eman ta zabal zazu

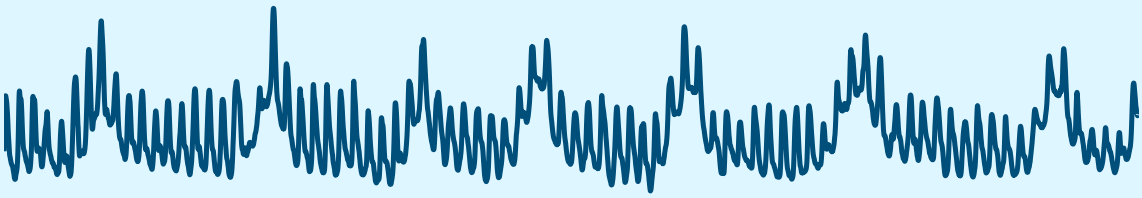
Universidad del País Vasco

Euskal Herriko Unibertsitatea

MACHINE LEARNING AND SIGNAL PROCESSING ALGORITHMS FOR THE ANALYSIS OF VENTILATION AND AIRWAY MANAGEMENT IN OUT-OF-HOSPITAL CARDIAC ARREST

Xabier Jaureguibeitia Lara

September, 2023



Machine learning and signal processing algorithms for the analysis of ventilation and airway management in out-of-hospital cardiac arrest

By

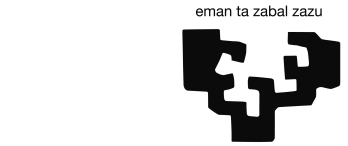
Xabier Jaureguibeitia Lara

Supervisors:

Unai Irusta Zarandona †

Elisabete Aramendi Ecenarro

Doctoral Thesis



Universidad del País Vasco
Euskal Herriko Unibertsitatea

Department of Communications Engineering

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis supervisors. Thank you Unai† for seeing something in me and inviting me to embark on this journey; I hope the final result lives up to the expectations. And thank you Eli for carrying on alone with the direction, and for dealing with the not-so-enthusiastic, possibly exasperating, version of me. This thesis would have not been possible without both of you constantly pushing forward.

Many thanks also to all the other members of the research group: To Erik, Iraia, Andoni and Jon; to former members Bea, Kike, Libe, Mariela and Gorka; to long-term visitors Trygve and Diego; and to the most recent incorporations Amaia and Imanol. Thank you for all the help and technical advice in the development of the thesis, but most importantly, thank you for the casual talks, the coffee-breaks, and all the so-much-needed human support in this often lonely road.

I would also like to extend my gratitude to all the other people and institutions that have made this thesis possible. To our research collaborators, especially professors Henry Wang, Ahamed Idris and their respective teams, for their invaluable contribution of data and their clinical perspective and advice. To Thomas Rea, Jason Coult and the rest of the UWCPR group at the University of Washington, for hosting me in Seattle during my research stay. And to the University of the Basque Country, the Basque Government, and the corresponding Spanish ministries, for their structural and financial support of the thesis.

And last but not least, thanks to my family and friends for their unwavering support and for enduring so many hours of incomprehensible jargon and quite a few existential crises. Do not worry, I will make sure to find new reasons for the beer to keep flowing.

ABSTRACT

Cardiac arrest is defined by the loss, often abrupt and unexpected, of the mechanical activity of the heart. Out-of-hospital cardiac arrest (OHCA) is a leading cause of death in developed countries, with an estimated incidence of 88.8 and 56.3 treated cases per 100,000 persons year in the USA and Europe, respectively. Despite substantial efforts, survival from OHCA remains awfully low, around or below 10%.

High-quality cardiopulmonary resuscitation (CPR), consisting of chest compressions and ventilations, maintains a minimal but critical flow of blood and oxygen to vital organs, and is crucial to improve survival from OHCA. The study of optimal CPR parameters has been an ongoing task for decades, and is largely made possible by signal processing algorithms. Current defibrillator devices acquire multiple biomedical signals during resuscitation, that can then be processed using custom or proprietary software to calculate CPR parameters of potential interest. In combination with detailed clinical information on the care and follow-up of the patients, this has allowed the definition of optimal ranges for important CPR metrics, such as the rate and depth of chest compressions. The optimal ventilation and airway management strategies, on the other hand, are mostly unknown, in large part due to limited measurement.

Capnography, which measures CO_2 concentration in respiratory gases, is the standard method to monitor ventilation in OHCA, but is rarely available until later phases of resuscitation, and cannot be used to measure air volumes. In contrast, thoracic impedance (TI) is acquired through defibrillation pads early on during resuscitation, and shows fluctuations due to air changes in the lungs that can also be used to identify ventilations and estimate air volumes. However, TI is also sensitive to several noise sources, including electrode motion and artifacts due to chest compressions, that may hinder the detection of ventilations. Few solutions have been proposed to detect ventilations in TI during concurrent chest compressions, and they make little or no use of potentially beneficial machine learning techniques.

