

Applying AI to Manage Acute and Chronic Clinical Condition



Rachael Hagan, Charles J. Gillan, and Murali Shyamsundar

Abstract Computer systems deployed in hospital environments, particularly physiological and biochemical real-time monitoring of patients in an Intensive Care Unit (ICU) environment, routinely collect a large volume of data that can hold very useful information. However, the vast majority are either not stored and lost forever or are stored in digital archives and seldom re-examined. In recent years, there has been extensive work carried out by researchers utilizing Machine Learning (ML) and Artificial Intelligence (AI) techniques on these data streams, to predict and prevent disease states. Such work aims to improve patient outcomes, to decrease mortality rates and decrease hospital stays, and, more generally, to decrease healthcare costs.

This chapter reviews the state of the art in that field and reports on our own current research, with practicing clinicians, on improving ventilator weaning protocols and lung protective ventilation, using ML and AI methodologies for decision support, including but not limited to Neural Networks and Decision Trees. The chapter considers both the clinical and Computer Science aspects of the field. In addition, we look to the future and report how physiological data holds clinically important information to aid in decision support in the wider hospital environment.

Keywords Healthcare · Predictive analytics · Artificial Intelligence · Clinical decision support

The chapter relates to the technical priorities Data Analytics of the European Big Data Value Strategic Research and Innovation Agenda [1]. It addresses the horizontal concern Data Analytics of the BDV Technical Reference Model and addresses the vertical concerns of Big Data Types and Semantics. The chapter further relates to the Reasoning and Decision Making cross-sectorial technology enablers of the AI, Data and Robotics Strategic Research, Innovation and Deployment Agenda [2].

R. Hagan (✉) · C. J. Gillan · M. Shyamsundar
Queen's University Belfast, Belfast, Northern Ireland, UK
e-mail: rhagan09@qub.ac.uk

1 Overview

The chapter begins by discussing Intensive Care medicine and the types of machines and data that is being recorded continuously, and thus producing ‘Big Data’. It proceeds to explain some of the challenges that can arise when working with such data, including measurement errors and bias. The subsequent section explains some of the common methodologies used to provide predictive analytics with examples given for both acute and chronic clinical conditions, and discuss our own work for the promotion of lung protective ventilation, to highlight the accuracies that can be achieved when pairing health ‘Big Data’ with common machine learning methodologies. The chapter concludes by discussing the future of this field and how we, as a society, can provide value to our healthcare systems by utilizing the routinely collected data at our disposal.

2 Intensive Care Medicine and Physiological Data

Intensive Care Units (ICU) offer expensive and labor-intensive treatments for the critically ill and are therefore a costly resource for the health sector around the world. They can also be referred to as Intensive Therapy Units (ITU) or Critical Care Units (CCU). UK estimates in 2007 highlighted that intensive care in the NHS costs £719 million per year [3]. Comparably, American studies reported in 2000 have shown that median costs per ICU stay can range between \$10,916 and \$70,501 depending on the length of stay [4]. A typical length of stay varies depending on the condition of the patient with studies showing the mean length of stay being 5.04 days [3], while the condition of the patient can change quickly and sometimes unpredictably.

Patients will normally be admitted to intensive care after a serious accident, a serious condition such as a stroke, an infection or for surgical recovery. Throughout their stay in ICU, these patients are monitored closely due to their critical condition, and on average require one nurse for every one or two patients. Many devices and tests may be used to ensure the correct level of care is provided. Patients will be placed on machines to monitor their condition, support organ function and allow for the detection of any improvements or deterioration. The functions of these machines can vary: from the monitoring of vital parameters such as heart rate via the patient monitor to the use of mechanical ventilators that provide respiratory function when a patient is not able to do so themselves.

The workload on the clinical staff in ICU is intense and so utilizing Big Data analytics will allow for healthcare providers to improve their efficiency through better management of resources, detection of decompensation and adverse events, and treatment optimization, among many benefits to both patient outcomes and hospital costs [5].

2.1 *Physiological Data Acquisition*

Health data recorded in ICU can be classified as ‘Big Data’ due to the volume, complexity, diversity and timeliness of the parameters [6], the aim being to turn these ‘Big Data’ records into a valuable asset in the healthcare industry.

As highlighted, patients requiring critical care are placed on numerous monitors and machines to help with and provide their care. Equipment can include, but is not limited to [7]:

- Patient monitoring system to monitor clinical parameters such as electrocardiogram (ECG), peripheral oxygen saturation, blood pressure, temperature.
- Organ support systems such as mechanical ventilator, extracorporeal organ support such as continuous renal replacement therapy.
- Syringe pump for the delivery of medicines.

These machines are all monitoring continuously (24×7) and thus representing one example of the emerging field of Big Data. It is important to wean patients off these machines as quickly as possible to avoid dependency and to lower the risk of infection. In addition to the organ support machines, the electronic health record includes laboratory data, imaging reports such as X-rays and CT scans, and daily review record.

2.1.1 *Time Series Data*

The human physiologic state is a time-dependent picture of the functions and mechanisms that define life and is amenable to mathematical modeling and data analysis. Physiology involves processes operating across a range of time scales resulting in different granularities. Parameters such as heart rate and brain activity are monitored on the millisecond level while others such as breathing have time windows over minutes and blood glucose regulation over hours. Analysis of instantaneous values in these circumstances is rarely of value. On the other hand, application of analytical techniques for time series offers the opportunity to investigate both the trends in individual physiological variables and the temporal correlation between them, thereby enabling the possibility to make predictions.

Features can also be extracted from these time series using packages such as the Python *tsfresh* software toolkit to use as input into Machine Learning models and to gain further insights into the relationship between the parameter and time [8]. By analyzing the time series we can make predictions on the future trajectory and alert care givers of possible issues in order to prevent complications.

Prior research has shown that these streams of data have very useful information buried in them, yet in most medical institutions today, the vast majority of the data collected is either dropped and lost forever, or is stored in digital archives and seldom reexamined [9].

In the past number of years there has been a rapid implementation of Electronic health records, or EHRs, around the world. EHRs are a common practice to record and store real-time, patient health data, enabling authorized users to track a patient from initial admission into the hospital, their deterioration/improvement, diagnoses, and all physiological parameters monitored and drugs given across all healthcare systems.

