7 |
| **Random Forest** | Min | 1 | 4 | 5 | 5 | 5 |
| | Max | 3 | 5 | 6 | 6 | 6 |

The results on the validation set are displayed on Table 36, Table 37 and Table 38 (see Annex). First, we have to note that unfortunately, the sensitivity could not be calculated for the take H6 and H12, as the failure patients (P36) do not have any valid measure for those measurements set. Overall, from those results, the same conclusion with the other set can be made. Meaning that the best model is still the Global H one, using the SVM. Interestingly though, we can observe that the prediction model has more difficulty to correctly identify the success patients.

## 5.7. Conclusion

In this chapter, the prediction models and methods used for the extubation prediction are explained. Dealing with a database with few examples and imbalance class distribution

highlights some challenges. Adaptation from usual framework had to be realized to address our problem. From the first result showed, we now have to observe if this initial trend in the results does continuous with the optimization that we bring forth in our model.

# 6. Optimization of prediction models

*This chapter display all of the prediction results, with the two different sorts of optimization. The best model is then better analyzed to attempt to understand how the prediction is realized. To finish, we compare our method with the one proposed by Longhini and his team* [68]. *Plus, a more overall comparison with the best prediction model from the scientific literature.*

## 6.1. Fine-tuning process

They are several solutions that exists to improve the results from machine learning algorithms. For instance, adding more examples would help to have a more generalized model which could handle more different situations. Thus, improving the prediction results. As this method is not an option for us, we are employing other means to improve our results.

In this part, we are focusing on optimizing the hyperparameters of the 3 estimators. There are no direct methods to find the optimal hyperparameters. Indeed, we do not know in advance how the hyperparameters will impact the results. A common method to find the best combination of hyperparameters is to do a grid search. This entails testing a list of values for each hyperparameters and to choose the combination values that return the best results. For our problem, the best result is defined as the highest sensitivity. The list of the chosen hyperparameters per estimator is displayed in Table 18.

Table 18 - Hyperparameters tested during the fine-tuning phase.

| | | |
|---|---|---|
| **Decision tree** | Criterion | Gini index, Shannon Entropy |
| | Max depth | [2, 3, 4, until leaf is pure] |
| **Random Forest** | Nb of estimators | [5, 25, 50, 75, 100] |
| | Criterion | Gini index, Shannon Entropy |
| | Max depth | [2, 3, 4, until leaf is pure] |
| **SVM** | Kernel | RBF, Polynomial 3, Sigmoid |
| | Regularization parameter | [1, 10, 100, 1000] |

For the Decision tree and the Random Forest, they are also other hyperparameter that we could have attempted to tune. Like the minimum of examples need to create a new node, or the minimum number of examples for a leaf. Though, those hyperparameters could not be tested as certain measurement set have very few examples, like for the single H model at H12 and H24 where only 2 and 3 examples are used for the training (cf. Table 13). Therefore, we used the default value for the minimum samples split (default = 2) and the default value for the minimum samples for leaf (default = 1).

It is important to note that the patient P33 to P37 were not included in the test set for this experiment. This allows to keep the validation set untouched for the optimization. Thus, avoiding to overfit our model to our whole dataset.

## 6.2. Fine-tuning: Prediction results

Now that the method to improve the prediction results by tuning the hyperparameters has been explained, we can observe the improvement from this optimization. But first, to illustrate what kind of variation the different hyperparameter induces, we look at Figure 43 to Figure 45, which shows every prediction result for the most promising model, Global H.

For the Decision tree, the entropy criterion unable to have the best sensitivity in contrast to the gini criterion. Regarding the depth of the tree, pruning the tree by restraining the depth to 2, allows to maximize the sensitivity for this estimator. On the other hand, the specificity is higher from the depth 3, but stays relatively the same from that point. This behavior on the sensibility is in accordance with the findings from the SVM. Indeed, both estimators performed better by avoiding the outliers present in the data, and thus having a more generalized model.

For the Random Forest, the same observation realized with the Decision tree estimators can be made regarding the depth.

For the SVM estimator, the sigmoid kernel is clearly the best one to use regardless of the regularization value. Regarding the regularization parameter, the lowest value results in the highest sensibility. Meaning that, by integrating the outliers through a high value of C, we are decreasing the prediction accuracy. Thus, it seems that we are better off to have a more generalized model that excludes outliers. Which results in a low value for the regularization parameter.

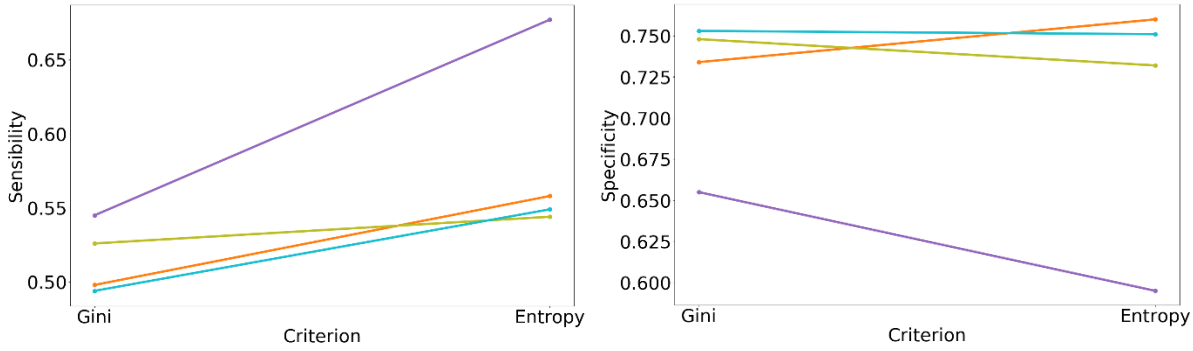


Figure 43 - Results during the fine-tuning phase for the Global H model – **Decision Tree** - On the right, the sensitivity; On the left, the specificity – In purple, depth=2; In orange, depth=3; In green, depth=4; In blue, depth= until pure.

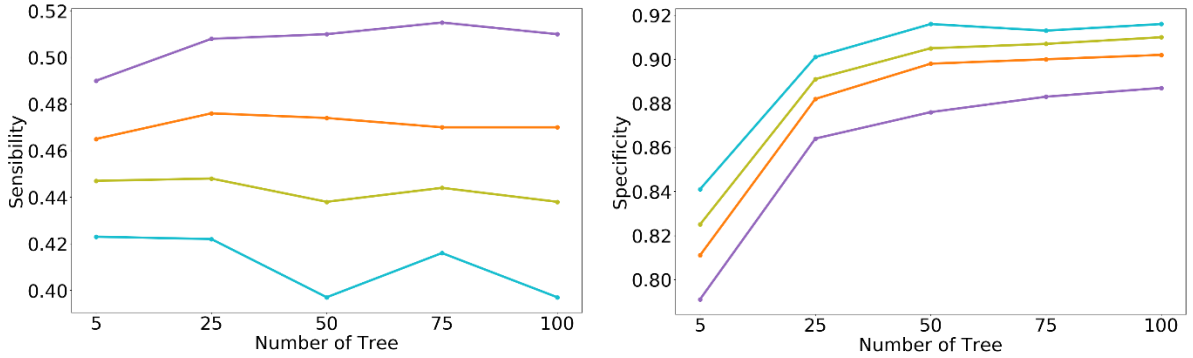


Figure 44 - Results during the fine-tuning phase for the Global H model – **Random Forest** - On the right, the sensitivity; On the left, the specificity – In purple, depth=2; In orange, depth=3; In green, depth=4; In blue, depth= until pure.

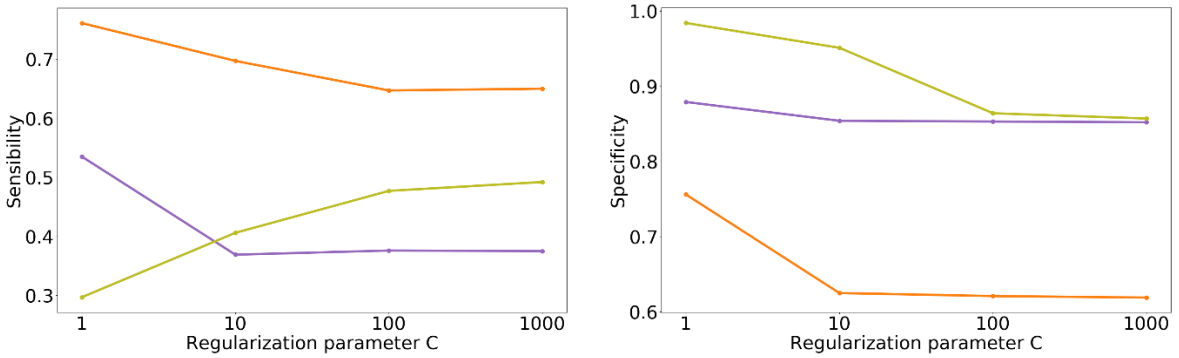


Figure 45 - Figure 44 - Results during the fine-tuning phase for the Global H model – SVM - On the right, the sensitivity; On the left, the specificity – In purple, RBF kernel; In orange, sigmoid kernel; In green, polynomial kernel.

The best hyperparameter for each model can be found in Table 19. As each different hypothesis have their database composed of different arrangement of EIT data, we can observe that most models resulted in different hyperparameter configuration.

Table 19 – Different Hyperparameters chosen through the fine-tuning of each model.

| | Decision Tree | | Random Forest | | | SVM | |
|---|---|---|---|---|---|---|---|
| | **Criterion** | **Depth** | **Criterion** | **Depth** | **Nb of estimators** | **Kernel** | **Regularization** |
| **Single H** | Gini | Until pure | Entropy | 2 | 5 | RBF | 1000 |
| **Variation H** | Entropy | Until pure | Entropy | 2 | 5 | Sigmoid | 1 |
| **Global H** | Entropy | 2 | Entropy | 2 | 75 | Sigmoid | 1 |

With the best hyperparameters that have been found, it is now possible to update the prediction results. The results can be seen in the Figure 46, Figure 47 and Figure 48 as well as the table including all results and variation (Cf Table 39, Table 40 and Table 41 in the Annex).

The overall percentage variation shows that we were able to improve the sensitivity for all three hypotheses, regardless of the inference model. The Global H model benefited the most from this optimization. With that optimization, it is clearly the best prediction for two main reasons:

1. The sensitivity increases over time, rendering the model more accurate the hours before the clinician's decision to re-intubate. This observation is even more true for the SVM estimators, as it never ceases to increase. While the Decision Tree and Random Forest shows a lack of accuracy for the last measurement set H24. The percentage of accurate prediction for each patient for the Decision Tree (Cf. Table 42), shows that the decrease is due to the patient P05. While for the Random Forest (Cf. Table 43), shows that decrease is due to P03 and P05.

2. The overall sensitivity regardless of the measurement set is superior for all estimator (Cf. Table 39, Table 40 and Table 41 in the Annex)

Using this model and between the different estimator, the SVM is the best one with an overall sensitivity of 0.76, then the Decision Tree with 0.68 and the Random Forest with 0.52.

From the figure below, we can also notice that this sensitivity optimization cost the degradation of the specificity for most models.

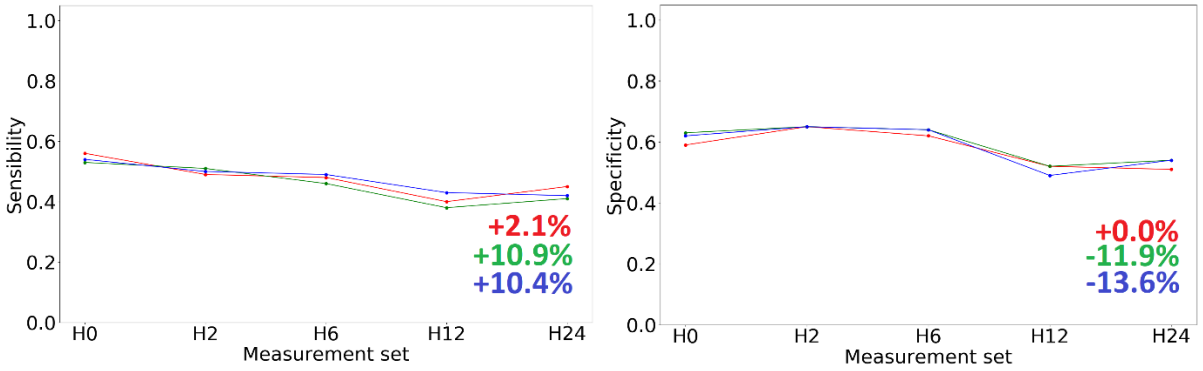


Figure 46 – Fine-tuning results - **Single H** - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM – The number represents the average percentage of variation between the fine-tuning results and the baseline.