The effectiveness of ventilation strategies may also be difficult to assess without parallel information on chest compressions. Even for this simpler task, current solutions may not adapt well to large OHCA datasets. The most ambitious studies in resuscitation are multicenter, and often include data from different defibrillator brands which can be difficult to harmonize using vendor-specific software. The many solutions in the literature are also usually designed for single specific signals, and seldom validated outside homogeneous datasets. On top of all this, ancillary but highly important procedures, such as the delineation of resuscitation efforts, are typically not addressed and left for manual review, often unfeasible for large datasets.

This thesis introduces new solutions for a more complete automatic analysis of ventilation and CPR in large OHCA datasets. Three novel machine learning-based solutions were developed for the detection of ventilations in TI with concurrent chest compressions: First, a solution for mechanical CPR, where the fixed and known compression frequency enabled the use of dedicated filters. Second, a solution for manual CPR, incorporating time series classification to exploit the similarity between contiguous ventilations, and a quality control stage to prevent erroneous feedback. And last, a deep learning-based solution, which did not rely on additional reference signals and could potentially be applied to a broader range of scenarios.

In addition, a unified methodology was proposed and evaluated for the automatic annotation of large multi-device datasets. State-of-the-art compression detection solutions were combined and adapted, and new procedures were introduced to automatize ancillary tasks such as signal loss control and delineation of resuscitation efforts.

Finally, these new solutions were used to characterize ventilation and CPR in a large multi-device dataset derived from the Pragmatic Airway Resuscitation Trial, the largest American clinical trial comparing the effectiveness of different airway management strategies. Impactful conclusions were drawn on the effect of airway management on ventilation characteristics and CPR quality, which could potentially help explain outcome differences in OHCA. Collectively, they contributed to advance knowledge in one of the hot topics in resuscitation science.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AED	automated external defibrillator
AHA	American Heart Association
ANN	artificial neural network
BVM	bag-valve-mask
CPR	cardiopulmonary resuscitation
CCR	chest compression rate
CCF	chest compression fraction
CC-CPR	continuous compressions cardiopulmonary resuscitation
CD	compression depth
CI	confidence interval
CO-CPR	compression only cardiopulmonary resuscitation
CV	cross-validation
ECG	electrocardiogram
EMS	emergency medical services
ERC	European Resuscitation Council
ETI	endotracheal intubation
IDR	interdecile range
IHCA	in-hospital cardiac arrest
ILCOR	International Liaison Committee of Resuscitation
IQR	interquartile range
LMS	least mean squares
LoA	levels of agreement
LT	laryngeal tube
MCIR	minimally interrupted cardiac resuscitation
MC-RAMP	multi-channel recursive adaptive matching pursuit

NFT	no-flow time
OHCA	out-of-hospital cardiac arrest
OOB	out-of-bag
PART	Pragmatic Airway Resuscitation Trial
PEA	pulseless electrical activity
PPV	positive predictive value
QS	quality score
RF	random forest
RCT	randomized controlled trial
RNN	recurrent neural network
ROC	Resuscitation Outcomes Consortium
ROSC	return of spontaneous circulation
SD	standard deviation
SE	sensitivity
SGA	supraglottic airway
SP	specificity
TI	thoracic impedance
VF	ventricular fibrillation
VR	ventilation rate
VT	ventricular tachycardia

1 INTRODUCTION

1.1 Out-of-hospital cardiac arrest

Cardiac arrest is the cessation of mechanical activity of the heart, which interrupts blood flow to the brain and other vital organs [1]. Cardiac arrest is often abrupt and unexpected, and potentially lethal within a short time — usually less than an hour — from the onset of symptoms, an event known as sudden cardiac death [2]. More than half of cardiac arrests occur outside the hospital setting [3], so their effective management is a complex process involving local emergency medical services (EMS), but in which community engagement is also critical [4]. The interventions performed to revert cardiac arrest are collectively referred to as resuscitation.

The etiology of cardiac arrest is heterogeneous, including medical causes but also external precipitants such as drowning or trauma [5]. About 90% of out-of-hospital cardiac arrests (OHCA) are attributed to medical causes [6, 7], and about two thirds to cardiac etiology [8, 9]. In western countries, coronary artery disease is the most common pathology underlying cardiogenic OHCA, accounting for about 75% of cases [10]. Secondary pathologies include cardiomyopathies (10-15%), valvular heart disease (1-5%), and inherited channelopathies (1-2%) [10]. The triggering mechanism is usually electrical instability leading to a lethal arrhythmia such as ventricular fibrillation (VF) or ventricular tachycardia (VT) [11]. However, these rhythms degrade rapidly [12], and most OHCA patients are nowadays in asystole or pulseless electrical activity (PEA) — both with worse prognosis than VF/VT — when first examined by EMS personnel [7, 13]. While an

underlying pathology is present in the large majority of cases, about 50% of cardiac-origin OHCA patients have no previous history of diagnosed heart disease [14].