2.1.2 Publicly Available Datasets

As part of a global effort to improve healthcare, numerous institutions have put together publicly available data sets based on their EHRs for people to use in research; enabling visualization, analysis, and model development.

One of the best-known and commonly used publicly available databases is MIMIC, the Multiparameter Intelligent Monitoring in Intensive Care database. Produced from the critical care units of the Beth Deaconess Medical Centre at Harvard Medical School, Boston, this large, freely available database consists of de-identified health-related data associated with over 40,000 patients who were in critical care between 2001 and 2012 [10]. After being approved access, users are provided with vital signs, medications, laboratory measurements, observations, and charted notes. Furthermore, waveforms are available for use, along with patient demographics and diagnoses.

PhysioNet offers access to a large collection of health data and related open-source software, including the MIMIC databases [11]. All data recorded in these publically available databases are anonymous, ensuring no patient can be identified from their data.

3 Artificial Intelligence in ICU

The broad field of Data Science has emerged within the discipline of Computer Science over the past 10 years, approximately. Its roots arguably lie in the work of Fayyad et al. [12], which defined a pipelined process for the extraction of abstract models from data using an amalgam of techniques, including Statistics and Machine Learning. Artificial Intelligence (AI) is the ability of a computer system to perform tasks that aim to reach the level of intelligent thinking of the human brain through applying machine learning methodologies, further details are discussed in the methodologies section.

3.1 Challenges

As in majority of real-world applications, there comes a series of challenges that arise when applying Big Data-driven analytics to such sensitive and intense data. We

must aim to build trust and quality between the decision support tools and the end users. It is always important to question if there is real clinical impact from carrying out such work and exploring AI methodologies for solving particular problems.

3.1.1 Data Integrity

Before building any Big Data-driven analytic system, it is important to ensure data has been collected and preprocessed correctly, with any errors handled accurately. It is crucial to ensure data harmonization, the process of bringing together all data formats from the different machines and tests into one database. Without this step, the algorithms will produce unreliable results, which in turn could result in a critically dangerous implementation and a lack of trust in the system. Errors in the data can be due to a variety of reasons:

- **Data input errors:** In intensive care, caregivers often have to input readings on the system. Fields will have a required metric in which the data should be entered and these may sometimes be ignored or mishandled. For example, height that should be entered in centimeters but a user may enter the reading in meters, e.g., 1.7 m instead of the 170 cm.
- **Sensor errors:** With the complexity and multitude of monitors, leads, and machines that a patient can be on at any given time, the sensors can sometimes fail or miss a reading. Sensors can be disconnected for a period of time to deliver medicines or for imaging tests. Patients' movements in the bed can cause a sensor error or unexpected result. These errors will present as outliers in the data and should be dealt with accordingly as to not throw off any predictive modeling.
- **Bias in the data:** AI methodologies are only as good as the data of which they are trained on. If this data contains, e.g., racial or gender biases, the algorithm will learn from this and produce similar results such as women not being given the same number of tests as men. Similarly, statistical biases can be present due to small sample numbers from underrepresented groups, for instance, an algorithm only being trained with White patients may not pick up the same diagnosis when presented with Hispanic patients [13]. Furthermore, selection biases exist when the selection of study participants is not a true representation of the true population, resulting in both cultural and social differences.
- **Bias in the study:** Studies can also contain information bias. These include measurement errors for continuous variables and misclassification for categorical variables. For example, a tumor stage being misdiagnosed would lead to algorithms being trained on incorrect data [14].

We must ensure to implement appropriate techniques such as imputation, smoothing and oversampling to prevent errors in our data and build trust with the user.

3.1.2 Alert Fatigue

When building decision support tools, it is crucial to ensure the integrity of alerts raised. Too many can result in alert fatigue, leading to alarms being switched off or ignored. Research has shown that there is on average 10.6 alarms per hour, too many for a single person to handle amidst their already busy work schedule [15].

In addition, 16% of health care professionals will switch off an alarm and 40.4% will ignore or silence the alarm [16].

This influx of alarms resulting in the monitoring system being turned off will lead to vital physiological problems being missed. We need to be confident that our systems are only producing true alerts and that false alerts generated are minimized.

It has further been highlighted that reducing the number of alerts repeated per patient will help reduce override rates and fatigue [17].

3.1.3 Bringing AI Systems to Clinical Trials

While we know these AI models, derived from Big Data-driven predictive analytics, can provide value to the healthcare industry, we must be able to test them in real time in order for them to be implemented and used throughout the world. This requires a clinical trial to be carried out using the system on patients in real time. While there exist the SPIRIT 2013 and CONSORT 2010 checklists for clinical trials, covering checklists for what you intended to do and what you actually did, respectively, neither of these include steps for AI implementation [18].

With the rise of AI, these guidelines have been extended in 2020 to include steps for reporting the quality of data, errors, clinical context, intended use, and any human interactions involved. These additional steps allow reviewers to better evaluate and compare the quality of research and thus systems created in the future [19, 20]. Other authors are further expanding these checklists to include reporting for diagnostic accuracies and prediction models [21, 22].

Researchers have analyzed the current work that has been published in this area [23]. With 93% of papers exploring model development to demonstrate potential for such systems, and only 5% validating these models with data from other centers. A further 1% of the work currently published reports on models that have actually been implemented and tested in real-time hospital systems and the final 1% of research have integrated their models in routine clinical practice which has proven to work. This summary highlights the large potential for the future of AI in healthcare systems, with huge opportunities for bringing forward accurate models to real-world clinical trials.

3.2 *AI Methodology*

Machine Learning methodologies are commonly used to enable machines to become ‘intelligent’ and can be classified as being supervised, predicting using a labeled dataset, i.e., a known output or target, or unsupervised, i.e., finding patterns or groupings in unlabeled data [24]. Common methodologies are utilized across the board for predictive and analytical purposes. This chapter focuses on commonly used supervised learning techniques; however, unsupervised methods can further be used to understand data. We can categorize supervised learning techniques into regression and classification models. Regression techniques aim to find relationships between one dependent variable and a series of other independent variables, e.g., a time series as previously discussed is common in physiological data. Classification techniques on the other hand attempt to label outcomes and draw conclusions from observed values, e.g., if patient has disease or not.