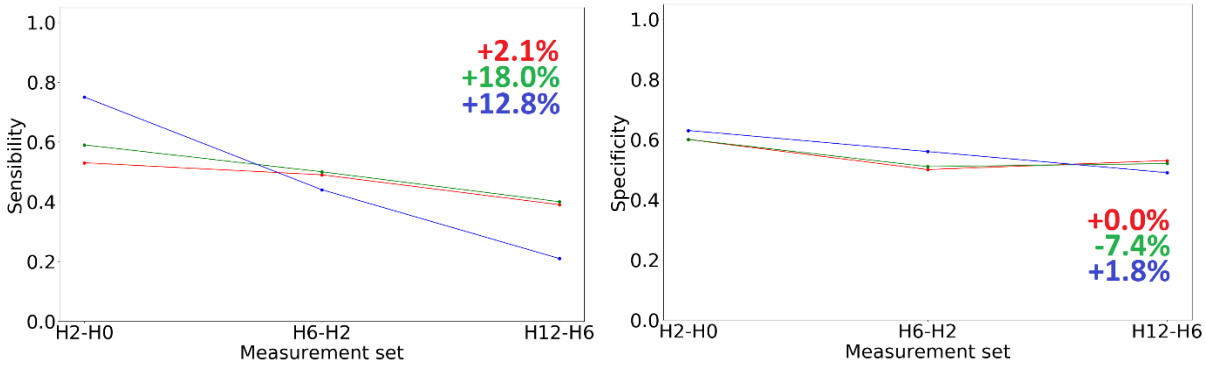


Figure 47 - Fine-tuning results - **Variation H** - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM – The number represents the average percentage of variation between the fine-tuning results and the baseline.

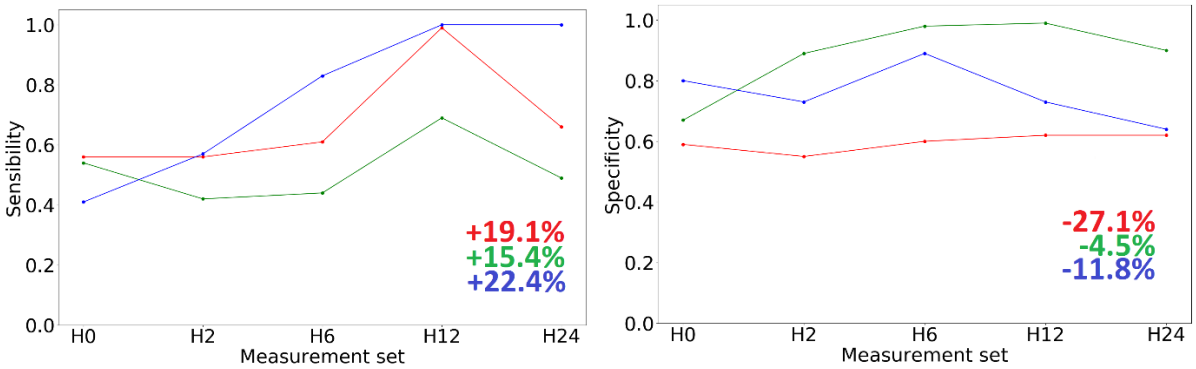


Figure 48 - Fine-tuning results - **Global H** - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM – The number represents the average percentage of variation between the fine-tuning results and the baseline.

For the depth of the Trees for both the Decision Tree and Random Forest, is close to 2 for H0 which contains few examples in comparison to the other set. Then, the depth reaches the limit impose by the hyperparameter. The best features are the same one seen from the Baseline (Cf. Table 35).

Table 20 – Fine-Tuning - Depth of each the Trees for the 100 trained estimators. The Random Forest results show the depth's average of the Tree present in the Forest.

| | | H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|---|---|
| **Decision Tree** | Min | 1 | 2 | 2 | 2 | 2 |
| | Max | 2 | 2 | 2 | 2 | 2 |
| **Random Forest** | Min | 1 | 2 | 2 | 2 | 2 |
| | Max | 2 | 2 | 2 | 2 | 2 |

We can also observe this optimization impact on the validation set. Fine-tuning results on the validation set

Table 44 to Table 46 in the Annex shows the results plus the variation between those results and the baseline results from the last chapter. The improvements on this set are mitigated as some results shows a decrease in sensitivity. Very surprising though, the Single H model with the SVM, performed quite well with a sensitivity equal to 0.66. But the specificity is only equal to 0.58. Using the same estimator on the Global H model, we have a sensitivity of 0.64 and a specificity of 0.73. For such a low difference of sensitivity (-4%) and significantly more specificity (+24%), we think the SVM is still the best model after this optimization.

## 6.3. Optimization of the EIT features

With the best hyperparameters chosen in the previous part, we focus on the EIT features to further improve the prediction. To this point, the totality of the EIT features is used in the learning model, thus 61 features. Though from those, they are certainly some of them that do not significantly contribute to the prediction. Indeed, not all features allow for a good distinction between the success and extubation failure. So, we test different mix of features from the 61, to find the one that maximize the sensitivity.

Ideally, the best method would have been to test every possible mix of features possible. Though, this method would have taken way too much time to compute with 61 features in total. Therefore, we went with other approaches which first scores the features and uses that to realize the different mix. We test3 different methods:

- Recursive feature elimination (RFE): this method trained the model with the initial set of features once. Then, it uses the features coefficient or features importance generated by the model during the training, to rank the features. The least important feature is then removed, and the process is repeated until we went through all features.

- Select from models (SFM): this method is also based on the coefficient applied to the features by the estimator. For each model training, we retrieve the feature coefficient. The one that are inferior to the median feature coefficient is set to 0. The coefficients are then saved for each feature and at the end we sum them all per features, to give us the overall feature ranking. Then like the RFE method, we iterate the training of the models by removing the features with the smallest overall coefficient.

- Correlation point-biserial (Pt biserial): features are ranked depending on their correlation score with the outcome. Then, we iterate the training by removing each time the least correlated feature. The usual correlation method like Pearson or Spearman cannot be applied in our case, as we are dealing with a continuous versus binary variable. Therefore, we use the Point-biserial correlation that is designed for such situation. The equation to calculate it can be seen below:

$$r_{pb} = \frac{mean(features(sucess)) - mean(features(failure))}{std(features)} \sqrt{\frac{nb_{success} * nb_{failure}}{nb_{total}}} \quad \text{Equation 27}$$

For all 3 methods, they find the best features out of all different measurement set (H0, H2, …) in order to have a more generalized ranking. The RFE and SFM can only be applied to the Decision tree and Random forest estimators, as the SVM does not score the features during the training. Otherwise, the 3 methods basically all based their approach on choosing the best features depending on their rank. The difference between them rests in their method to rank to features.

Like for the fine-tuning step, the patient P33 to the patient P37 were not included in the test set to avoid overfitting our model to our whole database.

## 6.4. Features optimization: Prediction results

### 6.4.1. Prediction results for patients failing the extubation

The sensibility results for each optimization methods can be seen in the Table 21. From the 3 methods, the point-biserial correlation resulted in the best sensibility improvement for the 3

estimators. The Figure 49 shows the variation of the sensitivity and the specificity for each estimator. We can observe that the Decision tree and Random forest peak rapidly, with 7 and 4 EIT features, respectively used by the 2 estimators. Then, the sensitivity decreases as they are more EIT features. The SVM, on the other hand, has a slow increase and peak at 42 EIT features.

Table 21 - Best number of features and sensibility depending on the optimization methods.

| Global | Decision Tree | | Random Forest | | SVM | |
|---|---|---|---|---|---|---|
| | Nb features | Max sensi | Nb features | Max sensi | Nb features | Max sensi |
| **RFE** | 3 | 0.78 | 5 | 0.60 | / | / |
| **SFM** | 4 | 0.84 | 4 | 0.62 | / | / |
| **Pt-Biserial** | 7 | 0.85 | 4 | 0.75 | 42 | 0.80 |

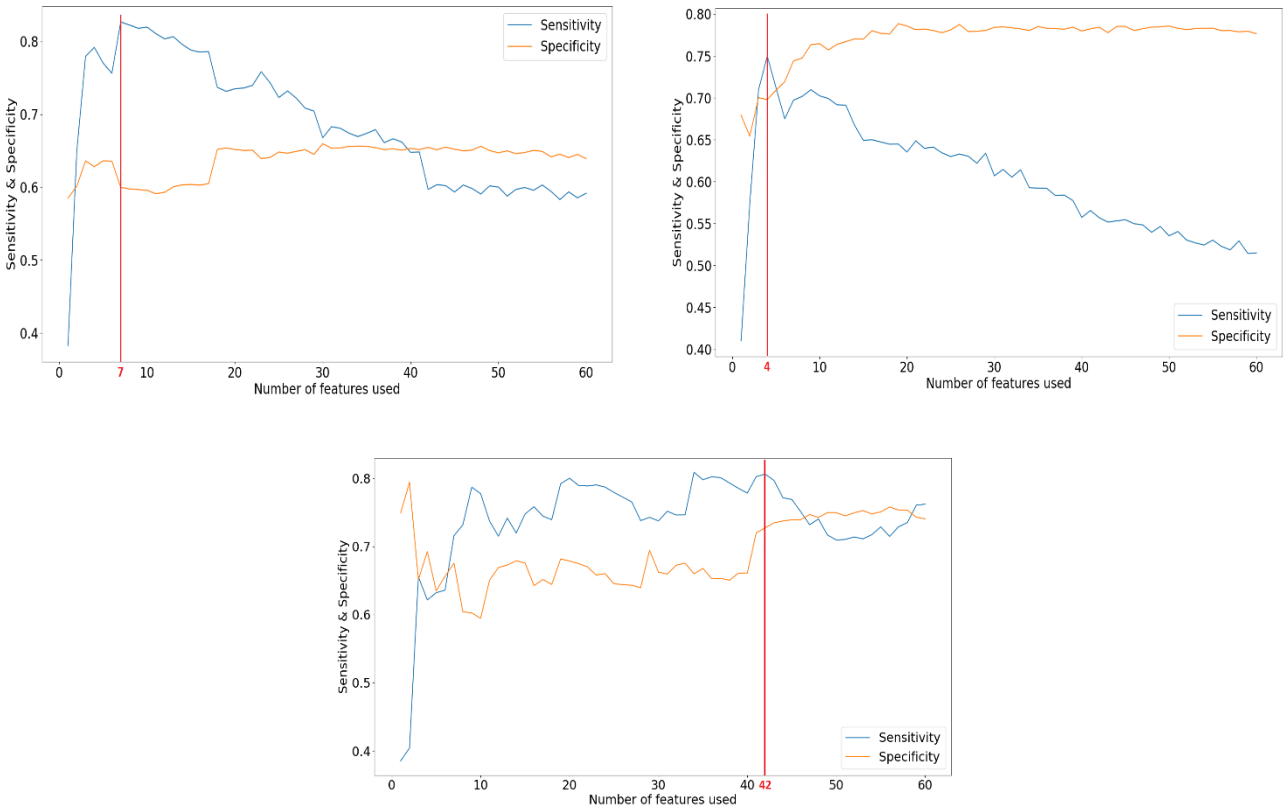


Figure 49 - Variation of the sensitivity and specificity with the point-biserial optimization method. On the top left, the Decision tree, on the top right, the Random forest and at the bottom, the SVM.

With the optimal EIT features for each estimator, it is now possible to have the final prediction results of our best model (cf. Table 47). The Decision tree and Random forest estimators significantly benefited from this optimization, with a 19.1% overall improvement for the first estimator and 28.3% for the second one. The improvements on the SVM have mainly impacted the prediction results for H2.

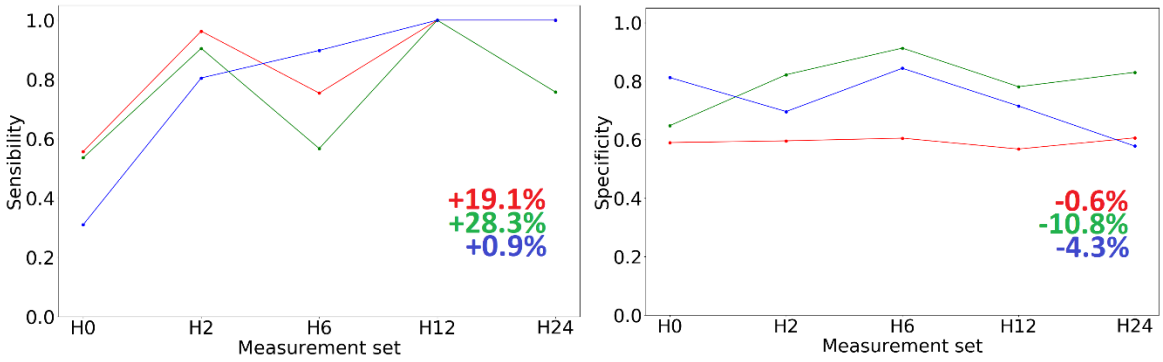


Figure 50 – Features optimization results - Global H - On the right, the sensitivity; On the left, the specificity – In red, Decision Tree; in green, Random Forest, in blue, SVM – The number represents the average percentage of variation between the features optimization results and the fine-tuning results.