Cardiac arrest is a leading cause of death in developed countries [3, 15]. However, the precise incidence of OHCA is difficult to estimate due to different definitions, inclusion criteria, and EMS configuration and protocols. Most OHCA episodes occur at home, about half are unwitnessed, and — although with large differences between EMS systems — resuscitation is attempted in only 50 – 60% of cases [7, 13]. The incidence of EMS-treated OHCA in the USA is estimated at 88.8 cases per 100,000 persons year, with a large variation (44.2 – 135.5) between states [13]. OHCA incidence in Europe is estimated at 56.3 EMS-treated cases, also with a large variability between countries (from 27.0 cases in Spain to 91.0 cases in the Czech Republic [7]). The incidence in the Basque Country can be estimated at 38.9 EMS-treated cases per 100,000 persons year [16].

After some improvement, survival to resuscitation has stagnated in the last decade [17]. Overall, it remains dismally low, around or below 10% [7, 13]. Its high incidence, unexpected and sudden nature, and low survival rates make OHCA a major public health problem that deserves attention.

1.2 Resuscitation therapies

Introduced by the American Heart Association (AHA) in 1991 [18], the concept of *chain of survival* (Figure 1.1) emphasizes the sequence of time-sensitive interventions that must be optimized to maximize survival from OHCA. Over the years, the chain has undergone some changes, such as the addition of a fifth link referred to post-resuscitation care [19], and a more recent sixth link on recovery and rehabilitation [3]. Nevertheless, the four original links of the chain encompass most pre-hospital interventions, and remain critical to this day:

- **Early access:** The first link of the chain stresses the importance of a rapid recognition of the cardiac arrest or its symptoms, and immediate activation of the EMS. The chances of successful

Figure 1.1. European Resuscitation Council version of the chain of survival [20], comprising four links: early recognition, early cardiopulmonary resuscitation (CPR), rapid defibrillation, and early advanced and post resuscitation care.

resuscitation can double if the arrest is witnessed, or even more if witnessed by EMS personnel [21]. A rapid EMS response is also an important contributor to survival [22].

- **Early cardiopulmonary resuscitation (CPR):** CPR consists of chest compressions and rescue breaths or ventilations to create a minimal flow of oxygenated blood to vital organs. While CPR alone is unlikely to terminate cardiac arrest, its application is crucial to slow down damage and buy time for other therapies to be applied. Bystander [1] CPR prior to EMS arrival has been associated with a two- to three-fold increase in survival [23, 24]. Both the AHA and the European Resuscitation Council (ERC) stress the importance of community awareness, education and new technologies [2] to increase the number of potential rescuers capable and willing to provide CPR [4, 26].
- **Early defibrillation:** A high number of cardiac arrests begin with a ventricular arrhythmia such as VF or VT [10]. While in such a state, the heart can be restored to its normal function by delivering an electrical shock known as defibrillation [27]. The time from arrest to defibrillation is critical, as these rhythms degrade rapidly, reducing the chances of return of spontaneous

1 The term *bystander* is used to refer to a person not responding as part of the EMS, but who participates to some extent in the chain of survival [5].

2 including, but not limited to, dispatcher-assisted CPR [25], and alarm/notification systems within first responder programs.

circulation (ROSC) and survival [12]. Shocks delivered within 3–5 minutes from collapse have been associated with very high survival rates of 50–70% [28, 29]. Such short response times are in part possible thanks to public access defibrillation programs [30], which enable defibrillation by bystanders prior to EMS arrival using automated external defibrillators (AED) [31]. Every minute that defibrillation is delayed, survival from VF OHCA is estimated to decrease by 10–12% in absence of treatment [32, 33], or by 3–4% if CPR is provided [34, 35].

- **Early advanced care:** Early CPR and defibrillation alone are often not enough to achieve or sustain spontaneous circulation. Advanced care encompasses additional interventions carried out by healthcare professionals, such as intravenous access and drug administration, or the insertion of an advanced airway [36]. The ERC version of the chain (see Figure 1.1) integrates into this link post-resuscitation care, which is initiated once sustained ROSC is achieved, and conducted primarily in-hospital [37].

The specific details and interventions associated with each link have evolved over time in the light of new scientific evidence and technical possibilities. Founded in 1992, the International Liaison Committee of Resuscitation (ILCOR) provides a forum for experts and organizations worldwide to examine and reach consensus on the body of science. The conclusions of the ILCOR are then adopted by regional organizations such as the AHA and ERC to update the *resuscitation guidelines*, which include the latest recommendations for resuscitation practice. The most recent AHA and ERC guidelines date from 2020 [3] and 2021 [38], respectively.

