Replicating human expertise of mechanical ventilation waveform analysis in detecting patient-ventilator cycling asynchrony using machine learning

Behnood Gholamia,*, Timothy S. Phan, Wassim M. Haddad, Andrew Cason, Jerry Mullis, Levi Price, James M. Bailey

Autonomous Healthcare, Inc., Hoboken, NJ, USA
School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA
Northeast Georgia Medical Center, Gainesville, GA, USA

ARTICLE INFO

Keywords:
Patient-ventilator asynchrony
Machine learning
Waveform analysis
Mechanical ventilation
Clinical decision support

ABSTRACT

Background: — Acute respiratory failure is one of the most common problems encountered in intensive care units (ICU) and mechanical ventilation is the mainstay of supportive therapy for such patients. A mismatch between ventilator delivery and patient demand is referred to as patient-ventilator asynchrony (PVA). An important hurdle in addressing PVA is the lack of a reliable framework for continuously and automatically monitoring the patient and detecting various types of PVA.

Methods: — The problem of replicating human expertise of waveform analysis for detecting cycling asynchrony (i.e., delayed termination, premature termination, or none) was investigated in a pilot study involving 11 patients in the ICU under invasive mechanical ventilation. A machine learning framework is used to detect cycling asynchrony based on waveform analysis.

Results: — A panel of five experts with experience in PVA evaluated a total of 1377 breath cycles from 11 mechanically ventilated critical care patients. The majority vote was used to label each breath cycle according to cycling asynchrony type. The proposed framework accurately detected the presence or absence of cycling asynchrony with sensitivity (specificity) of 89% (99%), 94% (98%), and 97% (93%) for delayed termination, premature termination, and no cycling asynchrony, respectively. The system showed strong agreement with human experts as reflected by the kappa coefficients of 0.90, 0.91, and 0.90 for delayed termination, premature termination, and no cycling asynchrony, respectively.

Conclusions: — The pilot study establishes the feasibility of using a machine learning framework to provide waveform analysis equivalent to an expert human.

1. Introduction

Clinical decision support systems are taking an ever-increasing role in the practice of healthcare. While providing clinical decision support using structured information has been widely investigated in the literature and is being translated to clinical practice [1], the next frontier is to address challenges in providing clinical decision support based on unstructured data (e.g., images and waveforms). Recent advances in machine learning have provided the opportunity to assist clinicians analyze unstructured data such as free text [2] and images [3]. One area that would benefit from automated analysis of waveforms is respiratory management for patients in the intensive care unit (ICU).

Acute respiratory failure due to infection, trauma, and major surgery is one of the most common problems encountered in the ICU and mechanical ventilation is the mainstay of supportive therapy for such patients. Patient-ventilator interaction is a critical component of the mechanical ventilation process. Specifically, the ventilator has to respond to a patient’s respiratory demand including triggering (i.e., initiation of inspiration) and cycling of breaths (i.e., transition from inspiration to expiration). A mismatch between ventilator delivery and patient demand is referred to as patient-ventilator asynchrony (PVA) [4].

Mechanical ventilation may be a distressing and noxious, albeit necessary, procedure for the critically ill patient. This is often manifested by PVA, which in extreme cases is referred to as “fighting the ventilator.” The incidence of PVA is high, with estimates ranging from 12 to 43%, and has been associated with failure to wean from ventilation, longer

* Corresponding author.
E-mail address: bgholami@autonomoushealthcare.com (B. Gholami).

https://doi.org/10.1016/j.compbiomed.2018.04.016
Received 26 January 2018; Received in revised form 2 April 2018; Accepted 21 April 2018
0010-4825/© 2018 Elsevier Ltd. All rights reserved.
duration of ventilation, and longer length of stay in critical care units [5-7]. While the association between PVA and poor patient outcome has been recognized, causality has not been established. However, it seems self-evident that the clinical provider should attempt to reduce PVA due to the ethical imperative of reducing patient distress.

