



HAL
open science

Ineffective Efforts in ICU Assisted Ventilation: Feature Extraction and Analysis Platform

Achilleas Chytas, Katerina Vaporidi, Dimitris Babalis, Dimitris Georgopoulos, Nicos Maglaveras, Ioanna Chouvarda

► **To cite this version:**

Achilleas Chytas, Katerina Vaporidi, Dimitris Babalis, Dimitris Georgopoulos, Nicos Maglaveras, et al.. Ineffective Efforts in ICU Assisted Ventilation: Feature Extraction and Analysis Platform. 12th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2016, Thessaloniki, Greece. pp.642-650, 10.1007/978-3-319-44944-9_57. hal-01557617

HAL Id: hal-01557617

<https://inria.hal.science/hal-01557617>

Submitted on 6 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Ineffective Efforts in ICU Assisted Ventilation: Feature Extraction and Analysis Platform

A. Chytas, K. Vaporidi, D. Babalis, D. Georgopoulos, N. Maglaveras,
Senior Member, IEEE, I. Chouvarda, *Member, IEEE*

Abstract. Intensive Care Unit (ICU) is a challenging environment, requiring continuous monitoring and treatment adaptations, raising the need for tools and platforms to support medical decisions. In this context, the focus of this work is in supporting clinicians in managing assisted ventilation intervention (AVI). In AVI the need for patient-ventilator coupling exists. Attention may be required in cases when patient's effort doesn't trigger the ventilator at all, and the assisted ventilation event is lost, i.e. when an ineffective effort (IE) event takes place. A high exposure to IEs has been related to adverse clinical outcomes. The purpose of this work is to create new features that complement the already existing IE index in terms of estimating the adverse effects of ventilation exposure. A series of tools varying from raw data handling to the creation of predictive models are created and implemented in a custom platform, utilizing open-source software.

1 Introduction

In an Intensive Care Unit (ICU) context, patient bio-signals are continuously monitored and displayed towards recognizing alerting events. These recordings of physiologic waveforms along with data coming from laboratory examinations and patient interaction (medications, medical procedures) can be difficult to be interpreted especially in an environment as demanding as the ICU. Thus there is a need for analysing and displaying the data in an easy-to-understand manner. Real-time analysis of patients' bio-signals can be used to detect conditions that precede medical complications using both domain expert knowledge and knowledge obtained by automated procedures [1].

The ventilation dy-synchronization is a known issue; a prime example are the incidences of ineffective triggering [2]. Ineffective triggering, (IT) is common, but factors affecting IT vary considerably and can be contributed to patient condition and factors related to the ventilation system. As a result, the IT frequency of appearances and overall distribution varies as well.

In the medical literature, the total time spent under ineffective effort has been proposed as an index related to adverse clinical events. However, these events are sometimes not equally distributed in time. Ineffective triggering of the ventilator is frequent but highly variable among patients and during the course of mechanical support for each patient. As previously reported [3], most patients have small (5minute) periods with high intensity of ineffective synchronization, i.e. ineffective efforts (IEs). It is still an open issue whether the cumulative effect of IE exposure during the total time in ventilation or the temporal patterns of IEs relate to patient deterioration.

The AEGLE [4] project was created by the need to provide the tools and the necessary data for physicians and researchers to explore new data and answer research questions. The datasets and tools described in this paper will enable clinicians to explore these new features and gain new insight in the related phenomena. In this work we propose a set of features that describe the morphology of IE event time-series and complement the IE index that might be proved helpful in better describing patients and estimating their hospital prognosis, and a web-based platform that is used to perform the analysis and evaluate the outcomes.

2 The AEGLE project Approach

The design and implementation of the 1st AEGLE platform prototype is currently underway, parallel to that, analytic tools such as the ones presented in this paper, are also being designed. Having the data providers, domain experts and analytics developers in different sites can be a hindrance in the design of novel analytics, since direct communication and information flow is of utmost importance especially in the early stages.

The solution was to expose directly to the domain experts the in-development analytics via a web-based platform, so that a feedback loop between the involved parties will be created having as a result a more efficient designing phase.

For the purpose of developing the PVI analytics we chose R, an open source and powerful scripting language. For the last few years it has been among the most common used software packages in scholarly articles, and in 2015 it became the 2nd most used, while also being in fourth position in terms of usage growth in the same domain [5]. That makes R the most popular free analytic software, having as an added bonus its' high flexibility in that it provides the users the ability to create custom analytics. At the same time, it benefits greatly from a large and active community that contributes to analytics and visualization libraries, and frameworks.