After being the best estimators so far, with those new results the SVM is no longer the best model in terms of pure sensitivity. Indeed, the Decision tree now has the highest score. While the Random Forest, has still the worst prediction from the 3 estimators. Though, we would argue that the SVM is still the best one for this application. To plead argue our case, we turn to the Table 22, which shows the average percentage of accurate prediction over the whole results cross-validation process. From it, we can observe from the patients that were not predicted 100% accurately that:

- Decision tree:
  - the prediction has a lower accuracy for the prediction of P02 before the failure which occurred 15 hours after the extubation for this patient;
  - for P03, the prediction is completely wrong at H6. This could have falsely informed the clinicians, resulting in the patient not receiving the proper treatment to avoid a failure;
  - the model is able to better predict the patient at H0 than for the SVM;
- SVM:
  - the prediction for P02 lacks accuracy for the measure H2. Though, it was 100% accurate the at the next occurrence, before the extubation;
  - the same observation can be made for P12, which had an accurate prediction hours before the failure. Plus, the low accuracy of the previous measures can be linked to the NIV treatment, which temporarily improve the ventilation;
  - measures that had prior NIV treatment (in yellow boxes) are more badly influence with the SVM. Indeed, the accuracy of those measures are lower than their respective counterpart using the Decision tree.

With those observations, the SVM seems more reliable in his prediction of the failure patient. Indeed, for it was always accurate hours before the actual ventilation distress inducing the re-

intubation. In addition to that, the SVM estimators have a better specificity, with an improvement of 22.9% from the Decision tree.

Table 22 - Percentage of accurate prediction for each patient. The grey boxes are takes that were not realized or not interpretable. The purple boxes are when measured that were not realized as the patient was re-intubated. The yellow cases marked when a patient received a NIV treatment.

| | | Decision tree | | | | | SVM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | H0 | H2 | H6 | H12 | H24 | H0 | H2 | H6 | H12 | H24 |
| Failure | P02 | 24 | 100 | 76 | | | 4 | 58 | 100 | | |
| | P03 | 10 | 89 | 0 | 100 | 100 | 4 | 97 | 95 | 100 | 100 |
| | P05 | 62 | 100 | 100 | | 100 | 89 | 93 | 100 | | 100 |
| | P06 | 68 | 88 | 100 | 100 | | 16 | 70 | 54 | 100 | |
| | P09 | 91 | | | | | 81 | | | | |
| | P14 | 38 | 91 | 100 | | | 5 | 64 | 100 | | |
| | P27 | 23 | | | | | 0 | | | | |
| | P29 | 86 | | | | | 100 | | | | |
| | P36 | 66 | 100 | | | 100 | 21 | 89 | | | 100 |
| Success | P01 | 69 | 89 | 100 | | | 78 | 100 | 100 | X | X |
| | P04 | 79 | 95 | 100 | | 3 | 99 | 24 | 94 | X | 0 |
| | P07 | 58 | 3 | 0 | | 100 | 94 | 60 | 91 | X | 0 |
| | P08 | 65 | 100 | 100 | 100 | | 99 | 100 | 100 | 100 | X |
| | P10 | 48 | 84 | | | | 55 | 27 | X | X | X |
| | P11 | 90 | 91 | 31 | 100 | 100 | 91 | 21 | 100 | 0 | 0 |
| | P12 | 82 | 100 | 90 | 100 | 100 | 99 | 97 | 100 | 54 | 100 |
| | P15 | 42 | 94 | 100 | | | 68 | 99 | 70 | X | X |
| | P16 | 49 | 84 | 0 | | 3 | 33 | 100 | 51 | X | 100 |
| | P17 | 47 | 84 | 100 | 0 | 3 | 83 | 100 | 100 | 100 | 45 |
| | P18 | 43 | 15 | 100 | | 100 | 83 | 98 | 100 | X | 100 |
| | P19 | 39 | 16 | 100 | 100 | 100 | 91 | 48 | 91 | 100 | 100 |
| | P20 | 51 | 79 | 9 | 0 | 3 | 80 | 19 | 0 | 72 | 0 |
| | P21 | 90 | 84 | 0 | 0 | 3 | 87 | 100 | 94 | 76 | 100 |
| | P22 | 30 | 4 | 100 | | 3 | 60 | 0 | 100 | X | 100 |
| | P23 | 79 | 95 | 9 | 100 | 100 | 98 | 100 | 100 | 100 | 100 |
| | P24 | 83 | | | 100 | 100 | 97 | | | 4 | 100 |
| | P25 | 26 | 88 | 10 | | 100 | 49 | 100 | 95 | | 100 |
| | P26 | 93 | 14 | 100 | 100 | 100 | 99 | 100 | 84 | 72 | 0 |
| | P28 | 29 | 84 | 11 | 0 | | 83 | 100 | 63 | 92 | |
| | P30 | 80 | 87 | 91 | 0 | | 87 | 75 | 100 | 72 | |
| | P31 | 41 | 13 | 0 | | 100 | 60 | 35 | 100 | | 100 |
| | P32 | 94 | 100 | 100 | 100 | | 96 | 90 | 100 | 100 | |
| | P33 | 40 | 4 | 100 | 0 | 73 | 87 | 43 | 100 | 71 | 0 |
| | P34 | 70 | 15 | 0 | | 3 | 69 | 96 | 0 | | 0 |
| | P37 | 48 | 4 | 100 | | | 90 | 14 | 100 | | |

We can also observe the impact of the features optimization on the validation set (cf. Features optimization results on the validation set

Table 48 in the Annex). The sensitivity results have been greatly improved for all three estimators. Especially for the measurement take H0 and H2. Though, this cost a significant decrease for the specificity. Indeed, we can observe with the SVM, that for H12 and H24, the success patients were always misclassified. From those sensitivity results, the SVM is the less performing one. The sensitivity with the Decision Tree is 11.7% better, as well as 10.2% for the Random Forest. On the other hand, the prediction is accurate in the measurement set preceding the re-intubation. It is interesting to observe that the SVM is the best worst prediction model for the validation set.

Those sensitivity results on the validation set are based on only one failure patient (P36). The results on the testing set include 5 failure patients at H2 and H6, and 2 for H12 and H24. As the prediction is more stable from the SVM than the Decision Tree (Cf. Figure 50). We still think than the SVM is better. As changing abruptly, the prediction by predicting a success (P03H6) could really endanger the patient outcome.

### 6.4.1. Prediction results for patients succeeding the extubation

In this part, we are interested to observe the model's performance regarding the prediction of the success group. To do that, we have defined patients succeeding the extubation as 1, and the one failing as 0. The Table 23 shows the results for the testing set, and in the Annex, the Table 49 shows the results for the validation set. From those results, we can observe that the SVM is still the best model for our case. Indeed, it has the highest sensitivity, thus resulting in the less false detection of success patients than with the other estimators.

Table 23 – Results for the success group - Feature optimization - Global model prediction.

| P01 -> P32 | Decision Tree | | Random Forest | | SVM | |
|---|---|---|---|---|---|---|
| | sensitivity | specificity | sensitivity | specificity | sensitivity | specificity |
| H0 | 0.55 | 0.61 | 0.61 | 0.59 | 0.81 | 0.32 |
| H2 | 0.63 | 0.92 | 0.74 | 0.93 | 0.69 | 0.79 |
| H6 | 0.63 | 0.62 | 0.78 | 0.53 | 0.85 | 0.88 |
| H12 | 0.57 | 1.00 | 0.65 | 1.00 | 0.77 | 1.00 |
| H24 | 0.60 | 1.00 | 0.70 | 0.74 | 0.67 | 1.00 |
| Mean | 0.60 | 0.83 | 0.70 | 0.76 | 0.76 | 0.80 |

### 6.4.2. Kappa results

For this evaluation, we wanted to compare two perspectives:

1. How does the clinician's assessment of the patient state compared to the extubation outcome?
2. How does our best prediction model compared to the extubation outcome?

In other terms, for the first perspectives, we want to compare the first option to the second one describes in the section 5.1. So, we are comparing the clinicians view that declare an extubation failure only on the measurement set preceding the failure, to the actual extubation outcome, where patients are defined as failure from the start.

For the second perspectives, we are comparing the prediction realized by our best prediction model (Global H using SVM), to the actual extubation outcome.

To realize the comparison of the different model, we compute the Cohen kappa coefficient for both perspectives, to then compare both methods of assessment (clinician versus our model). The results are shown in Table 24. From the clinician estimation, their strong assessments are for the success patients from which they are never wrong. Although their undoing is toward the assessments of the patients failing the extubation. Most clinicians would not notice the signs of ventilation distress before it would be too late, and thus had to re-intubate the patient. Although, experienced clinicians could have better insight to detect those signs before the actual failure. This explains why the Cohen kappa results are so low for the clinician perspective.

On the other hand, our model has proved effective in detecting the ventilation distress well ahead of the actual intervention to re-intubate the patient. Which shows in the high correlation of Cohen kappa results between our prediction and the actual extubation outcome. The low results for H24, rest in the bad detection of patients succeeding the extubation.

Table 24 – Comparison of the Cohen kappa results between the clinician and SVM-Global model estimation

| | H0 | H2 | H6 | H12 | H24 |
|---|---|---|---|---|---|
| Clinician | 0.00 | 0.43 | 0.00 | 0.25 | 0.52 |
| Our best model | 0.92 | 0.90 | 0.79 | 0.76 | 0.31 |

## 6.5. Exploring the prediction model

With the best model now selected, we want to better understand how it works. To do this, this part analyzes the EIT features that were selected to yield the highest sensitivity. Mainly the SVM estimators will be discussed here.

First, to give us a view of the best EIT features, the Table 25 display the 10 best EIT features depending on their Point-biserial correlation score. From it, 2 interesting information can be gathered:

- half of the features are one that are already present in the scientific literature and the other are the one added for this study. This demonstrates the effectiveness of our added EIT features, as they are among the one with high correlation with the extubation outcome;
- 40% of the features are their coefficient of variation counterparts, with the 2 best features among them. This enlightens the important to not only monitor the current ventilation state, but also the variation throughout the measurement.

Table 25 - 10 best EIT features depending on their Point-biserial correlation. The blue boxes are the features that were created for this study.

| Num | EIT features | H0 | H2 | H6 | H12 | H24 | Sum |
|---|---|---|---|---|---|---|---|
| 1 | CV_Allimage_Zone4 | 0.11 | 0.59 | 0.53 | 0.12 | 0.67 | 2.02 |
| 2 | CV_LungShape_left_compasity | 0.36 | 0.28 | 0.43 | 0.69 | 0.23 | 2.00 |
| 3 | Flow_expi | 0.30 | 0.52 | 0.41 | 0.20 | 0.52 | 1.95 |
| 4 | CV_LungShape_left_y | 0.37 | 0.21 | 0.44 | 0.72 | 0.16 | 1.92 |
| 5 | Allimage_Zone1 | 0.23 | 0.21 | 0.54 | 0.18 | 0.62 | 1.77 |
| 6 | CoV_RL_All | 0.25 | 0.37 | 0.49 | 0.17 | 0.46 | 1.75 |
| 7 | Frequency | 0.43 | 0.09 | 0.41 | 0.33 | 0.42 | 1.68 |
| 8 | Flow_inspi | 0.16 | 0.44 | 0.41 | 0.13 | 0.44 | 1.59 |
| 9 | Tidal_Zone1 | 0.08 | 0.15 | 0.54 | 0.28 | 0.51 | 1.56 |
| 10 | CV_LungArea_left | 0.34 | 0.36 | 0.11 | 0.52 | 0.21 | 1.54 |

The Figure 51 and Figure 52 shows the difference in the different EIT features present in the whole database, and the features present after the optimization of the SVM estimator (with 42 features). With all that information so far, we can draw a better picture of how the prediction models work:

- The prediction is predominantly based on the results of the distribution EIT features. Thus, reflecting the appearance of the lung through the EIT perspective. Indeed, they represent 83.3% of the 42 EIT features.
- The CV features plays an important role in the prediction. Indeed, they represent 45.9% of the features from the whole database, and 45.2% on the optimize database for the SVM.
- Though low in number, the EIT measures that are correlated to the frequency performed well. The $flow_{EIT}$ measures during the expiration is rank as the overall top 3 features, while the same metric for the inspiration is rank 8.
- After the optimization, there is only one volume-related feature which is the GI.