The most common approach to correcting patient-ventilator asynchrony is to increase sedation. The authors in Ref. [8] observed that to address severe breath-stacking, clinical staff increase sedation and/or analgesia in half the cases. However, increased sedation is associated with increased duration of ventilation, increased length of stay, increased incidence of delirium, and increased mortality. Specifically, based on the results of a randomized controlled trial, reduction in sedation use resulted in a reduction of mechanical ventilation and ICU length of stay [9].

Larger multi-center studies confirmed that daily interruption of sedation in conjunction with spontaneous breathing trials reduced length of stay and significantly reduced mortality [10]. In addition, a randomized clinical trial indicated that interruption of sedation in conjunction with physical and occupational therapy reduces the duration of delirium [11]. See Ref. [12] for a comprehensive discussion and review of the literature related to ICU sedation and delirium.

An alternative approach to reducing PVA is to adjust the ventilator mode or settings to match patient demand and respiratory cycles. The utility of this approach has been demonstrated in Ref. [8]. Specifically, in this study the authors observed that changes to ventilation settings was a more effective intervention for breath stacking asynchrony as compared to increase in sedation/analgesia.

While there is variability in nomenclature, PVA may be generally classified as trigger asynchrony, flow asynchrony, and cycling asynchrony [13]. In some cases, the recognition of PVA is obvious; however, in other cases both recognition and classification may be subtle, and in the clinical setting, recognition of asynchrony may be delayed. More accurate methods to detect asynchrony involves the addition of invasive components to measure esophageal pressure or diaphragm electrical activity [4]. The additional complexity has prohibited their widespread adoption, and hence, there is a need for a reliable non-invasive method to detect asynchrony.

Most prior studies of PVA have been conducted with off-line post hoc analysis of pressure and flow versus time recordings. Clinicians with knowledge of asynchrony rely on waveforms to detect subtle cases of PVA, which generally cannot be detected by only observing the patient without considering the ventilator waveforms. However, very few institutions can support one-on-one respiratory therapist deployment, and often nursing staff are not familiar with the analysis of pressure and flow versus time recordings needed to recognize asynchrony in the absence of frank physical manifestations of a patient “fighting the ventilator.”

If PVA can be reliably detected via a non-invasive and automated algorithm, then this information can be used within a clinical decision support system for respiration management. This system can show detected asynchrony events and their trends, and assist respiratory therapists in timely detection of asynchrony. It is expected that timely detection of asynchrony and its type will assist clinicians to identify the source of asynchrony and address it through adjustments of the ventilator settings and mode before increasing the patient’s sedation/analgesia.

Given these limitations for real-time monitoring by dedicated clinical staff, in this paper we investigate the feasibility of an automated waveform analysis algorithm that can replicate human expertise in detecting patient-ventilator cycling asynchrony. Our hypothesis is that the expertise of a human expert (or a panel of human experts) in analyzing mechanical ventilation waveforms can be replicated using a machine learning algorithm.

Rule-based as well as statistical asynchrony detection techniques have been developed for detecting ineffective triggering and double triggering [14-17]. However, the feasibility of machine learning-based systems to detect cycling asynchrony has not been investigated. Discussions on cycling asynchrony are generally qualitative and involve multiple features that may be found on pressure and flow versus time recordings. Hence, detection of cycling asynchrony provides an opportunity to “capture” human expertise that is presented qualitatively in the clinical literature. A summary of the proposed approach to replicate human expertise in detecting patient-ventilator cycling asynchrony using machine learning and evaluating its performance is shown in Fig. 1.

2. Materials and methods

2.1. Subjects

We analyzed mechanical ventilation waveform data from 11 patients admitted to the intensive care unit at the Northeast Georgia Medical Center, Gainesville, GA. This was a retrospective study, where collected waveform data did not include any patient identifiers, and hence, was exempt from IRB review. Our inclusion criteria for collection of data was that the patients were undergoing invasive ventilation in the pressure controlled-volume guaranteed (PCV-VG) mode of the GE Engstrom Carestation ventilator. We excluded children (patients under the age of 18), patients with tracheostomy, and those who received paralytic agents. Collected data did not include patient demographics and diagnosis.