1.3 Cardiopulmonary resuscitation

CPR has remained a central part of resuscitation since it was first introduced in its modern form, comprising chest compressions and ventilations, in 1960 [40]. Effective chest compressions can maintain limited but critical blood flow to the brain and other vital organs [41]. Ventilations may also play a crucial role in resuscitation by allowing for oxygen administration and CO_2 removal [42]. Finding the optimal CPR parameters that maximize survival has been object of study for

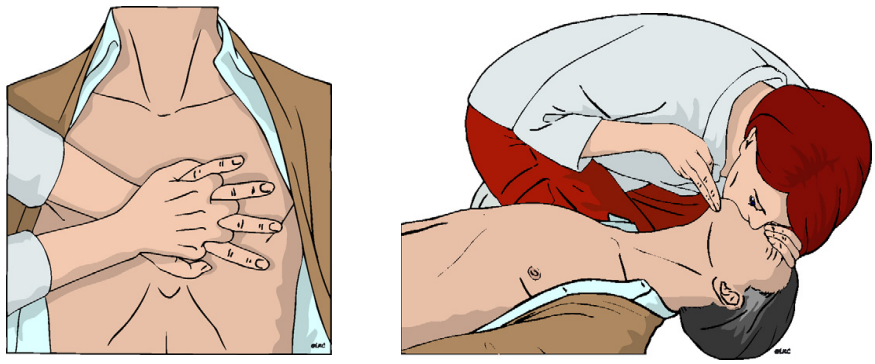


Figure 1.2. Conventional CPR, comprising chest compressions and rescue breaths. Chest compressions should be delivered in the middle of the chest, with both hands overlapped. To provide rescue breaths, open the victim's airway, pinch the nostrils shut, and create a mouth-to-mouth seal. Extracted from ERC guidelines 2010 [39].

decades, and continues to this day [43]. The resuscitation guidelines draw from all the evidence collected and describe how CPR should be performed both by bystanders and EMS personnel [17, 44, 36]:

Chest compressions should be applied in the middle of the chest (lower part of the sternum) using both hands (see Figure 1.2), at a rate between 100 and 120 min^{-1}, a depth between 5 and 6 cm, and minimizing interruptions. Leaning on the patient should also be avoided, as it may prevent full chest recoil and reduce the already limited venous return [45]. When two or more rescuers are available, it may be reasonable to switch compressors about every 2 min, since rescuer fatigue has been associated with a decrease in compression quality [46].

For prolonged resuscitation, EMS personnel may also consider the use of mechanical chest compression devices (see Figure 1.3), which are designed to deliver high-quality compressions in an autonomous manner. However, no benefit in survival has been yet demonstrated for these devices [47].

Regarding ventilation, compression only CPR (CO-CPR), with no positive pressure rescue breaths, is the standard recommendation for bystander CPR [17, 44]. However, if trained and willing, bystanders are encouraged to perform conventional CPR, alternating series of 30 chest compressions and 2 mouth-to-mouth rescue breaths (30:2 CPR, see Figure 1.2). EMS personnel should begin ventilating the patient

Figure 1.3. The two main types of automatic compression devices: pneumatic pistons, such as the LUCAS-3 (Stryker, Redmond, WA, USA; left image), and load-distributing bands, such as the AutoPulse (ZOLL medical, Chelmsford, MA, USA; right image).

using a bag-valve-mask (BVM) system, and then progress to more advanced airways as appropriate (see Figure 1.4). BVM ventilations should be delivered either following 30:2 CPR, or concurrently with continuous chest compressions at a rate of $10\,\mathrm{min}^{-1}$ (CC-CPR). Each ventilation should be delivered over one second, and insufflate a tidal volume around 600 mL [48], or enough to produce a visible chest rise. Compression pauses for ventilation should be minimized ($\leq 5\,\mathrm{s}$), and excessive ventilation should be avoided. Once an advanced airway is placed, ventilations should be delivered following CC-CPR at a rate of $10\,\mathrm{min}^{-1}$.

Figure 1.4. Ventilation devices. On the left, a standard bag-valve-mask (BVM) system. On the right, a King laryngeal tube (LT – Ambu-King Systems, Noblesville, IN, USA), advanced airway device.

1.3.1 CPR QUALITY AND RESUSCITATION OUTCOMES

High-quality compressions have been associated with improved OHCA outcomes. Observational studies suggest that too shallow compressions (< 4 cm) may result in reduced hospital admission and survival [49, 50]. No benefits have been observed for depths greater than 6 cm, which in turn have been linked to an increase in injuries such as contusions and broken ribs [51]. Similarly, reduced ROSC rates have been observed for chest compression rates (CCR) below $80\,\text{min}^{-1}$ [52]. Reduced ROSC and survival has also been reported for rates $> 140\,\text{min}^{-1}$, sharply decaying past $125\,\text{min}^{-1}$ [53]. Excessive CCR may prevent the heart from refilling between compressions [54], and is often accompanied by suboptimal compression depths [55]. Finally, interruptions in chest compressions have also been associated with diminished ROSC and survival, especially if immediately prior to a defibrillation shock [56, 57]. Reduced survival has been observed for chest compression fractions [3] (CCF) below 60% [59], or even 80% in case of prolonged resuscitation [60].