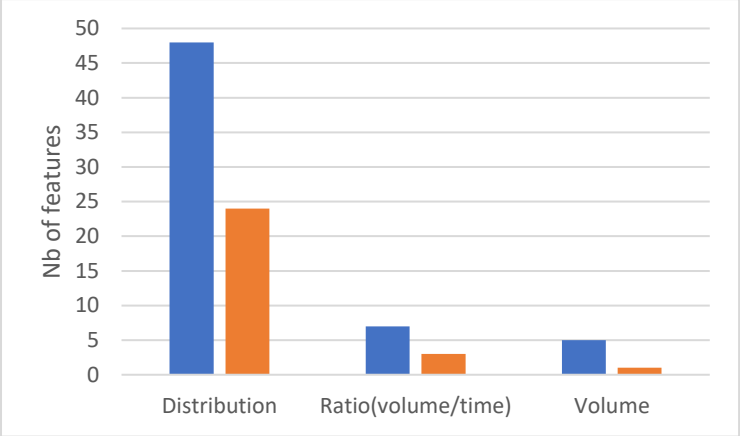


Figure 51 - Different sorts of features among all 61 of them. In blue, the original features and in orange, their coefficient variation counterpart.

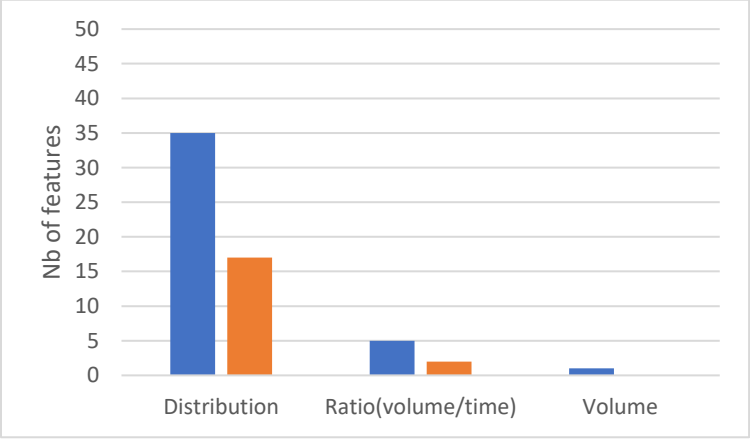


Figure 52 - Different sorts of features after the point-biserial feature optimization (42 features). In blue, the original features and in orange, their coefficient variation counterpart.

Seeing as the interpretation of the distribution from the EIT is one of the main parameters for the prediction, we can compare the prediction accuracy of different patients to their EIT images to test this assessment. We focus on the worst prediction for the success classes first (cf. Table

22). There is a total of 21 predictions that are below 50%, all measurement set included. By comparing the EIT images from those measures with ones that were correctly predicted, we identified that the predictor seems to recognize healthy lungs from the EIT images. Consequently, any EIT images that do not look like lungs are classified as extubation failure. To illustrate our argument, we show the Figure 53. Accurate prediction has EIT image that shows a good distribution of the ventilation on both lungs. While for the patient P20 and P34, has odd EIT image, with a ventilation that is more localized in certain area. The odd EIT images could come from various reasons: (1) the patient could have an acute ventilation distress which results in ventilation patterns that does not look like lungs (2) measurement noise from a bad electrodes placement or patient movement could generate artifacts during the image reconstruction.

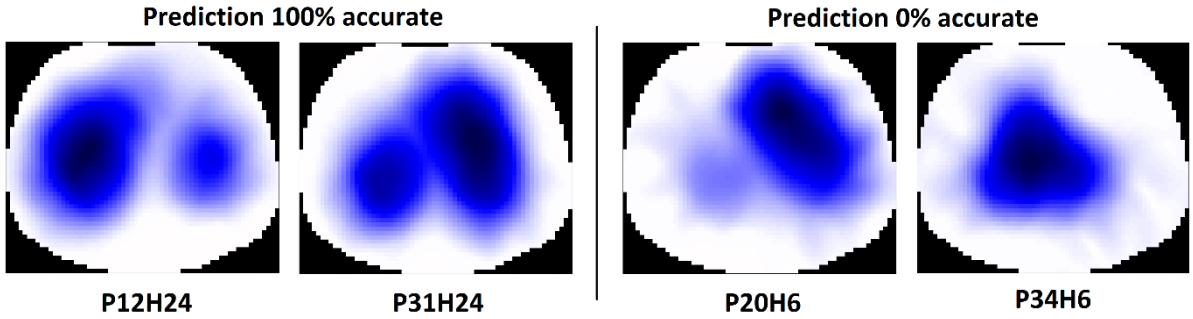


Figure 53 - Different tidal images for different patients that succeeded the extubation. On the left, patient that is rightfully classified; on the right, patients that were misclassified.

However, this theory does not cover for all misclassifications. Indeed, from the 21 incorrect predictions on the success class, there are 8 patients that are still misclassified while having normal EIT images (cf. Figure 54). Our first theory is mostly based on the numerous original distribution features, that allows to give an estimation of the lungs shape to the predictor. However, this leaves behind the all the distribution features with the coefficient of variation (CV). Plus, seeing as the $Flow_{EIT}$ has a high correlation with the outcome, they certainly play a role in the prediction result.

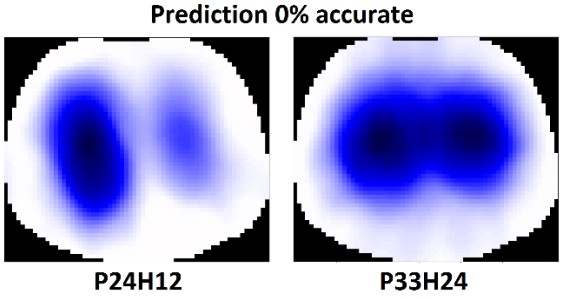


Figure 54 - Success patients that were badly predicted but have good ventilation distribution.

## 6.6. Gain from the additional EIT features.

In this part, we are interested in measuring the influence of the added EIT features. First, we can review the balance between the two sorts of features (cf. Figure 55). We now have a bit more of the new features in the database, as the balance changed with the optimization.

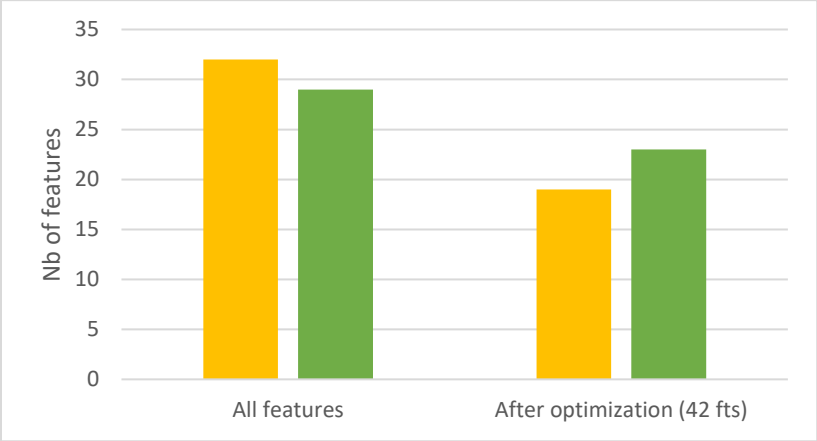


Figure 55 - Distribution of the usual EIT features finds in the science community (orange) versus the one that are added for this study (green).

Then, we want to really compare the performance of the new features, we create a new prediction model only based on the usual EIT features. We opted to only do the Global H model, as it was the best clear model. Except for that, the model follows the same optimization, with the fine-tuning and the features optimization. This process is realized for the 3 estimators. The result for each estimator can be seen in the Table 26. The decision tree uses 9 features, the Random Forest 5 features and the SVM, 23. Regardless of the estimator, the prediction is worst in all regards. TheTable 27, sum up the variation between this model and our best.

Table 26 - Prediction results for the Usual features prediction model after it was optimized.

| | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.56 | 0.53 | 0.48 | 0.55 | 0.63 | 0.58 |
| H2 | 0.44 | 0.65 | 0.56 | 0.79 | 0.40 | 0.59 |
| H6 | 0.72 | 0.69 | 0.57 | 0.81 | 0.62 | 0.63 |
| H12 | 0.93 | 0.60 | 0.64 | 0.84 | 0.96 | 0.63 |
| H24 | 0.88 | 0.61 | 0.77 | 0.89 | 0.96 | 0.58 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Mean | 0.71 | 0.62 | 0.60 | 0.78 | 0.72 | 0.60 |

Table 27 – Comparison of the different model using the 2 sets of EIT features.

| | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensi | Speci | Sensi | Speci | Sensi | Speci |
| Usual EIT features | 0.71 | 0.62 | 0.60 | 0.78 | 0.72 | 0.60 |
| With additional EIT features | 0.85 | 0.59 | 0.75 | 0.80 | 0.80 | 0.73 |
| **Variation [%]** | **17.5** | **-3.9** | **19.9** | **3.0** | **10.8** | **17.9** |

## 6.7. Comparison with the study from Longhini.

We wanted to compare our method with the one already proposed by Longhini et al [68], in order to compare both EIT methods. As mentioned in the state of the art, they based their prediction on only one EIT feature. After the extubation, the feature that yielded the best sensitivity is the GI. Therefore, we decided to use their same method and created a prediction model only based on the GI feature.

To find the best threshold for the EIT feature, their team used the Youden index [82]. This index is used to evaluate the performance of a diagnostic marker. In addition to the Youden index results, it is possible to calculate the optimal threshold.

By calculating the Youden index, like it is proposed by their team, we find the best GI threshold to uses is 0.62. Patients that have a GI superior to that value are classified as extubation failure. The result for this prediction model on our database is shown in the Table 28. Interestingly, the GI features did also perform well at the beginning of the extubation. Though, the performance then declines throughout the weaning observation.

This feature performed surprisingly well on his own, but overall offers less prediction performance.

Table 28 - Comparison between Longhini's method on our dataset and the SVM-Global H model.

| | **GI** | | **SVM** | |
|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.78 | 0.58 | 0.31 | 0.81 |
| H2 | 0.83 | 0.64 | 0.81 | 0.70 |
| H6 | 0.80 | 0.96 | 0.90 | 0.84 |
| H12 | 0.50 | 1.00 | 1.00 | 0.72 |
| H24 | 0.67 | 0.94 | 1.00 | 0.58 |
| Mean | 0.72 | 0.82 | 0.80 | 0.73 |

## 6.8. Overall results comparison with the other prediction methods.

To finish reviewing our proposed method, we compare our results with the best prediction model that were shown in the scientific literature. The Table 29 shows the 3 best models, with our best results. In this case, a comparison based on our dataset in not possible, because the dataset and the measures are different. Though still, it gives us an idea of what can be done. From it, we can observe that we were not able to outperform the prediction model already establish. Indeed, it seems that the RSBI is still the best prediction model, while the blood gas results, and diaphragm thickness are close second.

Table 29 - Comparison between our best prediction model versus the best model from the scientific literature.

| Authors | Measures | Sensitivity | Specificity |
|---|---|---|---|
| L.N. Segal et al. [37] | Rapid Shallow Breathing Index | 0.89 | 0.89 |
| Saugel et al. [24] | Corrected serum anion gap | 0.88 | 0.71 |
| E. DiNino [35] | Diaphragm thickness | 0.88 | 0.71 |
| Our best model (Global H - SVM) | EIT | 0.8 | 0.73 |

## 6.9. Conclusion

This whole chapter presented all of the different prediction results, through the different improvements that were realized. For the first optimization, we realized a grid-search to find the best mix of hyperparameter for each different prediction models. From the baseline results, the Global H hypothesis had seen the most improvement from the three. The best inference model was still the SVM on the testing and validation set.

After which, we took the best hypothesis models to further improve it by keeping only the features that yields the best sensitivity. From this optimization, we have seen that the Decision Tree has the best sensitivity, both in the testing and validation set. But has a low specificity. While the SVM is now second best. In the end, we chose the SVM as the best inference model, despite not having the best sensitivity. The strength of this inference model relies on his prediction stability for patients failing the extubation. Meaning that the prediction stays accurate and does not change once a patient is predicted as a failure. In contrast to the Decision Tree, which changes for a patient. For the SVM model, we then used 42 features, mostly composed of distribution features. We also saw that the introduce $Flow_{EIT}$ features played a significant role in the prediction has they have a high correlation with the outcome.