2.2. Data collection

All patient in this study were ventilated using a GE Engstrom Carestation connected to a GE EView, a module used for collecting ventilator data at a sampling rate of 25 samples per second. Recorded data segments, which were at least 12 h long, were downloaded from the EView to a laptop computer. Custom software was used to visualize the pressure, flow, and volume versus time recordings for further review and annotation by the human experts.

2.3. Asynchrony Type

In this study, we consider detecting the presence or absence of cycling asynchrony, that is, delayed termination, premature termination, or no cycling asynchrony. In delayed termination, the end of respiratory valve opens after the patient has already initiated exhalation. A

![Fig. 1. Summary of the approach to replicate and evaluate the human expertise in detecting cycling asynchrony.](https://example.com/fig1.png)
consequence of delayed termination is air trapping and ineffective triggering as a result of insufficient expiratory time and excessive tidal volume [13]. Premature termination occurs when the end of mechanical inspiration precedes the end of the patient’s neural inspiration, where the expiratory valve prematurely opens before the patient stops inhaling. Premature termination subjects the patient to increased risk of double triggering. In the case of a shortened expiration resulting in a double trigger, air trapping and auto-PEEP may occur, which may inhibit the ability of the patient to reach subsequent trigger thresholds [18].

2.4. Waveform Review by Human Experts

To investigate the accuracy of the proposed machine learning algorithm to reproduce expert human assessment for detecting cycling asynchrony using waveforms, data from 11 patients was used to compile 4–5 sets of 17–50 consecutive breath cycle segments in areas where there was some evidence of asynchrony. Specifically, a total of 125–126 breath cycles for each of the 11 patients were presented to the human experts. The waveforms used for analysis included flow, pressure, and volume versus time recordings. Segments of waveforms used for further review by the panel of experts were selected by 2 experts who reviewed the original pressure and flow versus time recordings to identify areas with evidence of cycling asynchrony. A panel of 5 human experts with knowledge of asynchrony reviewed a total of 1377 breaths to detect presence of cycling asynchrony (i.e., delayed termination, premature termination, or no cycling asynchrony) if at least 3 human experts agreed on a specific class label. Cases with no majority vote or cases with “unknown” label assigned by a majority of human reviewers were excluded from further analysis. A breakdown of breath cycles included and excluded in the analysis is given in Table 1.

Although three subjects were the source of 128 (74%) of the excluded breath cycles, the average yield for each patient (i.e., the number of included breath cycles in the final analysis divided by the number of breath cycles reviewed for each subject) was 87%, with the lowest yield of 63%. In addition, a perfectly balanced distribution of 1204 breath cycles over 11 subjects would be 9% per subject. In our analysis, breath cycles from each subject comprised 7–10% of the final number of breaths, which indicates that the overall set of breaths was not strongly imbalanced.

Before using a machine learning classification algorithm, relevant features need to be extracted from the waveforms. Once the features were extracted and the machine learning classification framework is chosen, a training set is used to identify the parameters of the classifier (i.e., the classifier is “trained” on data). Finally, the classifier is presented with unseen waveforms to evaluate its performance.

2.5.1. Feature Extraction

Cycling asynchrony can generally be detected by identifying certain
landmarks” on the flow and pressure versus time recordings. Specifically, in premature termination, the landmark includes a reversal of flow in the early part of the expiratory phase accompanied by a decrease in pressure. In delayed termination such landmarks include pressure “tenting” at the end of inspiration (sharp increase in pressure) [13] accompanied by decreased or reversal of flow. See Fig. 3 for sample breath cycles with premature termination and delayed termination asynchronies.