The question of whether the numerical models that are generated can actually be understood by humans has become a hot research topic. In the United States, DARPA has conducted significant investigative research in this field [25]. Other research teams have begun to define the concept of explainable AI with respect to several problem domains. They identify three classifications: opaque systems that offer no insight; interpretable systems where mathematical analysis of the algorithm is viable; and comprehensible systems that emit symbols enabling user-driven explanations of how a conclusion is reached [26].

We know from previous research that utilizing Big Data within the ICU can lead to many benefits to both patient and hospital. We can not only greatly improve the care given and thus patient outcomes, see Table 1, but also reduce hospital costs [5] and the stress levels of care givers [27]. McGregor and colleagues have demonstrated the viability of monitoring physiological parameters to detect sleep apnea in neo-natal ICU leading to the software architecture of the IBM InfoSphere product, which has now been extended into a Cloud environment (the Artemis project) making the functionality available at remote hospital sites [28, 29].

Patient deterioration can often be missed due to the multitude of and complicated relational links between physiological parameters. AI-driven multivariate analysis has the potential to ameliorate the work load of ICU staff. Multiple studies have shown AI to be comparable to routine clinical decision making, including ECG analysis, delirium detection, sedation, and identification of septic patients [28, 37–39].

AI-driven predictive analytics within healthcare most commonly uses supervised learning approaches, due to which we aim to base algorithms and decisions on previous examples and train our models with accurate outcomes, in particular regression analysis is used for time series data.

Below we review the more widely adopted Machine Learning methodologies developed [31], and examples are highlighted in Table 1, where these have been used in previous research.

Table 1 Highlighting applications of AI in healthcare

Area	Example	AI methodologies	Limitations
Mortality rates	Prediction of postoperative 30-day mortality [38]	CNN	Generalizability, completeness and accuracy of data
Deterioration	Prediction of patients at risk for deterioration within the next 6 h [39]	Random Forest	False alarms
Disease prediction	Sepsis prediction [40–42]	Gradient boosting Random forest Rules-based system	Generalization, explainability, data quality
Disease prediction	Prediction of kidney failure [35]	RNN	Unbalanced dataset
Data visualization	Display of optimum ventilator parameters to ensure compliance [43]	Regression	Lack of complete adaption
Patient outcomes	Prediction of ICU outcome based on data collected in first 48 h of admission [44]	Expert system ANN	Unbalanced dataset
Patient outcomes	Interpretable deep models for ICU prediction of acute lung injury (ALI) [45]	Gradient boosting	Generalization
Mortality	Prediction of mortality from respiratory distress among long-term mechanically ventilated patients [46]	Regression Decision trees	Limited data
Diagnosis	Comparison of classification models for the diagnosis of asthma [47]	Ensemble models ANN	Small dataset, Noisy waveforms
Risk prediction	System for predicting the risk of heart attacks [48]	Neural network Decision tree	Small dataset, generalization
Risk prediction	Predicting patients at risk for prolonged ventilation [49]	Gradient boosting decision trees	Information bias
Disease prediction	Early prediction of circulatory failure in ICU [50]	Gradient-boosted ensembles	False alarms, generalization
Outcome prediction	ECG analysis to predict atrial fibrillation [36]	CNN	Information bias
Complication prediction	Prediction of several severe complications post open heart surgery [51]	RNN	Generalization, Noisy data

3.2.1 Expert Systems

In a rule-based expert system, also known as a knowledge-based clinical decision support system, knowledge is represented by a series of IF-THEN rules that are created by knowledge acquisition, which involves observing and interviewing human experts and finalizing rules in format 'IF X happens THEN do Y'. These rules cannot be changed, learnt from or adapted to different environments, meaning the human experts must manually monitor and modify the knowledge base through careful management of the rules. Expert systems allow us to view and understand each of the rules applied by the system.

The systems can take over some mundane decision tasks and discussions of health care professionals, saving vital time and money. An example expert system was created for Diabetes patients to provide decision support for insulin adjustment based on simple rules depending on the regimen that patients were placed on [32].

3.2.2 Decision Trees

A decision tree can be used to visually represent a series of decisions used to reach a certain conclusion, useful when exploring medical reasons behind decisions made. The tree starts at the root, asking a specific question to split the data by the given condition. The tree then splits into branches and edges at each node representing the input variables, continually splitting by conditions until the final decision is achieved at the leaf node or the output variable.

They can also be referred to as Classification and Regression Trees (CART) as they can solve both classification and regression problems. A classification tree will arrive at a binary condition, leaf node, i.e., patient survives or not, whereas the regression trees will predict a certain continuous value, i.e., the heart rate of a patient.

The tree will not only explore the conditions used to split the data at each decision but also the features used and which features are the most significant at splitting the data, added as top-level nodes. Researchers have utilized simple decision tree models for the classification of patients with diabetes, among other disease states. Features include age, BMI, and both systolic and diastolic blood pressure of the patient to arrive at a decision whether the patient has diabetes or is healthy. Figure 1 shows how the features are split and decisions are made, in this circumstance [33].

When building decision tree models it is important to monitor the maximum depth of the tree to avoid overfitting and lengthy training times.

3.2.3 Ensemble Methods

To achieve the greatest predictive performance when working with complex problems, ensembles can be created. The decision trees are combined in different ways

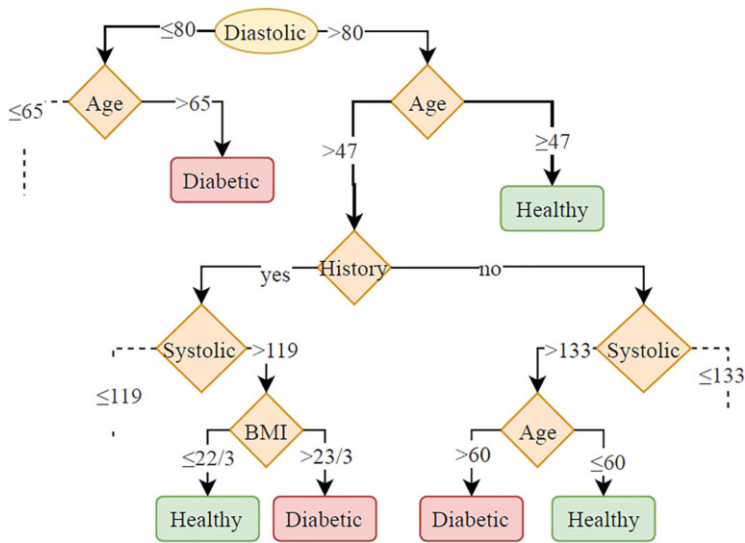


Fig. 1 Decision tree for the classification of diabetes [20]