Compared to chest compressions, the importance of ventilation in OHCA is much more debated, especially in the initial phases of resuscitation. Sufficient levels of oxygen may be available during the first minutes of arrest [42]. Excessive ventilation may also result in adverse hemodynamics, such as diminished venous return and cardiac output [61]. In line with this, improved survival rates have been observed following the adoption of CO-CPR as default form of bystander CPR [62, 63]. Increased survival has also been reported in the EMS setting after the implementation of minimally interrupted cardiac resuscitation (MICR) [64], where early ventilation is replaced by passive oxygenation [4]. However, lack of ventilation has also been associated with increased atelectasis (alveolar collapse), diminished pulmonary blood flow, and overall worse resuscitation outcomes [42]. While CO-CPR has led to higher bystander CPR and survival rates, increased survival has been observed among patients who received

3 *Chest compression fraction* (CCF) is defined as the proportion of resuscitation time during which chest compressions are delivered [58].

4 Animal studies suggest that, given an opened airway, intrathoracic pressure changes due to chest compressions may result in passive ventilation [65].

30:2 bystander CPR compared to those treated with CO-CPR [7, 63]. Whether passive ventilation can deliver effective air volumes is also debated [66]. The benefits of MICR and CO-CPR in the EMS setting have been observed mainly in patients with witnessed VF OHCA [64, 67], who may still present with oxygen reserves [42] and even some degree of agonal breathing [68]. The heterogeneity of cardiac arrest, and the usually "bundled" nature of resuscitation care (encompassing multiple interrelated interventions), make it difficult to ascertain the true impact of ventilation.

Even when ventilations are delivered, the optimal form of CPR is unknown. Observational studies and mathematical models suggest that 30:2 CPR may be superior to other previously used compression-to-ventilation ratios, such as 5:1 or 15:2 [69, 70]. Even so, pauses in chest compressions result in a loss of perfusion pressure that may override the benefits of ventilation [69]. Once an advanced airway is placed, CC-CPR enables ventilation while minimizing compression pauses. In this context, both extremely high ($\geq 30\,\mathrm{min}^{-1}$) and low ($\leq 2\,\mathrm{min}^{-1}$) ventilation rates — often referred to as hyper- and hypo-ventilation — have been associated with degraded hemodynamics and potentially worse outcomes [71, 72]. However, and although the resuscitation guidelines recommend a rate of $10\,\mathrm{min}^{-1}$, there is no conclusive evidence to establish an optimal ventilation rate [73, 74]. The effectiveness of CC-CPR with BVM ventilation is also debated. A large randomized controlled trial (RCT) comparing 30:2 CPR and CC-CPR with BVM ventilation found no differences in the primary survival outcome [75]. Secondary and per-protocol analyses, on the other hand, revealed improved outcomes linked to 30:2 CPR [75, 76]. One possible explanation could be increased rescuer fatigue during CC-CPR, leading to degraded compression quality [77]. Concurrent ventilations and chest compressions could also result in reversal air flow and diminished tidal volumes [78]. A recent study suggests that insufficient/ineffective air volumes could already be common in BVM ventilation during 30:2 CPR [79], an occurrence that could potentially be more frequent during CC-CPR. Questions like this remain largely unanswered, as current technology used in OHCA does not monitor nor measure insufflated air volumes.

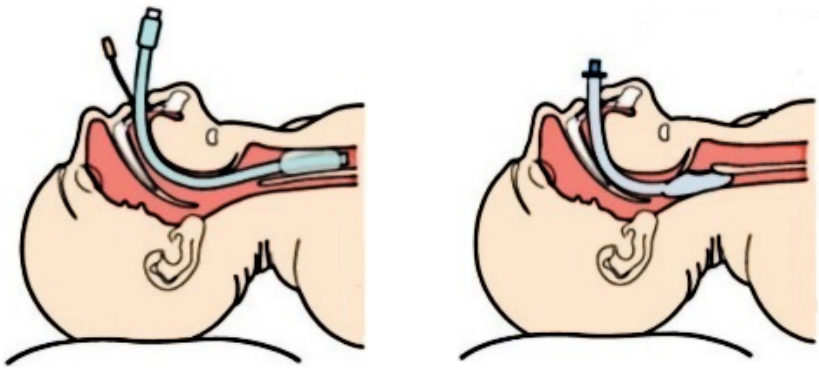


Figure 1.5. Advanced airway management techniques. On the left, endotracheal intubation (ETI). On the right, a supraglottic airway (SGA) device.