In order to detect simultaneous changes in the pressure and flow versus time recordings, we generated a derived waveform from pressure and flow versus time recordings referred to as the delta waveform. The delta waveform is defined as the difference between the normalized pressure and flow, where normalization of pressure and flow were performed by value of pressure and flow at peak flow, respectively. We extracted a series of features from the delta signal including the valley depth in the expiratory phase denoted by \( a \), estimated inspiratory time (i.e., duration of time where pressure is above 50% of difference between set pressure and PEEP) denoted by \( b \), and the ratio of the time flow is negative (denoted by \( c \)) over the estimated inspiratory time, that is, the ratio \( \frac{c}{b} \) (see Fig. 4).

2.5.2. Multi-Class Classification

We used random forests [20] to build a classifier using the available extracted features and ground truth labels provided by a majority vote of human experts. Random forests are a machine learning framework which extend the concept of decision trees. Specifically, a training set is used to generate a collection of decision trees. In order to classify a new observation, each decision tree classifies (“votes”) the new observation into a class (e.g., delayed termination, premature termination, or no cycling asynchrony). The class with the majority vote is selected as the most likely class. The proposed multi-class machine learning algorithm learns the complex mapping between morphological features of the delta waveforms (along with the pressure and flow versus time recordings) and the type of cycling asynchrony, given a set of ground truth labels.

If a larger training set is used, which requires manual labeling of a larger number of breaths, then the machine learning algorithm can automatically refine its learned mapping between features and classifications such that addition of more data can be used to further improve its performance. The ability of our machine learning-based system to improve performance with provision of additional data is an advantage over hard-coded, rule-based expert systems.

2.5.3. Model Evaluation

In order to assess the agreement of the machine learning algorithm with the labels determined by a majority vote of the human experts, we used a 10-fold cross validation framework [21]. Specifically, data (i.e., breath cycles and their corresponding class labels) was randomly assigned to 10 mutually exclusive sets, where 9 of the sets were used for training the machine learning algorithm and the remaining unseen set was used for testing. Next, the resulting class labels provided by the machine learning algorithm were compared to class labels provided by a majority vote of the human expert panel. This process was repeated 10 times until all sets were tested.

Classifier performance to detect cycling asynchrony was quantified using sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) to detect delayed termination, premature termination, or no cycling asynchrony. In addition, in order to quantify the agreement between the human expert labels and the machine learning algorithm, we used the kappa coefficient [22]; a statistical metric used to quantify the agreement between two raters classifying the same observation.

### Table 1

<table>
<thead>
<tr>
<th>Breath Cycles</th>
<th>Total</th>
<th>No Majority Vote</th>
<th>Low Volume</th>
<th>Double Triggering</th>
<th>Missing Features</th>
<th>Majority Voted as Unknown</th>
<th>Used for Analysis by Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1377</td>
<td>46</td>
<td>90</td>
<td>18</td>
<td>12</td>
<td>7</td>
<td>1204</td>
</tr>
</tbody>
</table>

Fig. 3. Left: Pressure and flow versus time recordings of a breath cycle with premature termination asynchrony. The reversal of flow in the early part of expiration is indicated by the arrow. Right: Pressure and flow waveforms of a breath cycle with delayed termination asynchrony. Pressure tenting and negative flow during inspiration are indicated by the arrow.
3. Results

A breakdown of the classification provided by the machine learning classifier compared to human expert labels is given by the confusion matrix in Table 2. Of the 1204 breath cycles with labels determined by majority vote of the human expert reviewers, 802 breath cycles (67%) were majority-voted to have neither of the cycling asynchronies considered, 232 (19%) were voted as delayed termination, and 170 (14%) were voted as premature termination. The performance and the agreement between the machine learning algorithm and human experts are summarized in Table 3.