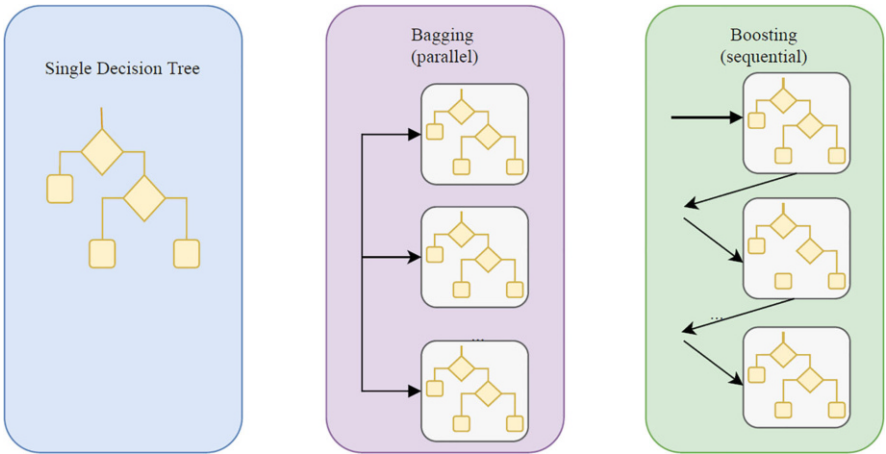


Fig. 2 Decision trees and ensemble methods

depending on the methodology used. The methods can be categorized as bagging or boosting (Fig. 2).

Boosting methods build the trees in a sequential way; for each predicted value multiple models or decision trees are made using different features and combinations of features, then weights are given to these models based on their error so that the final prediction is the most accurate. AdaBoost is an example of this method where very short trees are produced and higher weights are given to

more difficult decisions. GradientBoosting further combines boosting with gradient descent, allowing for the optimization of loss functions.

In contrast, for bagging methods, each model is given equal weights, they are combined in parallel and all of the predictions are averaged to get a final, most accurate decision, examples include Bagging and Extra Trees. Bagging can be extended to the RandomForest algorithm by randomly selecting the features used, and decision trees are built to have as many layers as possible.

3.2.4 Neural Networks

A neural network saves a lot of time when working with large amounts of data by combining variables, figuring out which are important and finding patterns that humans might not ever see. The neural network is represented as a set of interconnected nodes, connected by neurons.

They feed the weighted sum of the input values through an activation function, which takes the value and transforms it before returning an output. These activation functions in turn improve the way the neural network learns and allows for more flexibility to model complex relationships between the input and output (Fig. 3).

The neural network can be described as ‘Deep Learning’ when it has multiple hidden layers between the input and output layers. The neural network learns by

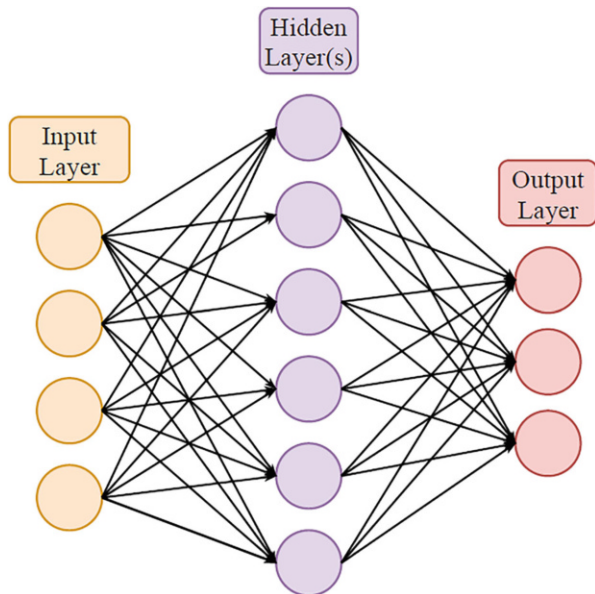


Fig. 3 Neural network structure

figuring out what it got wrong and working backwards to determine what values and connections made the prediction incorrect.

Additionally, there are different types of neural networks, and the list continues to expand as researchers propose new types. There are feedforward neural networks, where all the nodes only feed into the next layer from initial inputs to the output. Recurrent neural networks (RNN) make it possible to feed the output of a node back into the model as an input the next time you run it. Nodes in one layer can be connected to each other and even themselves. Furthermore, they work well with sequential data as they remember previous outputs. The long short-term memory form of the RNN enables the model to store information over longer periods of time, ideal for modeling time series data. Additionally, there are convolutional neural networks (CNN) which look at windows of variables rather than one at a time. Convolution applies a filter, or transformation, to the windows to create features. When working with large databases, pooling is a technique that takes a huge number of variables to create a smaller number of features. The neural network will then use the features generated by convolution and pooling to give its output.

The key parameters to distinguish neural networks are the number of layers and the shape of the input and the output layers.

Researchers have utilized many different formats of neural networks for prediction problems. A medical decision support tool for patient extubation was developed using a multilayer perception artificial neural network, a class of feedforward neural networks. The input layer consisted of 8 input parameters, defining 17 perceptions, and 2 perceptions in the output layer for prediction output. Perceptions can be described as a classifying decision. They explored the change in performance based on the number of perceptions in the hidden layer: 19 producing highly accurate results [34]. Other studies have shown that RNNs produce accurate results for the prediction of kidney failure [35] and CNNs have shown promise in the prediction of atrial fibrillation in ECG analysis [36],

However, it is difficult to explain predictions from neural networks to healthcare professionals due to having to understand a particular weight as a discrete piece of knowledge, although work is being done in the area [37].

With common limitations reoccurring, such as generalization, dataset size, noisy and unbalanced data, we highlight the importance of continuing research and building of larger datasets, across multiple centers, and further the exploration of methodologies to smooth noisy data in order to advance the work of producing accurate systems that can be implemented in real-world healthcare settings.

Imaging and waveforms are a further, huge division of physiological monitoring. Machines such as X-rays and CTs can produce images of internal organs to provide a greater insight into patient state. In addition, ECG waveforms and brain waves can be analyzed for diagnoses. Thus, signal processing is an integral part of understanding physiological data and getting the full picture of patient state.