3 Methodology

3.1 Data Collection

For the purposes of this study recordings from 108 patients for multiple days were obtained from the ICU clinic of the University Hospital of Heraklion, Crete (PAGNI) using an experimental protocol; Patient-Ventilator Interaction (PVI) Monitor [6], as well as a selection of fields from the hospital Electronic Health Records (EHR). The raw data produced is approximately 12Mb per patient per day as an average.

The data providers uploaded pseudo-anonymized data as files on a secure location. Later a more appropriate solution for datasets containing time-series was used by utilizing the NoSQL Apache Cassandra, although at this point it is implemented over workstation level resources.

3.2 Preprocessing

The main problem with the dataset that hinders a frequency based analysis is that the recordings are event driven, and thus the time difference between consecutives recordings varies, and in several cases significant. We applied a pre-processing tool that first utilizes general data cleaning methods and case specific rules (e.g. no breaths or attempts in breathing for 3 minutes is considered an artefact), and afterwards resamples the data to a fixed sampling rate of 30s [7][8].

Ineffective effort for more than 10% of total breath count is believed to cause problems and prolong hospitalization time [2][9]. We focused on the examination of periods in which the patient was experiencing serious troubles in breathing. We call such incidents IE events [7].

Table 1 INDICES OF 1st CATEGORY

Name and Description	Equation
Power: the sum of all IE, this is a helping index used to describe others	$Pow_{tot} = \sum IE$
Mean: the average of the IE. Also descriptive index for others	$Pow_m = \frac{1}{n} \sum IE$
Density: the percentage of samples in which the patient is experiencing at least one IE.	$Density = \frac{1}{N} \sum_{i=1}^N \begin{cases} 1, IE_i > 0 \\ 0, IE_i == 0 \end{cases}$ <p>where $i = 1, \dots, N$ the samples of the IE signal</p>
Mean to Density: Utilizing the previous indices we define this new one. For two patients having the same Mean score, the patient with the lowest Density score (same percentage of IE but more concentrated) will end up with a higher Mean to Density score, thus differentiating between those two.	$M2D = \frac{Pow_m}{Density}$
Area over x: this index is based on Density and represents the percentage of samples where the patient is experiencing at least x amount of IE.	$AreaOver_x = \frac{1}{N} \sum_{i=1}^N \begin{cases} 1, IE_i > x \\ 0, IE_i == x \end{cases}$
Variation: coefficient of variation for the IE signal.	$VarCof = 100 \frac{SD(IE)}{Pow_m}$
IE distance: median time distance between consecutives IE	
Max clean Area: percentage of the maximum IE free time period in reference to the total recording period.	

3.3 Feature Extraction for IE

A common used feature for the patient-ventilator interaction is the IE index.

$$\frac{\text{Ineffective Efforts}}{\text{Ineffective Efforts} + \text{Breaths}}, (1)$$

It is defined as the sum of IE to the combined sum of IE and breaths over a course of period, be it the totality of the recording, or smaller time segments (e.g. 1 hour to 5 minute). It gives us a general idea on how the patient fared over a period of time. It is reported that it IEs can lead to extended ventilation time, prolonged hospitalization and that it has an impact on mortality for patients with high IE index (>10%) [2][9]. A question that arises is whether or not a patient that might have a low overall IE index over the course of hours, has increased health risks because he is subject to short periods of intense IT activity. In such cases, even a much lower IE index might be related to patient deterioration. Also, a single feature might not be enough to sufficient describe the complexity of a signal distribution.

In order to address this issue, this paper suggests a set of indices that better describe the IE signal morphology. They can be divided into two categories:

1. Indices that are calculated based on IE signal morphology and are independent from the IE even definition (table 1).
2. Indices that are calculated based on the IE event definition (table 2), and thus can vary depending on the researcher's input.

Table 2 INDICES OF 2nd CATEGORY

Name and Description	Equation
Event Power: The sum of IE that belong to event period, this is also an index with limited predictive capabilities that is used to better describe other indices.	$Pow_e = \sum_{i=1}^n Pow_{e_i}$ <p style="text-align: center;">where n the number of events</p>
Event Duration: the total duration of all IE events combined.	$Dur_e = \sum_{i=1}^n Dur_{e_i}$ <p style="text-align: center;">where n the number of events</p>
Median Event Power: The median power of the IE events	$Pow_{e_m} = Pow_e / n$
Median Event duration: The median duration of the event	$Dur_{e_m} = Dur_e / n$
Event distance: median time distance between consecutives IE events, for recordings that did not have an event this value was the total recording time, for recordings that had a single event this value is	$\frac{Dur_{tot}}{2} - Dur_e$
Area concentrating power (x%): in this case the IE event , the percentage of time in which the x% percentage of power is concentrated. This requires to a) Locate the $i=1, \dots, n$ events, with Pow_{e_i}, Dur_{e_i} , b) Sort them by descending density d_i , c) Sum the m more dense events till we reach the	$d_i = \frac{Pow_{e_i}}{Dur_{e_i}}$ <p style="text-align: center;">Local event density</p>

desired x%, the index is calculated as the sum of their duration to the total recording duration. In most cases the desired x percentage is either not reached or is surpassed, so the index is scaled to the desired value.