For each fold of the 10-fold cross validation, a confusion matrix was produced, where row $i$ and column $j$ indicate the number of breaths where the majority of human experts chose asynchrony class $i$ and the algorithm predicted asynchrony class $j$, respectively. The diagonal of the confusion matrix provides the number of breaths with correct predictions in which the algorithm prediction matched the human expert classification. The confusion matrix for each fold was used to calculate the sensitivity as $TP/(TP + FN)$, specificity as $TN/(TN + FP)$, positive predictive value as $TP/(TP + FP)$, and negative predictive value as $TN/(TN + FN)$, where $TP$, $TN$, $FP$, and $FN$ represent the number of true positives, true negatives, false positives, and false negatives, respectively. The kappa coefficient was computed by comparing the labels provided by the majority vote of the human experts and the algorithm prediction. These metrics were then averaged across the 10 folds to obtain the values presented in Table 3.

The classifier can accurately detect the presence or absence of asynchrony and its type as reflected by high sensitivity, specificity, positive predictive value, and negative predictive value. Both sensitivity and specificity were high for all three classes of cycling asynchrony or no cycling asynchrony, ranging from 89% sensitivity for premature termination to 97% sensitivity for neither, and 93% specificity for neither to 99% specificity for premature termination. In addition, the machine learning algorithm showed a strong agreement with human observers as reflected by the high kappa coefficients ranging from 0.90 to 0.91 [23].

4. Discussion

Computer algorithms to detect patient-ventilator asynchrony have been mainly focused on the detection of ineffective efforts based on a series of features extracted from the flow and pressure waveforms [14–17]. Specifically, in Ref. [15] the ability of the proposed algorithm to detect ineffective efforts was compared to assessments provided by a panel of five human experts. Furthermore, the agreement of the algorithm to detect ineffective triggering was compared to a detection based on electrical activity of the diaphragm. The authors in Ref. [24] present a framework for detecting general asynchrony without the ability to detect the specific type of asynchrony using spectral analysis of the flow waveforms. In Ref. [25], a method is proposed to generate a signal reflective of a patient’s respiratory muscle pressure output. The application of this method to detect delay in cycling and ineffective efforts was investigated.

The authors in Ref. [19] present a framework to detect various types of asynchrony. Specifically, a framework is presented to detect cycling asynchrony only in pressure support ventilation mode by comparing the inspiration time with the mean inspiratory time in the previous 20 breath cycles. However, as the authors note, this approach has several limitations. First, no validation of the approach was performed to show that the detected cases of premature termination or delayed termination are in agreement with a clinician’s assessment or other forms of assessments. Furthermore, waveform “landmarks,” which human experts rely on to detect asynchrony, are not used. In addition, if the previous breath cycles used to calculate the mean inspiratory time are all asynchronous, then the mean inspiratory time is no longer an unbiased point of reference, which in turn will result in an unreliable asynchrony detection. Finally, this approach is developed for pressure support ventilation and cannot be used in other modes of ventilation including pressure control ventilation (which is investigated in this paper).

### Table 2
Confusion matrix used to compare machine learning prediction with expert human assessment.

<table>
<thead>
<tr>
<th>Expert Consensus</th>
<th>Predicted by Algorithm</th>
<th>None</th>
<th>Delayed Termination</th>
<th>Premature Termination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither (n = 802)</td>
<td>777</td>
<td>20</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Delayed Termination (n = 232)</td>
<td>12</td>
<td>219</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Premature Termination (n = 170)</td>
<td>17</td>
<td>2</td>
<td>151</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3
Classifier performance statistics and agreement with human experts.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value (PPV)</th>
<th>Negative Predictive Value (NPV)</th>
<th>Kappa Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither</td>
<td>0.97</td>
<td>0.93</td>
<td>0.96</td>
<td>0.94</td>
<td>0.90</td>
</tr>
<tr>
<td>Delayed Termination</td>
<td>0.94</td>
<td>0.98</td>
<td>0.91</td>
<td>0.99</td>
<td>0.90</td>
</tr>
<tr>
<td>Premature Termination</td>
<td>0.89</td>
<td>0.99</td>
<td>0.96</td>
<td>0.98</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Fig. 4. Representation of a number of extracted features used in the machine learning algorithm.
The ability of a machine learning framework to replicate human expertise to detect cycling asynchrony has not been previously investigated. Qualitatively we found that the machine learning system performed excellently in classifying cycling asynchronies consistent with the human experts. The results establish the feasibility of replicating human expertise in waveform analysis in detecting cycling asynchrony.