Furthermore, Machine Learning methodologies can be used for natural language processing to analyze lab notes and patient charts in order to summarize and detect common words and phrases used by care givers, in turn automating the process of flipping through hundreds of pages of records and saving vital time.

4 Use Case: Prediction of Tidal Volume to Promote Lung Protective Ventilation

Around 44.5% of patients in ICU are placed on mechanical ventilation to support respiratory function at any given hour [52]. However, the delivery of high tidal volume values can often lead to lung injury. Tidal volume being the amount of air delivered to the lungs; it is common knowledge amongst critical care providers that tidal volume values should be no greater than 8 ml per kg of ideal body weight. In our recent work we explored regressors for ensemble methods and the long short-term memory (LSTM) form of neural networks to predict tidal volume values to aid in lung protective ventilation [53].

Data acquisition took place at the Regional Intensive Care Unit (RICU), Royal Victoria Hospital, Belfast, over a 3-year period and the VILIAAlert system was introduced [54]. The data streams were monitored against the thresholds for lung protective ventilation and if thresholds were breached continuously, an alert was raised. We then turned our attention to predicting these alerts with the aim of preventing violations and protecting the patient's lungs. The VILIAAlert system ran for nearly 3 years, recording minutely tidal volume values for almost a thousand patients.

As discussed, noisy signals are common in ICU data. Time series often needs to be filtered to remove noise in the data and produce smooth signals. Methods such as moving average and exponential smoothing can be applied to the data to extract true signals, such as the work we carried out to extract true tidal volume trends [53]. Figure 4 shows how smoothing the time series, shown in blue, removes anomaly in the data, the large jumps, and extracts the true patient trend as shown in red. This work is related to the efforts of the international project known as the Virtual Physiological Human (<http://www.vph-institute.org>), which seeks to use individualized physiology-based computer simulations in all aspects of prevention, diagnosis, and treatment of diseases [55]. This computational approach has provided immediate insight into the COVID-19 pandemic [56].

We compare multiple regressor ensemble methods for initially predicting 15 min ahead. For each patient, we use the *tsfresh* toolkit to extract features to use as input into the regressor models in order to predict one time bin ahead and report the RMSE between the true observed values and the predicted values from our models. Table 2 reports RMSE calculated for 8 of the patients using each of the ensemble methods. In all models the maximum number of trees is set to 10. We can compare the depth of the bagging method trees: RandomForest being 32 ± 5 , ExtraTrees being 37 ± 4 , and Bagging being 33 ± 5 . In contrast, the boosting methods: AdaBoost and GradientBoosting, set trees of depth four by default. As expected, increasing the number of trees decreases the RMSE.

It is important to take computational time into consideration when choosing algorithms to make predictions in real time. From our experiments we found AdaBoost to give the best trade-off between RMSE and computation time.

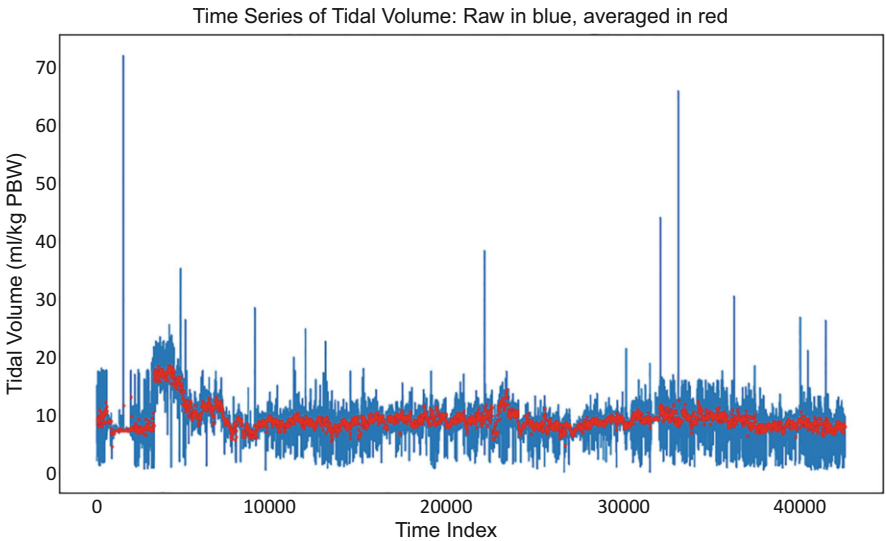


Fig. 4 Time series of a patient’s tidal volume profile. Raw minutely values in blue, 15-min averaged bins in red

Table 2 Comparison of regressor ensemble models performance for the prediction of patient’s tidal volume one time step ahead

Patient	No. data points	AdaBoost	RandomForest	Bagging	ExtraTrees	GradientBoosting
		RMSE	RMSE	RMSE	RMSE	RMSE
1	517	0.69	0.68	0.70	0.68	0.84
2	150	1.05	1.03	1.03	1.05	1.23
3	1358	0.38	0.38	0.38	0.36	0.39
4	40	1.05	0.99	1.02	1.00	0.95
5	162	0.34	0.32	0.32	0.34	0.41
6	178	1.38	1.28	1.35	1.39	1.32
7	1153	0.62	0.63	0.62	0.62	0.61
8	1245	0.83	0.84	0.84	0.84	0.91

One might expect that predicting further ahead in time would lead to larger RMSE values, however, the change in RMSE is small. We therefore explore using AdaBoost regression for the prediction of tidal volume values up to 1 h ahead, finding very little increase in RMSE values across patients.

As described in this chapter, a benefit of using ensemble models made from decision trees is that we can visualize the features and decisions used to make decisions. Figure 5 is one of the decision trees created when using the AdaBoost method.