$$ACPow_x = \frac{1}{Dur_{tot}} \sum_{i=1}^k Dur_{e_i}$$

$$ACPow_x = \frac{ACPow_{x' * x'}}{x}$$

Concentrated density (x%): in an effort to differentiate between patients that have low *AreaConPow_x* because they have several IE and most of them are concentrated in small places, or they have almost zero IE.

$$ConDensity_x = \frac{AreaConPow_x}{Density}$$

3.4 Exposing ICU Analytics via Web-based Platform

In order to make the developed tools accessible to physicians a custom web based ICU platform was created. The main goal was to present the analytic tools in a simplified way, hiding functionality from the user when the steps were already predefined while giving them the ability to parameterize tools when required. The platform supports two major functionalities (figure 1).

The first functionality is a module regarding data pre-processing and processing. Currently only a PVI dataset module is in place. The physicians can upload to the database pseudo-anonymized raw data as extracted by the PVI monitor, then query raw datasets to be pre-processed as previously defined. Afterwards, the user can select either of the two analysis pipelines currently in place, the feature extraction and/or the correlation between the ventilator recorded signals and their phase delay with each-other estimation, based on wavelet coherence [7]. The wavelet coherence analysis explores the time-series correlation with each other and produces a set of features that are expected to give insight regarding the physiological phenomena that take place at the vicinity of an IE event, either preceding it (potential causes) or following it (potential consequences). On each case, the user has the option to experiment using different thresholds regarding the IE events, thus affecting the outcome of the analysis.

The features extracted by the ventilator dataset are combined with a segment of the patients' clinical data that was retrieved offline from the hospital database.

The second functionality is to provide a set of analytics and exploratory visualization, offering a set of commonly used analytics. The platform is designed agnostic of the data type it is provided, although it requires them to be on a tabular format. The user can apply data cleaning methods (out of bound, missing values, etc.) either automatically or manually by the use of UI elements. There, the physicians can run the statistical analytic functions they select among the available ones and evaluate the clinical outcomes they choose. For this segment, R packages that implement well known and established algorithms were utilized.

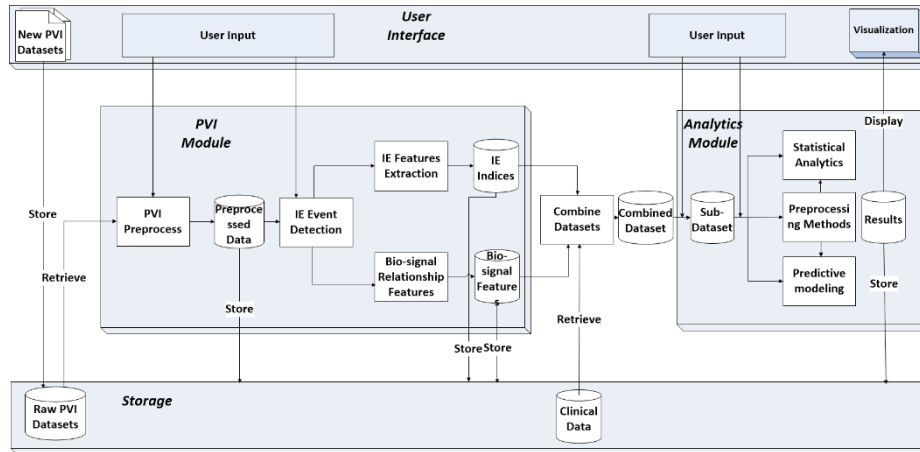


Figure 1. Overview of platform architecture

3.5 Web-based Platform Implementation

The web based platform was developed with Shiny, an R framework focused on creating web applications. An important part of the platform was the data visualization. In order to convey the information in a meaningful way based on the user ever-changing needs, we decided to focus on interactive visualization tools. Thus, we examined tools that are available as R packages. As it turned out there wasn't a single package that could cover the entirety of our needs, so we choose the packages described in table 3.