Our study has several limitations. First, we only investigated the cycling asynchronies, that is, premature termination and delayed termination, and did not include other asynchronies, such as ineffective triggering or breath stacking in the analysis. These other asynchronies could certainly be more clinically significant. However, their study will require a larger data set to ensure that an adequate number of asynchronous breaths for each type of asynchrony is included in the training set.

A second limitation of this study is that we did not have independent training and testing data sets, that is, training and testing sets were extracted from the same set of collected data. This also reflects the relatively limited scope of this initial investigation. However, we do note that the 10-fold cross validation framework we employed is well established and widely accepted in the statistics and machine learning literature. Specifically, the waveforms in the testing set are not included in the training set, and hence, the machine learning algorithm had not previously “seen” the data during training. We will investigate the performance of the machine learning algorithm with independent training and testing sets in a future study involving a larger number of subjects.

A third limitation of this study is that all patients were being ventilated in the pressure controlled-volume guaranteed mode of the GE Engstrom Carestation ventilator. Thus, our results could reflect this particular mode of ventilation and extrapolation to other modes should only be done with this limitation in mind. A forth limitation of the study is that breaths with substantially different morphology were not analyzed to detect cycling asynchrony. Specifically, the proposed machine learning algorithm relies on certain “landmarks” and features that may not be present in irregular breaths (i.e., breaths with substantially different morphology). In fact, if breath morphology is severely distorted, detecting certain landmarks associated with cycling asynchrony becomes highly challenging. As a result, irregular breaths were excluded from the analysis.

Finally, this study involved replicating the expertise of a panel of human expert reviewers to detect cycling asynchrony. Hence, the accuracy of the machine learning algorithm to detect true asynchronies is limited to the available training data as human expert reviewers can potentially provide incorrect labels. As investigated in a prior study, intensive care unit physicians showed a limited ability to detect patient-ventilator asynchrony using mechanical ventilation waveforms [26]. Here, we used the majority vote of the human expert panel to label breaths. There were cases were 3 or more experts identified an asynchrony type which was different than that chosen by other experts. It has been shown that the random forests algorithm is robust to “noisy” labels, that is, incorrect labels included in the training set [27]. The performance of a machine learning classifier trained on “ground truth” asynchrony labels generated by invasive measures of asynchrony (e.g., esophageal pressure) rather than asynchrony labels generated based on expert human knowledge needs to be investigated in a future study.

There are many important, and to date, unanswered questions related to PVA and the impact of this phenomenon on clinical outcome. While the association between PVA and poor patient outcome has been recognized, causality has not been established. It is quite plausible that asynchrony simply reflects more severe lung injury and that the underlying lung injury is the cause of poorer outcomes, not the asynchrony itself. And if there is a causative relationship between PVA, then the mechanism of this relationship is not known. It could be a result of increased sedation in response to the perception of asynchrony, a result of respiratory muscle fatigue due to excess work of breathing, or a result of excessive tidal volumes due to breath stacking or double triggering. Other mechanisms could certainly be postulated.

It is also unclear whether the association between asynchrony and poorer outcomes are true for all types of asynchrony. And it is also unclear how often asynchrony, other than breath stacking, can be corrected with adjustment of ventilator mode or settings. To answer these questions and to better understand the phenomenon, large data sets will need to be collected. However, as an inevitable consequence of limited human resources, most observations in critical care are not continuous in time but are rather “snapshots in time.” The development of machine learning-based systems to continuously monitor patients, as exemplified in this study, is one means to provide clinical decision support for the assistance of the bedside provider.