The features were extracted by *tsfresh* as the most significant features for this problem. It is interesting here to discuss what these features can mean for our

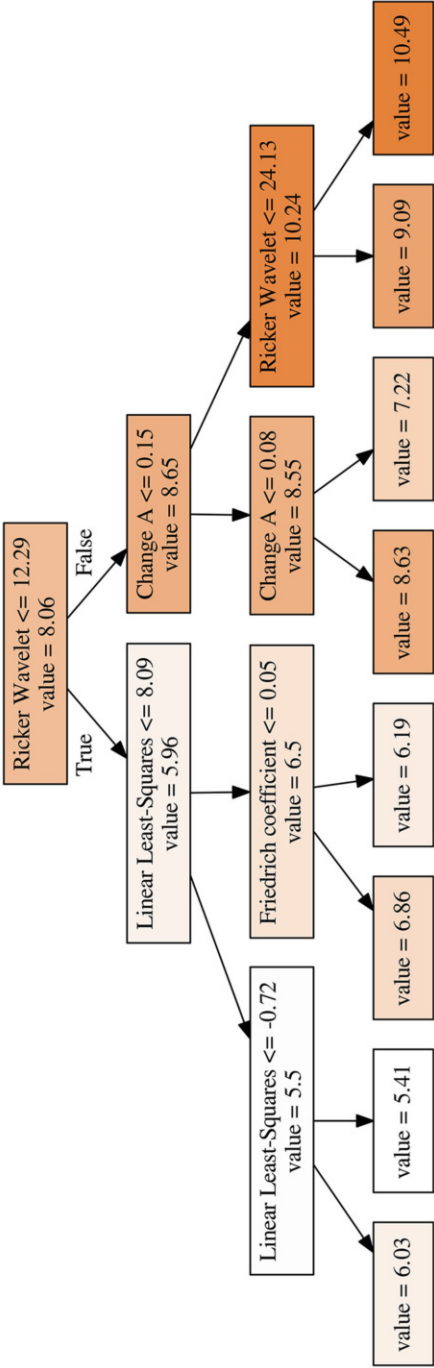


Fig. 5 One of the ten trees created by the AdaBoost method for patient 1 predictions

problem domain. The Ricker Wavelet is used to describe properties of a viscoelastic homogeneous media and the Friedrich coefficient aims at describing the random movement of a particle in a fluid. We can hypothesize from this finding that the amount of fluid in the lungs would be an impacting factor in how a patient’s tidal volume can change over time.

A comparison is then made applying long short-term memory neural networks to the same problem: predicting tidal volume 1 h ahead. Two models are created: ModelA has one hidden layer and ModelB has three hidden layers, with a 20% dropout layer between the second and third layers to avoid overfitting. In contrast to the regressor models that work with features extracted from the time series, the LSTM models use the time series values directly; requiring 70% of the time series to train the models. Each layer in our LSTM models has 50 nodes and both models use 20 input points to predict 4 ahead.

The RMSE values for our LSTM models are significantly greater than the AdaBoost method for predicting 1 h ahead, so we deem AdaBoost the better method of the two for this problem.

The VILIAAlert system alerted when four consecutive bins were greater than the 8 ml/kg tidal volume threshold. These alert times were stored in the database. We can thus work out the accuracy of our models to predict these alerts, showcasing the possibility of preventing threshold breaches and preventing injury. Table 3 highlights the predictive accuracy of the AdaBoost model for the 8 patients. Total Alerts being the total number of alerts recorded by the VILIAAlert system, TP being the true positives: the alerts that would have been predicted by the model, and FN being false negatives: the alerts that would not have been predicted. The accuracy is then calculated using:

$$\text{Accuracy} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

For the 84 alerts that were generated for patient 1, 81 would have been predicted using our models and thus those threshold breaches could have been prevented. These results showcase how Machine Learning algorithms, when paired with big

Table 3 Prediction accuracy of alerts using AdaBoost

Patient	Total alerts	TP	FN	Accuracy
1	84	81	3	0.96
2	25	23	2	0.92
3	0	0	0	1.00
4	2	1	1	0.50
5	3	2	1	0.67
6	11	3	8	0.27
7	0	0	0	1.00
8	167	142	25	0.85

data, can provide value in preventing lung injury during mechanical ventilation of intensive care patients.

4.1 The *ATTITUDE* Study

The *ATTITUDE* study operated by Queen's University, Belfast, and funded by The Health Foundation UK, aims to develop a clinical decision support tool to improve weaning protocols commonly used in clinical practice, and further understand the barriers in implementing evidence-based care from these tools tested in a proof-of-concept study carried out at the Royal Victoria Hospital ICU, Belfast. Improving patient care, outcomes, and mortality by reducing the duration of weaning can lead to reduced hospital stays and costs, and this study aims to find out if the use of clinical decision support tools can improve the quality of critical care practices.

5 Future of ML and AI in ICU

The methodology discussed in this chapter can, and must be, explored with various and extensive types and volumes of data, to investigate more disease states and clinical conditions. This data is currently being recorded worldwide in what are known as Electronic Health Records and these hold valuable insights which must be utilized to improve healthcare going forward.

From the European Big Data Value Strategic Research and Innovation Agenda [1] we understand the importance of Big Data and utilizing it to benefit both the economy and society. By exploring the already available EHRs we can provide societal benefit by improving patient outcomes and saving lives, and further economic benefits of saving millions in hospital costs, through shorter lengths of stays and disease prediction, among others. Healthcare, being one of the most important and largest sectors, can greatly impact the agenda of a data-driven economy across Europe. Data-driven predictive analytics, built using the methodologies discussed in the chapter, can produce clinical decision support tools, allowing for advanced decision making or automation of procedures. The Machine Learning methodologies can further provide greater insights into patient states and inform healthcare professionals with new information or possible reasoning that would not have been caught by a human. These new insights result in further research questions that can be explored.

This chapter can be aligned with the Big Data Reference model in various ways. The data recorded in ICU is extensive and in various different formats, from structured data and time series to imaging and text inputs. Advanced visualization tools can be used to add value to data by presenting it in user-friendly ways. Data analytics and Machine Learning can be applied for prediction and reasoning of

disease states. There exist further protocols and guidelines for the handling of patient information, ensuring efficient data protection and management.

The new AI, data and robotics partnership highlights the value opportunities that exist when transforming the healthcare sector by applying AI to produce value-based and patient-centric care in areas such as pandemic response, disease prevention, diagnosis decision support, and treatment [2].

High-performance computing is an integral part in deploying real-time predictive analytic models in intensive care. We must ensure our machines can process the data efficiently and quickly. Utilizing parallelism will provide speed up data processing and model predictions.

This chapter has explored the use of Big Data-driven analytics for acute and chronic clinical conditions to provide value to healthcare services. While there exists vast research carried out in certain disease states, such as Sepsis [30, 40–42], work is needed to provide greater in-depth analysis and insights into patients' complex physiologic state while in such critical conditions. Recent developments have led to a greater acceptance and excitement in this field, resulting in updated guidelines for testing AI models in real-life clinical trials to promote worldwide acceptance of the use of AI in healthcare.