Table 3 Visualization Packages

Visualization Packages	Description
GoogleVis	R interface to Google Charts API, a great collection of charts, but the time-series libraries were somewhat lacking[10]
Dygraphs	R interface to the 'dygraphs' JavaScript charting library. It focus on depicting time-series on highly customizable charts[11]
D3Heatmap	A single type of chart (Heatmap) that is highly customizable[12]

Specifically, for the Google charts, a custom function was created for the visualization of an entire column based data frame, as individual columns and the relation with each other (figure 1). For an N column data frame, a single mega-chart was created that consists of $N \times N$ individual charts. Each sub-chart C_{ij} ($i, j 1, \dots, N$) depicts the relationship between the variables residing in column i and column j of the data frame, with appropriate visualization that depends on the combination of the variable types. In the cases where $i == j$, a single variable is depicted.

4 Results

Both the platform and the algorithms are currently used by physicians at PAGNI, the functionality and the interface was evolved based on their input. The following two figures show instances of the platform running an exploratory statistical analysis (figure 2) and a data processing pipeline.

Based on analysis enabled by the platform, the optimal threshold for defining respiratory events that relate to adverse outcomes is found to be 10 IE per minute for a time span of at least 3 minutes. On the current state of this research, performing multi-variate analysis, adjusting for age and severity while setting the significance levels at .05 had as results that:

- three indices are found to be related to ICU mortality
- four indices are found to be related to hospital mortality
- two indices are found to be related to the number of days' spent ventilation.



Figure 2. Mass visualization of 5 variable and their combinations(3 numerical and 2 categorical). Red colored charts represent the diagonal (a single-variable chart).

5 Discussion

As the data collection phase proceeds, there will be an increase to the amount of data available (Volume), by including additional equipment recordings such as ICU monitors (recording bio-signals such as Electrocardiogram with a much higher sampling rate than the PVI) and also by increasing the amount of patients whose data are recorded. Provided that the research questions answered by the AEGLE platform yields useful

results regarding patient clinical outcome, our goal is to support streaming data and combine them with the acquired knowledge in order to build a decision support system that will require constant processing (the wavelet-based analysis previously described requires 2 to 10 seconds on workstation level resources depending on IE event duration) of live streaming data (Velocity) . Either way, we are gradually stepping into Big Data territory and performance will be an issue.

In order to address this, and as the AEGLE project moves towards the integrated platform, we are shifting our focus on the analytic integration and optimisation, so we can take advantage of the distributed storage and distributed processing that the combination of RHadoop and SparkR frameworks provide.

The knowledge acquired from this analysis, powered by the platform, can generate alarms thresholds that might be eventually applied on a clinical level.

References

- [1] M. Blount et al., "Real-Time Analysis for Intensive Care: Development and Deployment of the Artemis Analytic System," in *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 2, pp. 110-118, March-April 2010.
- [2] D. Georgopoulos, 'Ineffective efforts during mechanical ventilation: the brain wants, the machine declines', *Intensive Care Med*, vol. 38, no. 5, pp. 738-740, 2012.
- [3] D. Babalis, E. Lilitsis, E. Akoumianaki, et al, Incidence of Ineffective Efforts in mechanically ventilated critically ill patients, *ESICM 2011*
- [4] D. Soudris et al., "AEGLE: A big bio-data analytics framework for integrated health-care services," *Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS)*, 2015 International Conference on, Samos, 2015, pp. 246-253.doi: 10.1109/SAMOS.2015.7363682
- [5] Muenchen, R.: r4stats - 6/8/16: R Passes SAS in Scholarly Use, <http://r4stats.com/2016/06/08/r-passes-sas-in-scholarly-use-finally>.
- [6] M. Younes, L. Brochard, S. Grasso, et al, 'A method for monitoring and improving patient: ventilator interaction', *Intensive Care Med*, vol. 33, no. 8, pp. 1337-1346, 2007
- [7] Chytas, A., et al. "Ineffective efforts in ICU assisted ventilation: Exploring causalities via multiscale analysis." *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. IEEE*, 2015.
- [8] I. Chouvarda, D. Babalis, V. Papaioannou, et al, ' Multiparametric modeling of the ineffective efforts in assisted ventilation within an ICU ', *Medical & Biological Engineering & Computing*, vol. 54, issue 2, pp. 441-451, 2016
- [9] De Wit, Marjolein, et al. "Ineffective triggering predicts increased duration of mechanical ventilation*." *Critical care medicine* vol 37, no.10, pp. 2740-2745, 2009
- [10] Markus Gesmann and Diego de Castillo. Using the Google Visualisation API with R. *The R Journal*, 3(2):40-44, December 2011
- [11] D. Vanderkam, JJ Allaire, dygraphs: Interface to Dygraphs Interactive Time Series Charting Library. R package version 0.5, 2015
- [12] J. Cheng, T. Galili, d3heatmap: Interactive Heat Maps Using 'htmlwidgets' and 'D3.js'. R package version 0.6.1., 2015