At last, we compare ourselves with the other study that used EIT data to predict extubation failure. Longhini and his team found that the best EIT feature is the GI. Apply to our dataset, this feature outperforms at H0, and thus before the extubation. The sensitivity of the GI then declines with the other measurement set, while with the SVM improves. This shows that more

EIT features are needed when we went to observe and prediction patients' outcomes after the extubation.

# 7. Conclusion and future development

The EIT has since many improvements from the last decade, which have been possible through the development of open EIT toolkit like EIDORS, and the availability of new commercial EIT devices. In ICU, most researchers have focused their effort to develop new EIT features to better assist the clinicians during the ventilator treatment. In our approach, we wanted to assert whether the EIT could find a new application, in monitoring the patients after the extubation. We setup a clinical study to pursue this goal, in which we were able to recruit 37 patients in total, but 2 had EIT data that were too noisy. The goal was to recruit 50 patients, but the COVID crisis set us back a bit. Still, we make do with our database, which contains 25.7% of patients failing the extubation.

During this PhD, we developed our own EIT framework to pre-process, reconstruct and extract the EIT features. From our observation of the first patients and our discussion with clinicians, we decided to introduce 4 new EIT features. With 2 that are EIT copy of already existing metrics used in ICU. All those EIT features allowed to build a comprehensive model that attempt to portray the full picture of the ventilation. Through the 3 models tested, the Global H is the best clear one. Though between the estimators, the Decision tree and SVM both products good sensitivity results. But the SVM seems best as it is more reliable overall. The addition of the 4 features, greatly benefited the prediction. Indeed with it, we are able to reach the same range of results as the best model already proposed in the literature (cf. Table 1). Though, even with the different optimization that we put our model through, we did not outperform the already existing model. From the Table 29, we see that the best model is still from the RSBI, then in second the blood analysis and the diaphragm thickness that have the same prediction results. However, with this thesis, we were still able to prove the effectiveness of the EIT to predict the extubation failure. Plus, we proposed a new approach that focus more on giving clinicians better tools to assist them during the weaning phase.

Looking back at the thesis, there are several aspects that we could have done differently. For starters, we decided to separate the weaning phase of 48 hours into 7 measurements set. However, this is not taking full advantage of the EIT that can be used continuously. With such approach, it would have been possible to construct “real-time” model, which would signal any ventilation distress to the clinicians. But such approach also brings him lots of problems. The main one being that the contact impedance variation or electrode belt movement would certainly have added significant artifacts. Rendering the analysis complicated. Another possible improvement would have been to add the ICU features (blood analysis, lung ultrasound results,

…) to the prediction model. This approach was considered, but unlike the EIT data that are present in every measurement set, not all ICU features were measured each time. Thus, with thus approach, every prediction model at each $H_i$ would have been unique, which complicated the comparison. At least, we would have liked to test more complex prediction models, like the Long Short Term Memory neural network. This model takes as input vector instead of 1D value for each feature. This allows to add the notion of time to the model, which is a critic notion for our applications.

# 8. Publications

During this thesis, the following work have been published:

- V. Janiak, V. Joussellin, G. Tallec, C. Marsala, U. Saleem, M. Dres and A. Pinna., "Prediction of the extubation outcome through Electrical Impedance Tomography measurements," IEEE *Biocas 2022*.

- V. Bonny, V. Janiak, S. Spadaro, A. Pinna, A. Demoule, and M. Dres, "Effect of PEEP decremental on respiratory mechanics, gas exchange, pulmonary regional ventilation and hemodynamics in patients with SARS-Cov-2 associated Acute Respiratory Distress Syndrome," *Crit. Care*, vol. 24, no. 1, Dec. 2020

Workshop:

- PhD Workshop on Artificial Intelligence – Sorbonne Université - SCAI – Octobre 2022

# 9. Annex

## 9.1. Clinical study EXIT – Weaning outcome

Table 30 – All inclusions date for the EXIT patients. In green, the success; in orange, the failures.

| Patients | Inclusions date |
|---|---|
| P01 | 28/02/2020 |
| P02 | 04/03/2020 |
| P03 | 17/06/2020 |
| P04 | 25/06/2020 |
| P05 | 17/08/2020 |
| P06 | 01/09/2020 |
| P07 | 22/09/2020 |
| P08 | 17/02/2021 |
| P09 | 03/03/2021 |
| P10 | 10/03/2021 |
| P11 | 18/03/2021 |
| P12 | 01/04/2021 |
| P13 | 02/04/2021 |
| P14 | 19/04/2021 |
| P15 | 21/04/2021 |
| P16 | 10/05/2021 |
| P17 | 11/05/2021 |
| P18 | 12/05/2021 |
| P19 | 08/06/2021 |

| Patients | Inclusions date |
|---|---|
| P20 | 25/06/2021 |
| P21 | 30/06/2021 |
| P22 | 05/08/2021 |
| P23 | 12/08/2021 |
| P24 | 16/08/2021 |
| P25 | 19/08/2021 |
| P26 | 31/08/2021 |
| P27 | 01/09/2021 |
| P28 | 08/09/2021 |
| P29 | 15/09/2021 |
| P30 | 16/09/2021 |
| P31 | 28/09/2021 |
| P32 | 11/10/2021 |
| P33 | 18/11/2021 |
| P34 | 06/12/2021 |
| P35 | 24/01/2022 |
| P36 | 23/02/2022 |
| P37 | 17/03/2022 |

## 9.2. EIT features extracted from the images

Table 31 - Count of tidal and frames used after filtering the signal.

| | Count of EIT tidal images | | | | | | |
|---|---|---|---|---|---|---|---|
| | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
| P02 | 86 | 114 | 115 | | | | |
| P03 | 59 | 65 | 62 | 90 | 87 | | |
| P05 | 62 | 53 | 84 | | 76 | | |
| P06 | 118 | 85 | 74 | 90 | | | |
| P09 | 73 | 68 | 39 | | | | |
| P14 | 77 | | | | | | |
| P27 | 102 | | | | | | |
| P29 | 66 | 74 | | | 89 | | |
| P36 | 45 | 93 | 53 | | | | |
| P01 | 47 | 100 | 62 | | 79 | | |
| P04 | 119 | 100 | 75 | | 71 | | 82 |
| P07 | 39 | 52 | 45 | 63 | | | |
| P08 | 156 | | | | | | |
| P10 | 60 | 92 | | | | | |
| P11 | 56 | 60 | 76 | 52 | 56 | | 75 |
| P12 | 32 | 54 | 90 | 41 | 52 | | |
| P15 | 99 | 51 | 77 | | | | |
| P16 | 58 | 68 | 74 | | 61 | 72 | |
| P17 | 64 | 70 | 71 | 79 | 87 | 96 | 81 |
| P18 | 59 | 48 | 63 | | 53 | | 76 |
| P19 | 53 | 65 | 73 | 75 | 61 | | |
| P20 | 83 | 89 | 70 | 77 | 92 | | 89 |
| P21 | 59 | 77 | 133 | 88 | 81 | | 65 |
| P22 | 71 | 112 | 103 | | 68 | | |
| P23 | 62 | 74 | 73 | 57 | 64 | 84 | 81 |
| P24 | 27 | | | 53 | 47 | 60 | |
| P25 | 60 | 62 | 80 | | 67 | | 80 |
| P26 | 50 | 70 | 43 | 59 | 49 | 75 | 86 |
| P28 | 108 | 100 | 93 | 91 | | | |
| P30 | 58 | 79 | 42 | 75 | | | |
| P31 | 104 | 67 | 95 | | 110 | 135 | 92 |
| P32 | 59 | 52 | 44 | 45 | | | 52 |
| P33 | 64 | 74 | 72 | 47 | 73 | | 43 |
| P34 | 66 | 53 | 80 | | 74 | | |
| P37 | 47 | 58 | 61 | | | | |

| | Count of EIT frames | | | | | | |
|---|---|---|---|---|---|---|---|
| | H0 | H2 | H6 | H12 | H24 | H36 | H48 |
| P02 | 3330 | 5524 | 4609 | | | | |
| P03 | 3188 | 4117 | 2640 | 3556 | 3489 | | |
| P05 | 4050 | 923 | 3578 | | 3327 | | |
| P06 | 6370 | 4665 | 3097 | 4886 | | | |
| P09 | 919 | | | | | | |
| P14 | 3399 | 2409 | 1367 | | | | |
| P27 | 4881 | | | | | | |
| P29 | 3945 | | | | | | |
| P36 | 2715 | 2390 | | | 2666 | | |
| P01 | 4459 | 5707 | 3470 | | | | |
| P04 | 2261 | 5009 | 2388 | | 3382 | | |
| P07 | 6622 | 4134 | 2729 | | 3046 | | 3596 |
| P08 | 2461 | 3048 | 1869 | 3814 | | | |
| P10 | 2001 | 3100 | | | | | |
| P11 | 2376 | 2539 | 3148 | 2698 | 1690 | | 2019 |
| P12 | 2648 | 3185 | 5562 | 2695 | 5432 | | |
| P15 | 3079 | 2651 | 3012 | | | | |
| P16 | 3992 | 2732 | 3420 | | 1953 | 3446 | |
| P17 | 2278 | 2772 | 1877 | 2333 | 3894 | 4345 | 3133 |
| P18 | 3556 | 3724 | 3678 | | 2112 | | 3432 |
| P19 | 1906 | 3813 | 5184 | 4444 | 3458 | | |
| P20 | 2194 | 2085 | 2459 | 2169 | 4224 | | 4146 |
| P21 | 2760 | 3690 | 6621 | 2685 | 3364 | | 2940 |
| P22 | 4356 | 6608 | 3245 | | 2372 | | |
| P23 | 3295 | 3441 | 1946 | 3556 | 2659 | 4860 | 3148 |
| P24 | 2575 | | | 3002 | 2134 | 3702 | |
| P25 | 1334 | 1622 | 4092 | | 3475 | | 5751 |
| P26 | 3459 | 2894 | 4516 | 2987 | 3108 | 4089 | 5751 |
| P28 | 5913 | 3984 | 3200 | 4441 | | | |
| P30 | 1783 | 3520 | 1267 | 3172 | | | |
| P31 | 3134 | 2861 | 4220 | | 4598 | 5070 | 3001 |
| P32 | 4216 | 4065 | 3527 | 2153 | | | 2485 |
| P33 | 3328 | 4577 | 4504 | 2224 | 3911 | | 2180 |
| P34 | 2624 | 2270 | 2572 | | 2393 | | |
| P37 | 2141 | 3488 | 4864 | | | | |

## 9.3. Framework to construct and test the prediction models.

### 9.3.1. Baseline results on the testing set

Table 32 - Baseline results - Single model prediction results.

| P01->P32 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.47 | 0.61 | 0.45 | 0.69 | 0.31 | 0.8 |
| H2 | 0.53 | 0.61 | 0.59 | 0.72 | 0.66 | 0.63 |
| H6 | 0.45 | 0.62 | 0.4 | 0.74 | 0.47 | 0.74 |
| H12 | 0.39 | 0.51 | 0.18 | 0.58 | 0.22 | 0.58 |
| H24 | 0.47 | 0.55 | 0.43 | 0.59 | 0.49 | 0.59 |
| **Mean** | **0.46** | **0.58** | **0.41** | **0.66** | **0.43** | **0.67** |

Table 33 - Baseline results - Variation model prediction results.

| P01->P32 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H2-H0 | 0.48 | 0.59 | 0.56 | 0.63 | 0.58 | 0.6 |
| H6-H2 | 0.49 | 0.5 | 0.44 | 0.53 | 0.45 | 0.56 |
| H12-H6 | 0.4 | 0.52 | 0.25 | 0.57 | 0.21 | 0.5 |
| **Mean** | **0.46** | **0.54** | **0.41** | **0.58** | **0.41** | **0.55** |

Table 34 - Baseline results - Global model prediction results.

| P01->P32 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.49 | 0.61 | 0.44 | 0.69 | 0.34 | 0.74 |
| H2 | 0.47 | 0.64 | 0.29 | 0.97 | 0.65 | 0.85 |
| H6 | 0.33 | 0.86 | 0.41 | 1 | 0.72 | 0.96 |
| H12 | 0.96 | 0.76 | 0.53 | 1 | 0.49 | 0.93 |
| H24 | 0.51 | 0.87 | 0.52 | 0.96 | 0.75 | 0.75 |
| **Mean** | **0.55** | **0.75** | **0.44** | **0.92** | **0.59** | **0.85** |