Acknowledgments Rachael Hagan acknowledges funding for a PhD studentship from the Department for the Economy, Northern Ireland, and Dr. Murali Shyamsundar is an NIHR Clinical Scientist fellow.

References

1. Zillner, S., Curry, E., Metzger, A., et al. (2017). *European big data value strategic research & innovation agenda*. Big Data Value Association.
2. Zillner, S., Bisset, D., Milano, M., Curry, E., García Robles, A., Hahn, T., Irgens, M., Lafrenz, R., Liepert, B., O'Sullivan, B., & Smeulders, A. (Eds.). (2020). *Strategic research, innovation and deployment agenda – AI, data and robotics partnership*. Third release. Brussels. BDVA, euRobotics, ELLIS, EurAI and CLAIRE.
3. Ridley, S., & Morris, S. (2007). Cost effectiveness of adult intensive care in the UK. *Anaesthesia*, 62, 547–554. <https://doi.org/10.1111/j.1365-2044.2007.04997.x>
4. Teno, J. M., Fisher, E., & Hamel, B. (2000). Decision-making and outcomes of prolonged ICU stays in seriously ill patients. *Journal of the American Geriatrics Society*, 48, 70–74.
5. Bates, D. W., Saria, S., Ohno-Machado, L., et al. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, 33, 1123–1131. <https://doi.org/10.1377/hlthaff.2014.0041>
6. Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe* (pp. 1–303). <https://doi.org/10.1007/978-3-319-21569-3>
7. Bhutkar, G., Deshmukh, S., & Detection, D. (2015). Vital medical devices in intensive care unit. <https://doi.org/10.13140/RG.2.1.4671.6247>
8. Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time series FeatuRe extraction on basis of scalable hypothesis tests (tsfresh – A Python package). *Neurocomputing*, 307, 72–77. <https://doi.org/10.1016/j.neucom.2018.03.067>

9. Drews, F. A. (2008). Patient monitors in critical care: Lessons for improvement. In *Advances in patient safety: New directions and alternative approaches* (pp. 1–13). Rockville, MD: Agency for Healthcare Research and Quality. <https://doi.org/NBK43684> [bookaccession].
10. Johnson, A., Pollard, T., Shen, L., et al. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*.
11. Goldberger, A. L., Amaral, L. G., et al. (2000). Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101, 215–220.
12. Fayyad, U., Piatetsky-shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *American Association for Artificial Intelligence*, 17, 37–54.
13. Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing bias in artificial intelligence in health care. *JAMA*, 322, 2377–2378. <https://doi.org/10.1001/jama.2019.18058>
14. Hammer, G. P., Du Prel, J. B., & Blettner, M. (2009). Avoiding Bias in observational studies. *Dtsch Arztebl*, 106, 664–668. <https://doi.org/10.3238/arztebl.2009.0664>
15. Bridi, A. C., Louro, T. Q., & Da Silva, R. C. L. (2014). Clinical alarms in intensive care: Implications of alarm fatigue for the safety of patients. *Revista Latino-Americana de Enfermagem*, 22, 1034–1040. <https://doi.org/10.1590/0104-1169.3488.2513>
16. Thangavelu, S., Yunus, J., Ifeakor, E., et al. (2015). Responding to clinical alarms: A challenge to users in ICU/CCU. In *Proceedings – International Conference on Intelligent Systems, Modelling and Simulation, ISMS* (pp. 88–92).
17. Ancker, J. S., Edwards, A., Nosal, S., et al. (2017). Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Medical Informatics and Decision Making*, 17, 1–9. <https://doi.org/10.1186/s12911-017-0430-8>
18. Wicks, P., Liu, X., & Denniston, A. K. (2020). Going on up to the SPIRIT in AI: Will new reporting guidelines for clinical trials of AI interventions improve their rigour? *BMC Medicine*, 18, 4–6. <https://doi.org/10.1186/s12916-020-01754-z>
19. Statement, C. (2020). The SPIRIT-AI extension. 26. <https://doi.org/10.1038/s41591-020-1037-7>
20. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. 26.
21. Sounderajah, V., Ashrafian, H., Aggarwal, R., et al. (2020). Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI steering group. *Nature Medicine*, 26, 807. <https://doi.org/10.1038/s41591-020-0941-1>
22. Collins, G. S., & Moons, K. G. M. (2019). Reporting of artificial intelligence prediction models. *Lancet*, 393, 1577–1579. [https://doi.org/10.1016/S0140-6736\(19\)30037-6](https://doi.org/10.1016/S0140-6736(19)30037-6)
23. Fleuren, L. M., Thorat, P., Shillan, D., et al. (2020). Machine learning in intensive care medicine: Ready for take – off? *Intensive Care Medicine*, 46, 1486–1488. <https://doi.org/10.1007/s00134-020-06045-y>
24. Deo, R. C. (2015). Basic science for clinicians. *Circulation*, 132, 1920–1930. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>
25. Turek, M. (2020). *DARPA Project Explainable Artificial Intelligence (XAI)*. Accessed November 5, 2020, from <https://www.darpa.mil/program/explainable-artificial-intelligence%0A>
26. Doran, C., Does, T. R. W., & Ai, E. (2018). What does explainable AI really mean? A new conceptualization of perspectives. In *CEUR Workshop Proceedings*, 2071.
27. Bannach-Brown, A., Przybyła, P., Thomas, J., et al. (2019). Machine learning algorithms for systematic review: Reducing workload in a preclinical review of animal studies and reducing human screening error. *Systematic Reviews*, 8, 1–12. <https://doi.org/10.1186/s13643-019-0942-7>
28. Thommandram, A., Pugh, J. E., Eklund, J. M., et al. (2013). Classifying neonatal spells using real-time temporal analysis of physiological data streams: Algorithm development. In *IEEE EMBS Special Topics Conferenc on Point-of-Care Healthcare Technologies Synerg Toward Better Glob Heal PHT 2013* (pp. 240–243). <https://doi.org/10.1109/PHT.2013.6461329>
29. Blount, M., Ebling, M., Eklund, J. M., et al. (2010). Real-time analysis for intensive care. *IEEE Engineering in Medicine and Biology Magazine*, 29(2), 110–118.