**Ethical Approval**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For this type of study formal consent is not required.

**Summary**

Acute respiratory failure due to infection, trauma, and major surgery is one of the most common problems encountered in the ICU and mechanical ventilation is the mainstay of supportive therapy for such patients. Patient-ventilator interaction is a critical component of the mechanical ventilation process. Specifically, the ventilator has to respond to a patient’s respiratory demand including triggering (i.e., initiation of inspiration) and cycling of breaths (i.e., transition from inspiration to expiration). A mismatch between ventilator delivery and patient demand is referred to as patient-ventilator asynchrony (PVA).

Clinicians with knowledge of asynchrony rely on waveforms to detect subtle cases of PVA. However, very few institutions can support one-on-one respiratory therapist deployment, and often nursing staff are not familiar with the analysis of pressure and flow versus time recordings needed to recognize asynchrony in the absence of frank physical manifestations of a patient “fighting the ventilator.”

If PVA can be reliably detected via a non-invasive and automated algorithm, then this information can be used within a clinical decision support system for respiration management. This system can show detected asynchrony events and their trends, and assist respiratory therapists in timely detection of asynchrony. It is expected that timely detection of asynchrony and its type will assist clinicians to identify the source of asynchrony and address it through adjustments of the ventilator settings and mode before increasing the patient’s sedation-anaesthesia.

Given these limitations for real-time monitoring by dedicated clinical staff, in this paper we investigate the feasibility of an automated waveform analysis algorithm that can replicate human expertise in detecting patient-ventilator cycling asynchrony. Our hypothesis is that the expertise of a human expert (or a panel of human experts) in analyzing mechanical ventilation waveforms can be replicated using a machine learning algorithm.

The waveforms (pressure, flow, and volume as a function of time) from 11 mechanically ventilated patients were independently analyzed by 5 different expert observers, who were asked to identify cycling asynchronies, and asynchrony labels were only valid if there was agreement among at least 3 of the 5 observers. In order to assess the agreement of machine learning algorithm with the labels determined by a majority vote of the human experts, we used a 10-fold cross validation framework.

The classifier can accurately detect presence or absence of asynchrony and its type as reflected by high sensitivity, specificity, positive predictive...
value, and negative predictive value. Both sensitivity and specificity were high for all three classes of cycling asynchrony or no cycling asynchrony, ranging from 99% sensitivity for premature termination to 97% sensitivity for neither, and 93% specificity for neither to 99% specificity for premature termination. In addition, the machine learning algorithm showed a strong agreement with human observers as reflected by the high kappa coefficients ranging from 0.90 to 0.91.

Funding

This work was supported by the National Science Foundation grant IIP-1456404 awarded to Autonomous Healthcare, Inc.

Conflict of Interest

B.G. has stock ownership in Autonomous Healthcare Inc. T.S.P. has stock options in Autonomous Healthcare. B.G. and T.S.P. are inventors on a provisional patent application (assigned to Autonomous Healthcare) for a patient-ventilator asynchrony detection system. W.M.H. has stock ownership in Autonomous Healthcare. J.M.B. has stock options in Autonomous Healthcare and serves as its Chief Medical Officer. A.C., J.M., L.P. have received support from Autonomous Healthcare. A.C. is a part-time employee of Autonomous Healthcare.

Acknowledgements

The assistance of Candace Cox, Joseph Briggs, and Leonard T. Barrett in data collection is acknowledged.