Table 35 - Most and least used EIT features for the Global H hypothesis.

| | | Decision Tree | Random Forest |
|---|---|---|---|
| **H0** | Most used features | CV_LungShape_right_compasity | CV_LungShape_right_compasity |
| | Least used Feature | $\Delta Z_{tidal}$ | $\Delta Z_{tidal}$ |
| **H2** | Most used features | CV_LungShape_right_compasity | CV_LungShape_right_compasity |
| | Least used Feature | $\Delta Z_{dynamic}$ | $\Delta Z_{tidal}$ |
| **H6** | Most used features | $Flow_{EIT\text{-}Expiration}$ | $Flow_{EIT\text{-}Expiration}$ |
| | Least used Feature | LungShape_right_compasity | $\Delta Z_{tidal}$ |
| **H12** | Most used features | $Flow_{EIT\text{-}Expiration}$ | $Flow_{EIT\text{-}Expiration}$ |
| | Least used Feature | $CV\_CoV\_FB_{dynamic}$ | $\Delta Z_{dynamic}$ |
| **H24** | Most used features | $Flow_{EIT\text{-}Expiration}$ | $Flow_{EIT\text{-}Expiration}$ |
| | Least used Feature | $\Delta Z_{tidal}$ | $\Delta Z_{dynamic}$ I |

### 9.3.2. Baseline results on the validation set

Table 36 - Baseline results on the validation set - Single model prediction results.

| P33->P37 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.30 | 0.71 | 0.22 | 0.80 | 0.27 | 0.75 |
| H2 | 0.60 | 0.42 | 0.80 | 0.43 | 0.60 | 0.31 |
| H6 | / | 0.54 | / | 0.60 | / | 0.72 |
| H12 | / | 0.55 | / | 0.65 | / | 0.67 |
| H24 | 0.56 | 0.54 | 0.62 | 0.60 | 0.71 | 0.65 |
| **Mean** | **0.49** | **0.55** | **0.55** | **0.62** | **0.53** | **0.62** |

Table 37 - Baseline results on the validation set - Variation model prediction results.

| P33->P37 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H2-H0 | 0.42 | 0.45 | 0.37 | 0.38 | 0.43 | 0.28 |
| H6-H2 | / | 0.48 | / | 0.51 | / | 0.54 |
| H12-H6 | / | 0.52 | / | 0.52 | / | 0.59 |
| **Mean** | **0.42** | **0.49** | **0.37** | **0.47** | **0.43** | **0.47** |

Table 38 - Baseline results on the validation set - Global model prediction results.

| P33->P37 | Decision tree | | Random forest | | SVM | |
|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| H0 | 0.34 | 0.69 | 0.23 | 0.79 | 0.29 | 0.74 |
| H2 | 0.49 | 0.54 | 0.16 | 0.90 | 0.66 | 0.62 |
| H6 | / | 0.81 | / | 1.00 | / | 1.00 |
| H12 | / | 0.01 | / | 1.00 | / | 1.00 |
| H24 | 0.84 | 0.84 | 0.99 | 0.99 | 1.00 | 0.10 |
| **Mean** | **0.56** | **0.58** | **0.46** | **0.94** | **0.65** | **0.69** |

## 9.4. Fine-tuning results

### 9.4.1. Fine-tuning results on the testing set

Table 39 – Fine-tuning results – Single model prediction results.

| P01->P32 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.56 | 19.1 | 0.59 | -3.3 | 0.53 | 17.8 | 0.63 | -8.7 | 0.54 | 74.2 | 0.62 | -22.5 |
| H2 | 0.49 | -7.5 | 0.65 | 6.6 | 0.51 | -13.6 | 0.65 | -9.7 | 0.5 | -24.2 | 0.65 | 3.2 |
| H6 | 0.48 | 6.7 | 0.62 | 0.0 | 0.46 | 15.0 | 0.64 | -13.5 | 0.49 | 4.3 | 0.64 | -13.5 |
| H12 | 0.4 | 2.6 | 0.52 | 2.0 | 0.38 | 111.1 | 0.52 | -10.3 | 0.43 | 95.5 | 0.49 | -15.5 |
| H24 | 0.45 | -4.3 | 0.51 | -7.3 | 0.41 | -4.7 | 0.54 | -8.5 | 0.42 | -14.3 | 0.54 | -8.5 |
| **Mean** | **0.47** | **2.1** | **0.58** | **0.0** | **0.46** | **10.9** | **0.59** | **-11.9** | **0.48** | **10.4** | **0.59** | **-13.6** |

Table 40 – Fine-tuning results - Variation model prediction results.

| P01->P32 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H2-H0 | 0.53 | 10.4 | 0.6 | 1.7 | 0.59 | 5.4 | 0.6 | -4.8 | 0.75 | 29.3 | 0.63 | 5.0 |
| H6-H2 | 0.49 | 0.0 | 0.5 | 0.0 | 0.5 | 13.6 | 0.51 | -3.8 | 0.44 | -2.2 | 0.56 | 0.0 |
| H12-H6 | 0.39 | -2.5 | 0.53 | 1.9 | 0.4 | 60.0 | 0.52 | -8.8 | 0.21 | 0.0 | 0.49 | -2.0 |
| **Mean** | **0.47** | **2.1** | **0.54** | **0.0** | **0.5** | **18.0** | **0.54** | **-7.4** | **0.47** | **12.8** | **0.56** | **1.8** |

Table 41 – Fine-tuning results - Global model prediction results.

| P01->P32 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.56 | 14.3 | 0.59 | -3.3 | 0.54 | 22.7 | 0.67 | -2.9 | 0.41 | 20.6 | 0.8 | 8.1 |
| H2 | 0.56 | 19.1 | 0.55 | -14.1 | 0.42 | 44.8 | 0.89 | -8.2 | 0.57 | -12.3 | 0.73 | -14.1 |
| H6 | 0.61 | 84.8 | 0.6 | -30.2 | 0.44 | 7.3 | 0.98 | -2.0 | 0.83 | 15.3 | 0.89 | -7.3 |
| H12 | 0.99 | 3.1 | 0.62 | -18.4 | 0.69 | 30.2 | 0.99 | -1.0 | 1 | 104.1 | 0.73 | -21.5 |
| H24 | 0.66 | 29.4 | 0.62 | -28.7 | 0.49 | -5.8 | 0.9 | -6.2 | 1 | 33.3 | 0.64 | -14.7 |
| **Mean** | **0.68** | **19.1** | **0.59** | **-27.1** | **0.52** | **15.4** | **0.88** | **-4.5** | **0.76** | **22.4** | **0.76** | **-11.8** |

### 9.4.2. Percentage of good prediction per failure patients

Table 42 - Percentage of accurate prediction for each patient. The grey boxes are takes that were not realized or not interpretable. The purple boxes are when measured that were not realized as the patient was re-intubated. The yellow cases marked when a patient

| | Fine-Tuning - Decision Tree | | | | |
|---|---|---|---|---|---|
| | h0 | h2 | h6 | h12 | h24 |
| P02 | 22 | 70 | 0 | | |
| P03 | 40 | 64 | 0 | 98 | 100 |
| P05 | 65 | 62 | 100 | | 33 |
| P06 | 81 | 21 | 100 | 100 | |
| P09 | 82 | | | | |
| P14 | 47 | 70 | 100 | | |
| P27 | 54 | | | | |
| P29 | 71 | | | | |
| P36 | 31 | 65 | | | 84 |

Table 43 - Percentage of accurate prediction for each patient. The grey boxes are takes that were not realized or not interpretable. The purple boxes are when measured that were not realized as the patient was re-intubated. The yellow cases marked when a patient

| | Fine-Tuning - Random Forest | | | | |
|---|---|---|---|---|---|
| | h0 | h2 | h6 | h12 | h24 |
| P02 | 4 | 51 | 0 | | |
| P03 | 15 | 34 | 0 | 0 | 0 |
| P05 | 89 | 92 | 98 | | 8 |
| P06 | 88 | 0 | 74 | 100 | |
| P09 | 96 | | | | |
| P14 | 25 | 1 | 31 | | |
| P27 | 29 | | | | |
| P29 | 90 | | | | |
| P36 | 25 | 17 | | | 100 |

### 9.4.3. Fine-tuning results on the validation set

Table 44 - Fine-tuning results on the validation set - Single model prediction results.

| P33->P37 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.32 | 5.0 | 0.69 | -2.5 | 0.39 | 74.2 | 0.65 | -19.1 | 0.71 | 157.3 | 0.64 | -14.6 |
| H2 | 0.64 | 6.8 | 0.41 | -3.5 | 0.68 | -15.3 | 0.45 | 4.7 | 0.57 | -5.8 | 0.30 | -3.2 |
| H6 | / | / | 0.54 | -0.2 | / | / | 0.55 | -8.3 | / | / | 0.68 | -5.6 |
| H12 | / | / | 0.54 | -1.3 | / | / | 0.58 | -10.3 | / | / | 0.65 | -2.9 |
| H24 | 0.57 | 1.8 | 0.55 | 2.6 | 0.55 | -11.8 | 0.53 | -11.4 | 0.72 | 1.4 | 0.65 | 0.8 |
| **Mean** | **0.51** | **4.5** | **0.55** | **-1.0** | **0.54** | **-1.9** | **0.55** | **-10.3** | **0.66** | **25.6** | **0.58** | **-5.6** |

Table 45 - Fine-tuning results on the validation set - Variation model prediction results.

| P33->P37 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H2-H0 | 0.41 | -3.1 | 0.44 | -3.1 | 0.42 | 12.0 | 0.43 | 12.6 | 0.26 | -40.1 | 0.38 | 37.9 |
| H6-H2 | / | / | 0.49 | 2.5 | / | / | 0.48 | -5.9 | / | / | 0.55 | 1.5 |
| H12-H6 | / | / | 0.49 | -5.7 | / | / | 0.53 | 2.9 | / | / | 0.46 | -22.2 |
| **Mean** | **0.41** | **-3.1** | **0.48** | **-2.2** | **0.42** | **12.0** | **0.48** | **2.3** | **0.26** | **-40.1** | **0.46** | **-1.3** |

Table 46 - Fine-tuning results on the validation set - Global model prediction results.

| P33->P37 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.35 | 0.87 | 0.69 | -0.6 | 0.27 | 15.9 | 0.79 | 0.4 | 0.10 | -63.5 | 0.82 | 10.7 |
| H2 | 0.64 | 30.49 | 0.38 | -29.8 | 0.28 | 82.1 | 0.56 | -38.1 | 0.80 | 21.9 | 0.67 | 8.3 |
| H6 | / | / | 0.57 | -29.5 | / | / | 1.00 | -0.2 | / | / | 0.67 | -33.3 |
| H12 | / | / | 0.01 | -20.0 | / | / | 0.99 | -1.4 | / | / | 1.00 | 0.0 |
| H24 | 0.85 | 0.12 | 0.35 | -58.3 | 1.00 | 0.6 | 0.63 | -36.4 | 1.00 | 0.0 | 0.48 | 377.0 |
| **Mean** | **0.61** | **9.17** | **0.40** | **-31.0** | **0.52** | **12.4** | **0.79** | **-15.3** | **0.64** | **-1.9** | **0.73** | **5.0** |

## 9.1. Features optimization: Prediction results

### 9.1.1. Features optimization results on the testing set

Table 47 - Features optimization - Global model prediction results.

| P01->P32 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.56 | -0.5 | 0.59 | 0.0 | 0.54 | -0.7 | 0.65 | -3.3 | 0.31 | -24.4 | 0.81 | 1.5 |
| H2 | 0.96 | 72.0 | 0.60 | 8.4 | 0.91 | 115.5 | 0.82 | -7.6 | 0.81 | 41.2 | 0.70 | -4.7 |
| H6 | 0.75 | 23.6 | 0.61 | 0.8 | 0.57 | 28.9 | 0.91 | -6.8 | 0.90 | 8.2 | 0.84 | -5.2 |
| H12 | 1.00 | 1.0 | 0.57 | -8.4 | 1.00 | 44.9 | 0.78 | -21.1 | 1.00 | 0.0 | 0.72 | -2.1 |
| H24 | 1.00 | 51.5 | 0.61 | -2.3 | 0.76 | 54.7 | 0.83 | -7.8 | 1.00 | 0.0 | 0.58 | -9.7 |
| **Mean** | **0.85** | **19.1** | **0.59** | **-0.6** | **0.75** | **28.3** | **0.80** | **-10.8** | **0.80** | **0.9** | **0.73** | **-4.3** |

### 9.1.1. Features optimization results on the validation set

Table 48 - Features optimization on the validation set - Global model prediction results.