1.4 Advanced airway management

Advanced airway management provides a more direct conduit to the lungs, which in theory should allow for more efficient ventilation and oxygenation [42]. As shown in Figure 1.4, two types of advanced airway techniques can be distinguished: endotracheal intubation (ETI) and supraglottic airways (SGA). ETI has been the standard advanced airway technique in OHCA for decades [80], and involves the insertion — with aid of either direct or video laryngoscopy [36] — of a flexible tube down to the patient's trachea. The tracheal tube is usually ended in an inflatable cuff to seal the trachea, preventing air leakage and protecting the lungs from gastric contents. More recently, SGAs have been adopted by many EMS agencies as primary method of advanced airway management. Compared to ETI, SGAs require less training, and may allow for faster insertion [42]. In general, SGAs provide a pathway to the laryngopharinx, and include cuffs or other methods to isolate the oropharynx (which connects to the mouth and nose) and the esophagus. Well-known commercial SGA devices include the esophageal tracheal Combitube (ETC; Covidien-Nellcor, Boulder, CO, USA), the laryngeal mask airway (LMA; LMA North America, San Diego, CA, USA), the i-gel (Intersurgical Ltd, Wokingham, UK), and the King laryngeal tube (LT; Ambu-King Systems, Noblesville, IN, USA; see Figure 1.4).

While the placement of an advanced airway is standard practice in post-resuscitation care [17, 36], its use during ongoing resuscitation is more controversial. Advanced airways may offer advantages over BVM ventilation (e.g., better oxygenation and lung protection), but may also be associated with long interruptions in chest compressions and potential injury during the insertion process [42]. Advanced airways are also prone to failed insertion efforts and dislodgement, leading to further interruptions, and also periods of hypoxia in the absence of ventilation [80]. ETI-related interruptions often exceeding 40 s — or even 100 s if accounting for multiple insertion efforts — have been reported in the past [81]. More recent studies suggest shorter interruptions [82], but still frequently above the 5 s recommended in the ERC guidelines [36]. Interruptions could also be similar between the different airway devices [82]. In terms of survival, observational studies have often found BVM superior to advanced airways [83, 84]. However, these results have been questioned for potential biases, as advanced airways may be associated with prolonged resuscitation (typically with poorer outcomes), and may be rarely used if ROSC is achieved early [42, 85]. In contrast, other studies have associated early advanced airway insertion with improved survival in patients with non-shockable rhythms [86, 87]. A recent RCT in a physician-based EMS system found no significant outcome differences between ETI and BVM ventilation [88]. No RCTs have been conducted directly comparing BVM and SGAs.

The optimal advanced airway strategy is also unclear. ETI offers a more direct pathway to the lungs, but this could be outweighed by the easier handling of SGAs. Compared to ETI, SGAs have been associated with higher first-attempt insertion success [82, 89]. ETI may also be more prone to tube misplacement, such as esophageal insertion, which has been reported in rates up to 17% in paramedic-based EMS systems [36]. Observational studies suggest that ETI may be associated with higher survival than SGAs [84, 90], but this has not been supported by recent RCTs. The AIRWAYS-2 trial, which compared initial ETI and i-gel SGA strategies, found no significant outcome differences between both strategies [91]. The Pragmatic Airway Resuscitation Trial (PART), which compared initial ETI and King LT SGA strategies, found better outcomes associated with the

SGA strategy [92]. The results of these RCTs, however, are difficult to contextualize due to important crossovers and low ETI insertion success rates, of 69% and 52% respectively [17]. By contrast, the ILCOR recommends that ETI should only be performed within EMS systems with a high insertion success (of 95% or higher within two insertion attempts, based on ERC expert consensus) [36]. Thus, the optimal airway management strategy could be highly dependent not only on the patient's status and characteristics, but also on the training and experience of the EMS personnel [17].

Considering the many unknown and controversial aspects about ventilation strategies and effects, further research is needed to better understand the pros and cons of the different airway devices, and to effectively tailor the ventilation strategy to different scenarios.

1.5 CPR monitoring and analysis

Resuscitation guidelines emphasize the importance of high-quality CPR to maximize survival from cardiac arrest. However, important deviations from the recommendations have been frequently reported in the literature, including shallow compressions [93, 94], inadequate CCR [43, 52], low CCF [43, 93, 94], or severe hyperventilation [71, 94, 95]. Current defibrillator devices allow the acquisition of care and CPR information during resuscitation, which can later be used both in observational studies and for case debriefing within quality enhancement programs [43, 96].