References


Behnood Gholami received the B.Sc. degree in mechanical engineering from the University of Tehran, Tehran, Iran, in 2003, the M.A.Sc. degree in mechanical engineering from Concordia University, Montreal, Canada, in 2005, and the M.S. degrees in mathematics and aerospace engineering, and the Ph.D. degree in aerospace engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, in 2009 and 2010, respectively. He was a Postdoctoral Fellow at the School of Electrical and Computer Engineering, Georgia Tech, from 2010 to 2011 and the Brigham and Women’s Hospital, Harvard Medical School from 2011 to 2012. He is currently the Founder and CEO of Autonomous Healthcare, Inc., Hoboken, NJ. His current research interests include machine learning and application of dynamical systems and control theory to biomedicine.

Timothy S. Phan received the B.S. degree in Electrical and Computer Engineering, B.S. degree in Biomedical Engineering, and M.S. degree in Electrical Engineering from Rutgers University, New Brunswick, NJ. He is currently the Director of Engineering at Autonomous Healthcare, Inc. His research interests include cardiovascular system dynamics, mathematical modeling of physiological systems, and noninvasive phenotyping of cardiopulmonary-vascular system coupling with machine learning and medical device design.

Vassim M. Haddad received the B.S., M.S., and Ph.D. degrees in mechanical engineering from Florida Tech in 1983, 1984, and 1987. Since 1994 he has been with the School of Aerospace Engineering at Georgia Tech, where he holds the rank of Professor, the David Lewis Chair in Dynamical Systems and Control, and Chair of the Flight Mechanics and Control Discipline. He also holds a joint Professor appointment with the School of Electrical and Computer Engineering at Georgia Tech. Dr. Haddad has made numerous contributions to the development of nonlinear control theory and its application to aerospace, electrical, and biomedical engineering. His interdisciplinary research in systems and control is documented in over 630 archival journal and conference publications, and 7 books in the areas of science, mathematics, medicine, and engineering. Dr. Haddad is an NSF Presidential Faculty Fellow; a member of the Academy of Nonlinear Sciences; an IEEE Fellow; and the recipient of the 2014 AIAA Pendray Aerospace Literature Award.

Andrew Cason received the B.S. degree in political science from the University of Georgia, Athens, GA in 1994. In 1998, he received the associate degree in respiratory therapy from the Gwinnett Technical College, Lawrenceville, GA. He has been a respiratory therapist since 1998 and has 16 years of critical care experience. His research interests include optimization of patient-ventilator interaction and its impact on decreasing ICU length of stay, morbidity and mortality.

Jerry Mullis received the associate degree in respiratory therapy from the Heart of Georgia Technical College in 2006. He has been a respiratory therapist since 2006 and has over 11 years of critical care experience. Currently, he serves as the dayshift respiratory supervisor at the Northeast Georgia Medical Center, Gainesville, GA. His research interests include optimization of patient-ventilator interaction, airway pressure release ventilation (APRV) for trauma patients, and extracorporeal membrane oxygenation (ECMO) criteria and their correlation with outcome.

Levi Price attended the North Georgia Technical College in Clarkesville, GA and became a Nationally Registered Emergency Medical Technician in 2003. He then joined the North-east Georgia Medical Center, Gainesville, GA as an emergency medical technician in 2004.
After developing an interest in respiratory therapy, he received the associate degree in respiratory therapy from Gwinnett Technical College, Lawrenceville, GA in 2011. He currently serves as a critical care respiratory therapist in the Trauma ICU at the Northeast Georgia Medical Center. His research interests include optimization of patient-ventilator interaction and improving patient compliance while being mechanically ventilated.

James M. Bailey received his B.S. degree from Davidson College in 1969, a Ph.D. in Chemistry (Physical) from the University of North Carolina at Chapel Hill in 1973, and the M.D. degree from Southern Illinois University School of Medicine in 1982. After receiving his M.D. degree he completed a residency in anesthesiology and then a fellowship in cardiac anesthesiology at the Emory University School of Medicine affiliated hospitals. From 1986 to 2002 he was an Assistant Professor of Anesthesiology and then Associate Professor of Anesthesiology at Emory, where he also served as director of the critical care service. He is currently Medical Director of Critical Care Medicine at the Northeast Georgia Physicians Group.