| P33->P37 | Decision Tree | | | | Random Forest | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sensitivity | | specificity | | sensitivity | | specificity | | sensitivity | | specificity | |
| | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] | Result | Var[%] |
| H0 | 0.71 | 105.5 | 0.50 | -27.5 | 0.76 | 182.2 | 0.58 | -27.1 | 0.53 | 410.6 | 0.59 | -27.7 |
| H2 | 0.98 | 52.8 | 0.11 | -70.0 | 0.90 | 215.5 | 0.23 | -57.9 | 0.85 | 5.6 | 0.53 | -20.8 |
| H6 | / | / | 0.67 | 17.0 | / | / | 0.78 | -21.5 | / | / | 0.67 | 0.0 |
| H12 | / | / | 0.00 | -100.0 | / | / | 0.14 | -86.1 | / | / | 0.00 | -100.0 |
| H24 | 1.00 | 18.3 | 0.38 | 7.4 | 0.99 | -1.1 | 0.40 | -35.9 | 1.00 | 0.0 | 0.00 | -100.0 |
| Mean | 0.90 | 58.9 | 0.33 | -34.6 | 0.88 | 132.2 | 0.43 | -45.7 | 0.79 | 138.7 | 0.36 | -49.7 |

### 9.1.1. Results for the success group - Features optimization on the validation set

Table 49 - Validation set - Results for the success group - Features optimization - Global model prediction.

| P33 -> P37 | Decision Tree | | Random Forest | | SVM | |
|---|---|---|---|---|---|---|
| | sensitivity | specificity | sensitivity | specificity | sensitivity | specificity |
| H0 | 0.48 | 0.74 | 0.55 | 0.79 | 0.82 | 0.20 |
| H2 | 0.11 | 0.98 | 0.24 | 0.92 | 0.52 | 0.84 |
| H6 | 0.67 | 0.00 | 0.78 | 0.00 | 0.67 | 0.00 |
| H12 | 0.00 | 0.00 | 0.14 | 0.00 | 0.92 | 0.00 |
| H24 | 0.38 | 1.00 | 0.42 | 0.99 | 0.50 | 1.00 |
| Mean | 0.33 | 0.55 | 0.42 | 0.54 | 0.69 | 0.41 |

[37]

[38]

# 10. References


[1] A. S. Slutsky, "History of Mechanical Ventilation. From Vesalius to Ventilator-induced Lung Injury," *Am. J. Respir. Crit. Care Med.*, vol. 191, no. 10, pp. 1106–1115, May 2015, doi: 10.1164/rccm.201503-0421PP.

[2] B. Ibsen, "Intensive therapy: background and development. 1966," *Int. Anesthesiol. Clin.*, vol. 37, no. 1, pp. 1–14, 1999.

[3] A. B. Baker, "Artificial respiration, The history of an idea," *Med. Hist.*, vol. 15, no. 4, pp. 336–351, Oct. 1971, doi: 10.1017/S0025727300016896.

[4] D. C. Barber and B. H. Brown, "Applied potential tomography," p. 12.

[5] A. S. Slutsky and V. M. Ranieri, "Ventilator-Induced Lung Injury," *N. Engl. J. Med.*, vol. 369, no. 22, pp. 2126–2136, Nov. 2013, doi: 10.1056/NEJMra1208707.

[6] M. Dres *et al.*, "Coexistence and Impact of Limb Muscle and Diaphragm Weakness at Time of Liberation from Mechanical Ventilation in Medical Intensive Care Unit Patients," *Am. J. Respir. Crit. Care Med.*, vol. 195, no. 1, pp. 57–66, Jan. 2017, doi: 10.1164/rccm.201602-0367OC.

[7] G. M. Albaiceta *et al.*, "The central nervous system during lung injury and mechanical ventilation: a narrative review," *Br. J. Anaesth.*, vol. 127, no. 4, pp. 648–659, Oct. 2021, doi: 10.1016/j.bja.2021.05.038.

[8] G. Béduneau *et al.*, "Epidemiology of Weaning Outcome according to a New Definition. The WIND Study," *Am. J. Respir. Crit. Care Med.*, vol. 195, no. 6, pp. 772–783, Mar. 2017, doi: 10.1164/rccm.201602-0320OC.

[9] K. E. A. Burns *et al.*, "Ventilator Weaning and Discontinuation Practices for Critically Ill Patients," *JAMA*, vol. 325, no. 12, p. 1173, Mar. 2021, doi: 10.1001/jama.2021.2384.

[10] M. J. Tobin, "Extubation and the Myth of 'Minimal Ventilator Settings,'" *Am. J. Respir. Crit. Care Med.*, vol. 185, no. 4, pp. 349–350, Feb. 2012, doi: 10.1164/rccm.201201-0050ED.

[11] E. W. Ely *et al.*, "Effect on the Duration of Mechanical Ventilation of Identifying Patients Capable of Breathing Spontaneously," *N. Engl. J. Med.*, vol. 335, no. 25, pp. 1864–1869, Dec. 1996, doi: 10.1056/NEJM199612193352502.

[12] T. D. Girard and G. R. Bernard, "Mechanical Ventilation in ARDS," *Chest*, vol. 131, no. 3, pp. 921–929, Mar. 2007, doi: 10.1378/chest.06-1515.

[13] S. K. Epstein and R. L. Ciubotaru, "Independent Effects of Etiology of Failure and Time to Reintubation on Outcome for Patients Failing Extubation," *Am. J. Respir. Crit. Care Med.*, vol. 158, no. 2, pp. 489–493, Aug. 1998, doi: 10.1164/ajrccm.158.2.9711045.

[14] "Weaning from mechanical ventilation," *Réanimation*, vol. 17, no. 1, pp. 74–97, Feb. 2008, doi: 10.1016/j.reaurg.2007.10.001.

[15] V. M. Boniatti *et al.*, "The Modified Integrative Weaning Index as a Predictor of Extubation Failure," *Respir. Care*, vol. 59, no. 7, pp. 1042–1047, Jul. 2014, doi: 10.4187/respcare.02652.

[16] A. M. Llamas-Álvarez, E. M. Tenza-Lozano, and J. Latour-Pérez, "Diaphragm and Lung Ultrasound to Predict Weaning Outcome," *Chest*, vol. 152, no. 6, pp. 1140–1150, Dec. 2017, doi: 10.1016/j.chest.2017.08.028.

[17] A. W. Thille, F. Boissier, H. Ben Ghezala, K. Razazi, A. Mekontso-Dessap, and C. Brun-Buisson, "Risk Factors for and Prediction by Caregivers of Extubation Failure in ICU Patients: A Prospective Study*," *Crit. Care Med.*, vol. 43, no. 3, pp. 613–620, Mar. 2015, doi: 10.1097/CCM.0000000000000748.

[18] A. W. Thille *et al.*, "Easily identified at-risk patients for extubation failure may benefit from noninvasive ventilation: a prospective before-after study," *Crit. Care*, vol. 20, no. 1, p. 48, Dec. 2016, doi: 10.1186/s13054-016-1228-2.

[19] G. Hernández *et al.*, “Effect of Postextubation High-Flow Nasal Cannula vs Conventional Oxygen Therapy on Reintubation in Low-Risk Patients: A Randomized Clinical Trial,” *JAMA*, vol. 315, no. 13, p. 1354, Apr. 2016, doi: 10.1001/jama.2016.2711.
[20] A. W. Thille *et al.*, “Comparison of the Berlin Definition for Acute Respiratory Distress Syndrome with Autopsy,” *Am. J. Respir. Crit. Care Med.*, vol. 187, no. 7, pp. 761–767, Apr. 2013, doi: 10.1164/rccm.201211-1981OC.
[21] E. W. Ely *et al.*, “Mechanical Ventilator Weaning Protocols Driven by Nonphysician Health-Care Professionals,” *Chest*, vol. 120, no. 6, pp. 454S-463S, Dec. 2001, doi: 10.1378/chest.120.6_suppl.454S.
[22] K. L. Yang and M. J. Tobin, “A Prospective Study of Indexes Predicting the Outcome of Trials of Weaning from Mechanical Ventilation,” *N. Engl. J. Med.*, vol. 324, no. 21, pp. 1445–1450, May 1991, doi: 10.1056/NEJM199105233242101.
[23] A. Savi *et al.*, “Weaning predictors do not predict extubation failure in simple-to-wean patients,” *J. Crit. Care*, vol. 27, no. 2, p. 221.e1-221.e8, Apr. 2012, doi: 10.1016/j.jcrc.2011.07.079.
[24] B. Saugel *et al.*, “Prediction of extubation failure in medical intensive care unit patients,” *J. Crit. Care*, vol. 27, no. 6, pp. 571–577, Dec. 2012, doi: 10.1016/j.jcrc.2012.01.010.
[25] M. Wysocki *et al.*, “Reduced breathing variability as a predictor of unsuccessful patient separation from mechanical ventilation*:,” *Crit. Care Med.*, vol. 34, no. 8, pp. 2076–2083, Aug. 2006, doi: 10.1097/01.CCM.0000227175.83575.E9.
[26] M.-Y. Bien *et al.*, “Comparisons of predictive performance of breathing pattern variability measured during T-piece, automatic tube compensation, and pressure support ventilation for weaning intensive care unit patients from mechanical ventilation*:,” *Crit. Care Med.*, vol. 39, no. 10, pp. 2253–2262, Oct. 2011, doi: 10.1097/CCM.0b013e31822279ed.
[27] F. Gobert *et al.*, “Predicting Extubation Outcome by Cough Peak Flow Measured Using a Built-in Ventilator Flow Meter,” *Respir. Care*, vol. 62, no. 12, pp. 1505–1519, Dec. 2017, doi: 10.4187/respcare.05460.
[28] F. M. Kutchak *et al.*, “Reflex cough PEF as a predictor of successful extubation in neurological patients,” *J. Bras. Pneumol.*, vol. 41, no. 4, pp. 358–364, Aug. 2015, doi: 10.1590/S1806-37132015000004453.
[29] E. Vivier *et al.*, “Inability of Diaphragm Ultrasound to Predict Extubation Failure,” *Chest*, vol. 155, no. 6, pp. 1131–1139, Jun. 2019, doi: 10.1016/j.chest.2019.03.004.
[30] H. Yu *et al.*, “Early prediction of extubation failure in patients with severe pneumonia: a retrospective cohort study,” *Biosci. Rep.*, vol. 40, no. 2, p. BSR20192435, Feb. 2020, doi: 10.1042/BSR20192435.
[31] B.-P. Dubé, M. Dres, J. Mayaux, S. Demiri, T. Similowski, and A. Demoule, “Ultrasound evaluation of diaphragm function in mechanically ventilated patients: comparison to phrenic stimulation and prognostic implications,” *Thorax*, vol. 72, no. 9, pp. 811–818, Sep. 2017, doi: 10.1136/thoraxjnl-2016-209459.
[32] M. Dres and A. Demoule, “Diaphragm dysfunction during weaning from mechanical ventilation: an underestimated phenomenon with clinical implications,” *Crit. Care*, vol. 22, no. 1, p. 73, Dec. 2018, doi: 10.1186/s13054-018-1992-2.
[33] F. Mojoli, B. Bouhemad, S. Mongodi, and D. Lichtenstein, “Lung Ultrasound for Critically Ill Patients,” *Am. J. Respir. Crit. Care Med.*, vol. 199, no. 6, pp. 701–714, Mar. 2019, doi: 10.1164/rccm.201802-0236CI.
[34] A. Ferré *et al.*, “Lung ultrasound allows the diagnosis of weaning-induced pulmonary oedema,” *Intensive Care Med.*, vol. 45, no. 5, pp. 601–608, May 2019, doi: 10.1007/s00134-019-05573-6.
[35] E. DiNino, E. J. Gartman, J. M. Sethi, and F. D. McCool, “Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation,” *Thorax*, vol. 69, no. 5, pp. 431–435, May 2014, doi: 10.1136/thoraxjnl-2013-204111.