A defibrillator electronic file may include, among others, a log of device-operated events (such as shocks), continuous and/or periodic recordings of various physiological signals, and a synchronized audio record where the EMS personnel verbalizes the status and care of the patient. Generally, the device manufacturer will also provide software to retrieve and analyze this information, often embedding proprietary algorithms to process the stored signals and derive CPR metrics of common interest. In some cases, lightweight algorithms may also be incorporated in the defibrillator device itself or other field equipment, enabling real-time feedback to the rescuer. Although a direct benefit on survival has not yet been demonstrated [97], the use

Figure 1.6. Defibrillator devices. On the left, a Lifepak LP1000 automated external defibrillator (AED; Stryker/Physio-Control, Redmond, WA, USA). On the right, a HeartStart MRx monitor-defibrillator (Philips Healthcare, Andomer, MA, USA).

of feedback devices has been associated with improved CPR quality, as in the sense of adherence to guideline recommendations [49, 98].

In addition to its own capabilities, defibrillator software usually allows exporting all or part of the collected data to an open format, thus enabling the use of custom analysis tools. This has led to the development of many signal processing solutions to extract CPR information, both on chest compressions [99, 100, 101, 102, 103, 104] and ventilations [99, 105, 106, 107, 108, 109], from different individual or combined signal sources. In general, the extent and reliability of the information that can be obtained is highly dependent on the range of signals acquired, which in turn depends on the brand and model of the defibrillator device. The most important defibrillator commercial brands include Philips Healthcare (Andover, MA, USA), Stryker/Physio-Control (Redmond, WA, USA), and ZOLL Medical (Chelmsford, MA, USA). Defibrillator models can be broadly divided into AEDs and monitor-defibrillators (see Figure 1.6).

1.5.1 Defibrillator devices and signal acquisition

AEDs are the most basic devices used during resuscitation. The primary purpose of AEDs is to enable rapid defibrillation, so they are designed for minimally-trained users such as bystanders [30, 31]. Modern AEDs provide guidance to the rescuer during resuscitation, including audio instructions for pad placement and to stop/resume

CPR. More importantly, they include logic to automatically analyze the heart rhythm of the patient and determine whether defibrillation is appropriate [110, 111]. The standard AED acquires two different signals through the defibrillation pads: the electrocardiogram (ECG), which monitors the electrical activity of the heart and enables rhythm analysis, and the thoracic impedance (TI), which measures the body resistance to current flow. To obtain the TI, a high frequency current (usually of 20 – 100 kHz and 1 – 5 mA) is injected through the pads and the resulting voltage drop is measured; the TI is then computed by direct application of Ohm's law [112]. The original purpose of TI in AEDs was to assess pad-skin contact and to adjust defibrillation energy, but it has also been demonstrated sensitive to both chest compressions and ventilations [113, 114, 115].

The most advanced AED models, designed for professional first responders such as policemen or firefighters, may also be equipped with stand-alone or integrated CPR assist pads (see Figure 1.7). These gadgets include accelerometers and/or force sensors, whose readings can be processed to derive important CPR parameters such as the CCR or the compression depth [103]. This processing can usually be done in real time, allowing immediate feedback to the rescuer.

Beyond AEDs, monitor-defibrillators are more sophisticated devices used in and out of the hospital by EMS physicians and other

Figure 1.7. CPR assist pads. On the left, a Q-CPR pluggable assist pad (Laerdal, Stavanger, Norway; Philips Healthcare, Andover, MA, USA), incorporating both force and acceleration sensors. On the right, an AED PRO defibrillator with CPR-D-padz accelerometer-integrated pads (ZOLL medical, Chelmsford, MA, USA).

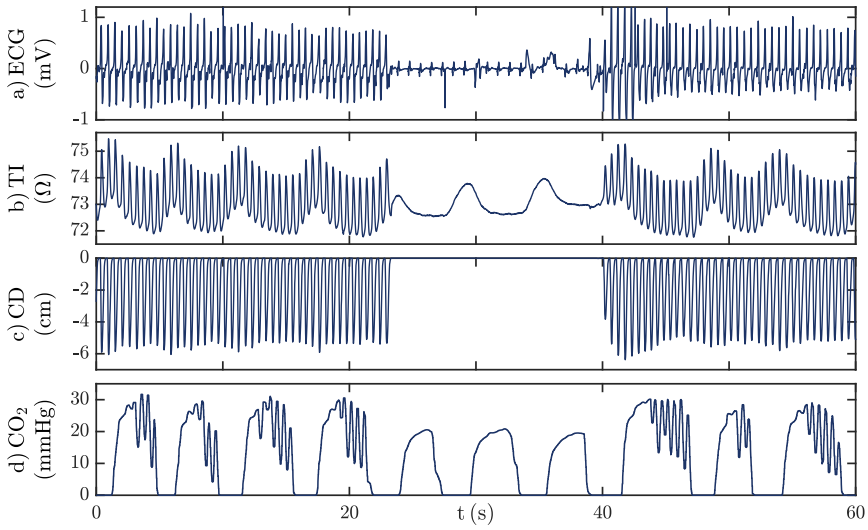


Figure 1.8. Example of some of the signals acquired by monitor-defibrillators. From top to bottom: a) electrocardiogram (ECG), b) thoracic impedance (TI), c) compression depth (CD, constructed from chest force and acceleration), and d) capnogram (CO_2). Rapid fluctuations corresponding to chest compressions can be observed in both TI and CD (and to a minor extent in ECG and CO_2). Slower fluctuations due to ventilations can also be observed in TI and CO_2.