30. Nemati, S., Holder, A., Razmi, F., et al. (2018). An interpretable machine learning model for accurate prediction of Sepsis in the ICU. *Critical Care Medicine*, 46, 547–553. <https://doi.org/10.1097/CCM.0000000000002936>
31. Kubat, M. (2017). *An introduction to machine learning* (2nd ed.). Springer.
32. Ambrosiadou, B. V., Goulis, D. G., & Pappasa, C. (1996). Clinical evaluation of the DIABETES expert system for decision support by multiple regimen insulin dose adjustment. *Computer Methods and Programs in Biomedicine*, 49, 105–115.
33. Sayadi, M., Zibaenezhad, M., Mohammad, S., & Ayatollahi, T. (2017). Simple prediction of type 2 diabetes mellitus via decision tree modeling. *International Cardiovascular Research Journal*, 11, 71–76.
34. Kuo, H. J., Chiu, H. W., Lee, C. N., et al. (2015). Improvement in the prediction of ventilator weaning outcomes by an artificial neural network in a medical ICU. *Respiratory Care*, 60, 1560–1569. <https://doi.org/10.4187/respcare.03648>
35. Tomašev, N., Glorot, X., Rae, J. W., et al. (2019). A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*, 572, 116–119. <https://doi.org/10.1038/s41586-019-1390-1.A>
36. Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., et al. (2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction. *Lancet*, 394, 861–867. [https://doi.org/10.1016/S0140-6736\(19\)31721-0](https://doi.org/10.1016/S0140-6736(19)31721-0)
37. Montavon, G., Samek, W., & Müller, K. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73, 1–15. <https://doi.org/10.1016/j.dsp.2017.10.011>
38. Fritz, B. A., Cui, Z., Zhang, M., et al. (2019). Deep-learning model for predicting 30-day postoperative mortality. *British Journal of Anaesthesia*, 123, 688–695. <https://doi.org/10.1016/j.bja.2019.07.025>
39. Higgins, T., Freese-man-freeman, L., & Henson, K. (2019). Mews ++: Predicting clinical deterioration in admitted patients using a novel inpatient deterioration: Can artificial intelligence predict who will be transferred to the ICU? 48:2019.
40. Li, X., Xu, X., Xie, F., et al. (2020). A time-phased machine learning model for real-time prediction of sepsis in critical care. *Critical Care Medicine*, E884–E888. <https://doi.org/10.1097/CCM.0000000000004494>
41. Ginestra, J. C., Giannini, H. M., Schweickert, W. D., et al. (2019). Clinician perception of a machine learning-based early warning system designed to predict severe Sepsis and septic shock. *Critical Care Medicine*, 47, 1477–1484. <https://doi.org/10.1097/CCM.0000000000003803>
42. Komorowski, M., Celi, L. A., Badawi, O., & Gordon, A. C. (2018). The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine*, 24. <https://doi.org/10.1038/s41591-018-0213-5>
43. Eslami, S., de Keizer, N. F., Abu-Hanna, A., et al. (2009). Effect of a clinical decision support system on adherence to a lower tidal volume mechanical ventilation strategy. *Journal of Critical Care*, 24, 523–529. <https://doi.org/10.1016/j.jcrc.2008.11.006>
44. Jalali, A., Bender, D., Rehman, M., et al. (2016). Advanced analytics for outcome prediction in intensive care units. In *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS* (pp. 2520–2524). <https://doi.org/10.1109/EMBC.2016.7591243>
45. Che, Z., Purushotham, S., Khemani, R., & Liu, Y. (2016). Interpretable deep models for ICU outcome prediction. In *AMIA. Annual Symposium proceedings AMIA Symposium 2016* (pp. 371–380).
46. Boverman, G., & Genc, S. (2014). Prediction of mortality from respiratory distress among long-term mechanically ventilated patients. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC 2014* (pp. 3464–3467). <https://doi.org/10.1109/EMBC.2014.6944368>

47. Emanet, N., Öz, H. R., Bayram, N., & Delen, D. (2014). A comparative analysis of machine learning methods for classification type decision problems in healthcare. *Decision Analysis*, 1, 1–20. <https://doi.org/10.1186/2193-8636-1-6>
48. Florence, S., Amma, N. G. B., Annappoorani, G., & Malathi, K. (2014). Predicting the risk of heart attacks using neural network and decision tree. *International Journal of Innovative Research in Computer and Communication Engineering*, 2, 7025–7030.
49. Parreco, J., Hidalgo, A., Parks, J. J., et al. (2018). Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement. *The Journal of Surgical Research*, 228, 179–187. <https://doi.org/10.1016/j.jss.2018.03.028>
50. Hyland, S. L., Faltys, M., Hüser, M., et al. (2019). Machine learning for early prediction of circulatory failure in the intensive care unit. *Nature Medicine*, 26. <https://doi.org/10.1038/s41591-020-0789-4>
51. Meyer, A., Zverinski, D., Pfahringer, B., et al. (2018). Machine learning for real-time prediction of complications in critical care: A retrospective study. *The Lancet Respiratory Medicine*, 6, 905–914. [https://doi.org/10.1016/S2213-2600\(18\)30300-X](https://doi.org/10.1016/S2213-2600(18)30300-X)
52. Harrision, D. (2014). ICNARC. Number of mechanically ventilated patients during 2012.
53. Hagan, R., Gillan, C. J., Spence, I., et al. (2020). Comparing regression and neural network techniques for personalized predictive analytics to promote lung protective ventilation in intensive care units. *Computers in Biology and Medicine*, 126, 104030. <https://doi.org/10.1016/j.compbimed.2020.104030>
54. Gillan, C. J., Novakovic, A., Marshall, A. H., et al. (2018). Expediting assessments of database performance for streams of respiratory parameters. *Computers in Biology and Medicine*, 100, 186–195. <https://doi.org/10.1016/j.compbimed.2018.05.028>
55. Viceconti, M., & Hunter, P. (2016). The virtual physiological human: Ten years after. *Annual Review of Biomedical*.
56. Das, A., Saffaran, S., Chikhani, M., et al. (2020). In silico modeling of coronavirus disease 2019 acute respiratory distress syndrome: Pathophysiologic insights and potential management implications. *Critical Care Explorations*, 2, e0202. <https://doi.org/10.1097/ccx.0000000000000202>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