[36] A. Soummer *et al.*, “Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress*:,” *Crit. Care Med.*, vol. 40, no. 7, pp. 2064–2072, Jul. 2012, doi: 10.1097/CCM.0b013e31824e68ae.
[37] L. N. Segal *et al.*, “Evolution of pattern of breathing during a spontaneous breathing trial predicts successful extubation,” *Intensive Care Med.*, vol. 36, no. 3, pp. 487–495, Mar. 2010, doi: 10.1007/s00134-009-1735-6.
[38] A. Perren *et al.*, “Patients’ prediction of extubation success,” *Intensive Care Med.*, vol. 36, no. 12, pp. 2045–2052, Dec. 2010, doi: 10.1007/s00134-010-1984-4.
[39] M.-H. Hsieh, M.-J. Hsieh, C.-M. Chen, C.-C. Hsieh, C.-M. Chao, and C.-C. Lai, “An Artificial Neural Network Model for Predicting Successful Extubation in Intensive Care Units,” *J. Clin. Med.*, vol. 7, no. 9, p. 240, Aug. 2018, doi: 10.3390/jcm7090240.
[40] A. R. Baptistella *et al.*, “Prediction of extubation outcome in mechanically ventilated patients: Development and validation of the Extubation Predictive Score (ExPreS),” *PLOS ONE*, vol. 16, no. 3, p. e0248868, Mar. 2021, doi: 10.1371/journal.pone.0248868.
[41] H. Taud and J. F. Mas, “Multilayer Perceptron (MLP),” in *Geomatic Approaches for Modeling Land Change Scenarios*, M. T. Camacho Olmedo, M. Paegelow, J.-F. Mas, and F. Escobar, Eds. Cham: Springer International Publishing, 2018, pp. 451–455. doi: 10.1007/978-3-319-60801-3_27.
[42] M. J. Tobin and A. Jubran, “Variable performance of weaning-predictor tests: role of Bayes’ theorem and spectrum and test-referral bias,” *Intensive Care Med.*, vol. 32, no. 12, pp. 2002–2012, Dec. 2006, doi: 10.1007/s00134-006-0439-4.
[43] K. Y. Aristovich, B. C. Packham, H. Koo, G. S. dos Santos, A. McEvoy, and D. S. Holder, “Imaging fast electrical activity in the brain with electrical impedance tomography,” *NeuroImage*, vol. 124, pp. 204–213, Jan. 2016, doi: 10.1016/j.neuroimage.2015.08.071.
[44] M. Akhtari-Zavare and L. A. Latiff, “Electrical Impedance Tomography as a Primary Screening Technique for Breast Cancer Detection,” *Asian Pac. J. Cancer Prev.*, vol. 16, no. 14, pp. 5595–5597, Sep. 2015, doi: 10.7314/APJCP.2015.16.14.5595.
[45] R. Li, J. Gao, Y. Li, J. Wu, Z. Zhao, and Y. Liu, “Preliminary Study of Assessing Bladder Urinary Volume Using Electrical Impedance Tomography,” *J. Med. Biol. Eng.*, vol. 36, no. 1, pp. 71–79, Feb. 2016, doi: 10.1007/s40846-016-0108-1.
[46] A. Adler *et al.*, “Whither lung EIT: Where are we, where do we want to go and what do we need to get there?,” *Physiol. Meas.*, vol. 33, no. 5, pp. 679–694, May 2012, doi: 10.1088/0967-3334/33/5/679.
[47] E. Teschner, M. Imhoff, and S. Leonhardt, “Electrical impedance tomography: The realisation of regional ventilation monitoring 2nd edition.”
[48] C. Putensen, B. Hentze, S. Muenster, and T. Muders, “Electrical Impedance Tomography for Cardio-Pulmonary Monitoring,” *J. Clin. Med.*, vol. 8, no. 8, p. 1176, Aug. 2019, doi: 10.3390/jcm8081176.
[49] K. Sudha, M. Israil, S. Mittal, and J. Rai, “Soil characterization using electrical resistivity tomography and geotechnical investigations,” *J. Appl. Geophys.*, vol. 67, no. 1, pp. 74–79, Jan. 2009, doi: 10.1016/j.jappgeo.2008.09.012.
[50] A. Perrone, V. Lapenna, and S. Piscitelli, “Electrical resistivity tomography technique for landslide investigation: A review,” *Earth-Sci. Rev.*, vol. 135, pp. 65–82, Aug. 2014, doi: 10.1016/j.earscirev.2014.04.002.
[51] D. Silvera-Tawil, D. Rye, M. Soleimani, and M. Velonaki, “Electrical Impedance Tomography for Artificial Sensitive Robotic Skin: A Review,” *IEEE Sens. J.*, vol. 15, no. 4, pp. 2001–2016, Apr. 2015, doi: 10.1109/JSEN.2014.2375346.
[52] K. Liu *et al.*, “Artificial Sensitive Skin for Robotics Based on Electrical Impedance Tomography,” *Adv. Intell. Syst.*, vol. 2, no. 4, p. 1900161, Apr. 2020, doi: 10.1002/aisy.201900161.
[53] Md. R. Islam and Md. A. Kiber, “Electrical Impedance Tomography imaging using Gauss-Newton algorithm,” in *2014 International Conference on Informatics, Electronics &*

*Vision (ICIEV)*, Dhaka, Bangladesh, May 2014, pp. 1–4. doi: 10.1109/ICIEV.2014.6850719.
[54] K. Wu, J. Yang, X. Dong, F. Fu, F. Tao, and S. Liu, "Comparative Study of Reconstruction Algorithms for Electrical Impedance Tomography," p. 4.
[55] V. Sarode, S. Patkar, and A. N Cheeran, "Comparison of 2-D Algorithms in ElT based Image Reconstruction," *Int. J. Comput. Appl.*, vol. 69, no. 8, pp. 6–11, May 2013, doi: 10.5120/11860-7642.
[56] B. Grychtol, G. Elke, P. Meybohm, N. Weiler, I. Frerichs, and A. Adler, "Functional Validation and Comparison Framework for EIT Lung Imaging," *PLoS ONE*, vol. 9, no. 8, p. e103045, Aug. 2014, doi: 10.1371/journal.pone.0103045.
[57] A. Adler *et al.*, "GREIT: a unified approach to 2D linear EIT reconstruction of lung images," *Physiol. Meas.*, vol. 30, no. 6, pp. S35–S55, Jun. 2009, doi: 10.1088/0967-3334/30/6/S03.
[58] H. Li *et al.*, "Optimized Method for Electrical Impedance Tomography to Image Large Area Conductive Perturbation," *IEEE Access*, vol. 7, pp. 140734–140742, 2019, doi: 10.1109/ACCESS.2019.2944209.
[59] A. Adler, R. Amyot, R. Guardo, J. H. T. Bates, and Y. Berthiaume, "Monitoring changes in lung air and liquid volumes with electrical impedance tomography," *J. Appl. Physiol.*, vol. 83, no. 5, pp. 1762–1767, Nov. 1997, doi: 10.1152/jappl.1997.83.5.1762.
[60] C. Ngo *et al.*, "Linearity of electrical impedance tomography during maximum effort breathing and forced expiration maneuvers," *Physiol. Meas.*, vol. 38, no. 1, pp. 77–86, Jan. 2017, doi: 10.1088/1361-6579/38/1/77.
[61] N. Coulombe, H. Gagnon, F. Marquis, Y. Skrobik, and R. Guardo, "A parametric model of the relationship between EIT and total lung volume," *Physiol. Meas.*, vol. 26, no. 4, pp. 401–411, Aug. 2005, doi: 10.1088/0967-3334/26/4/006.
[62] S. D. Reinartz *et al.*, "EIT monitors valid and robust regional ventilation distribution in pathologic ventilation states in porcine study using differential DualEnergy-CT (ΔDECT)," *Sci. Rep.*, vol. 9, no. 1, p. 9796, Dec. 2019, doi: 10.1038/s41598-019-45251-7.
[63] D. R. Hess, "Recruitment Maneuvers and PEEP Titration," *Respir. Care*, vol. 60, no. 11, pp. 1688–1704, Nov. 2015, doi: 10.4187/respcare.04409.
[64] C. Gomez-Laberge, J. H. Arnold, and G. K. Wolf, "A Unified Approach for EIT Imaging of Regional Overdistension and Atelectasis in Acute Lung Injury," *IEEE Trans. Med. Imaging*, vol. 31, no. 3, pp. 834–842, Mar. 2012, doi: 10.1109/TMI.2012.2183641.
[65] T. H. Becher *et al.*, "Assessment of respiratory system compliance with electrical impedance tomography using a positive end-expiratory pressure wave maneuver during pressure support ventilation: a pilot clinical study," *Crit. Care*, vol. 18, no. 6, p. 679, Dec. 2014, doi: 10.1186/s13054-014-0679-6.
[66] O. C. Radke, T. Schneider, A. R. Heller, and T. Koch, "Spontaneous Breathing during General Anesthesia Prevents the Ventral Redistribution of Ventilation as Detected by Electrical Impedance Tomography," *Anesthesiology*, vol. 116, no. 6, pp. 1227–1234, Jun. 2012, doi: 10.1097/ALN.0b013e318256ee08.
[67] P. Blankman, D. Hasan, M. S. van Mourik, and D. Gommers, "Ventilation distribution measured with EIT at varying levels of pressure support and Neurally Adjusted Ventilatory Assist in patients with ALI," *Intensive Care Med.*, vol. 39, no. 6, pp. 1057–1062, Jun. 2013, doi: 10.1007/s00134-013-2898-8.
[68] F. Longhini *et al.*, "Electrical impedance tomography during spontaneous breathing trials and after extubation in critically ill patients at high risk for extubation failure: a multicenter observational study," *Ann. Intensive Care*, vol. 9, no. 1, p. 88, Dec. 2019, doi: 10.1186/s13613-019-0565-0.
[69] Z. Zhao, K. Möller, D. Steinmann, I. Frerichs, and J. Guttmann, "Evaluation of an electrical impedance tomography-based global inhomogeneity index for pulmonary

ventilation distribution," *Intensive Care Med.*, vol. 35, no. 11, p. 1900, Nov. 2009, doi: 10.1007/s00134-009-1589-y.
[70] Z. Zhao, S. Pulletz, I. Frerichs, U. Müller-Lisse, and K. Möller, "The EIT-based global inhomogeneity index is highly correlated with regional lung opening in patients with acute respiratory distress syndrome," *BMC Res. Notes*, vol. 7, no. 1, p. 82, Dec. 2014, doi: 10.1186/1756-0500-7-82.
[71] R. Fluss, D. Faraggi, and B. Reiser, "Estimation of the Youden Index and its Associated Cutoff Point," *Biom. J.*, vol. 47, no. 4, pp. 458–472, Aug. 2005, doi: 10.1002/bimj.200410135.
[72] A. Adler and W. R. B. Lionheart, "Uses and abuses of EIDORS: an extensible software base for EIT," *Physiol. Meas.*, vol. 27, no. 5, pp. S25–S42, May 2006, doi: 10.1088/0967-3334/27/5/S03.
[73] F. Thürk *et al.*, "Effects of individualized electrical impedance tomography and image reconstruction settings upon the assessment of regional ventilation distribution: Comparison to 4-dimensional computed tomography in a porcine model," *PLOS ONE*, vol. 12, no. 8, p. e0182215, Aug. 2017, doi: 10.1371/journal.pone.0182215.
[74] F. Thürk *et al.*, "Influence of reconstruction settings in electrical impedance tomography on figures of merit and physiological parameters," *Physiol. Meas.*, vol. 40, no. 9, p. 094003, Sep. 2019, doi: 10.1088/1361-6579/ab248e.
[75] G. Hahn, J. Dittmar, A. Just, M. Quintel, and G. Hellige, "Different approaches for quantifying ventilation distribution and lung tissue properties by functional EIT," *Physiol. Meas.*, vol. 31, no. 8, pp. S73–S84, Aug. 2010, doi: 10.1088/0967-3334/31/8/S06.
[76] I. Frerichs *et al.*, "Chest electrical impedance tomography examination, data analysis, terminology, clinical use and recommendations: consensus statement of the TRanslational EIT developmeNt stuDy group," *Thorax*, vol. 72, no. 1, pp. 83–93, Jan. 2017, doi: 10.1136/thoraxjnl-2016-208357.
[77] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *J. Chemom.*, vol. 18, no. 6, pp. 275–285, Jun. 2004, doi: 10.1002/cem.873.
[78] P. Probst, M. Wright, and A.-L. Boulesteix, "Hyperparameters and Tuning Strategies for Random Forest," 2018, doi: 10.48550/ARXIV.1804.03515.
[79] N. Altman and M. Krzywinski, "Ensemble methods: bagging and random forests," *Nat. Methods*, vol. 14, no. 10, pp. 933–934, Oct. 2017, doi: 10.1038/nmeth.4438.
[80] M. Sabzevari, G. Martínez-Muñoz, and A. Suárez, "Vote-boosting ensembles," *Pattern Recognit.*, vol. 83, pp. 119–133, Nov. 2018, doi: 10.1016/j.patcog.2018.05.022.
[81] W. S. Noble, "What is a support vector machine?," *Nat. Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006, doi: 10.1038/nbt1206-1565.
[82] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950, doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.