healthcare professionals. These devices enable manual control of defibrillation (although most of them can also operate in AED mode) and provide extensive monitoring capabilities, including the acquisition and visual representation of multiple physiological signal waveforms. These signals can be interpreted by expert EMS personnel and used as a reference for decision making throughout the resuscitation episode. In general, monitor-defibrillators have access to a much wider range of signals than AEDs, including the 12-lead ECG[5] and other supplementary signals such as blood pressures, pulse oximetry,[6], or capnography. Capnography in particular, which measures the concentration of CO_2 in exhaled gases, has seen increased used in recent times due to its many applications in resuscitation. Capnogram readings may serve as both surrogate measures of CPR quality

5 ECG leads are acquired by means of electrodes placed in different body positions, and monitor different electrical planes of the heart.

6 Pulse oximetry is a non-invasive measure of oxygen saturation in blood, acquired through light sensors typically placed on finger or ear.

and early indicative of ROSC, and can also be used to verify tracheal tube placement and to monitor ventilation rates [17, 36]. Figure 1.8 shows some of the typical signals acquired by monitor-defibrillators.

1.5.2 Cardiac arrest registries

In addition to the defibrillator electronic file, healthcare personnel collect many other clinically relevant variables related to the cardiac arrest episode. This information is then compiled into cardiac arrest registries, which may contain data from hundreds or even thousands of OHCA episodes, and which constitute an invaluable resource for the study and advancement of the treatment of cardiac arrest [116].

In an effort to enable comparison between EMS-systems, resuscitation strategies and outcomes, the Utstein template was introduced in 1991 as a standardized format for uniform reporting of OHCA and resuscitation data [18]. The current and refined Utstein template [5] covers variables related to the patient (age, sex, ...), the arrest (initial rhythm, etiology, ...), the response/treatment both by bystanders (witnessed arrest, bystander CPR, ...) and EMS-personnel (response time, first defibrillation time, ...), and outcomes, the latter including scene (ROSC time, survival to hospital admission, ...), hospital (survival to discharge, neurological status, ...) and follow up (survival at 30 days, ...) outcomes. According to EMS and ongoing RCT protocols, other non-standard variables may also be collected, such as timestamps for specific drug administrations or airway insertion efforts.

The combination of clinical and CPR information enables the study of the influence of different interventions and CPR parameters on resuscitation outcomes, while also controlling for other demographic, interventional or CPR-related confounding factors [43].

Among the most important cardiac arrest registries worldwide we can find both the Resuscitation Outcomes Consortium (ROC) [117] and the Cardiac Arrest Registry to Enhance Survival (CARES) [118] in North America, and the European Registry of Cardiac Arrest (EuReCa) [119] in Europe.

1.6 Motivation for the thesis work

The optimal ventilation and airway management strategies in resuscitation are largely unknown. This is partly due to limited measuring capabilities, as underscored by very recent technologies and pilot studies [120, 121]. Resuscitation guidelines recommend the use of capnography to monitor ventilation rates [36]. However, capnography provides limited information on ventilation: first, because it is rarely acquired before the placement of an advanced airway [109]; and second, because it misses potentially relevant metrics such as tidal/insufflated air volumes [79, 122]. TI is sensitive to air volume changes in the lungs, so it has also been proposed to monitor ventilation [114]. Compared to capnography, TI is available much earlier during resuscitation, as soon as defibrillation pads are attached; and while not very accurate, it can also provide surrogate measures of tidal volumes [115]. Few algorithms have been proposed to characterize ventilation based on the TI [99, 105, 106, 109]; however, they have not been validated in large OHCA datasets, and make limited or no use of modern machine learning techniques that have proven effective in other resuscitation-related problems [123, 124, 125].

Besides the characterization of ventilation itself, the effectiveness of ventilation strategies may be difficult to assess without concurrent information on chest compressions. Most commercial software enables calculation of chest compression metrics of typical interest. However, this software is vendor-specific, making it difficult to harmonize analyses when different defibrillator brands are involved (as is the case in most multicenter studies and clinical trials). Alternatively, many solutions have been proposed in the literature to characterize chest compressions [99, 100, 101, 126, 102]; but again, they are based on single specific signals, and have rarely been validated outside limited and homogeneous datasets. Unified methodologies are needed to tackle the analysis of large multi-device datasets, which can be adapted to the retrospective analysis of registries in general, and to the specifics of clinical trials in particular.

This thesis provides new insights for a more complete automatic analysis of ventilation and CPR quality in large OHCA datasets. First, through the application of signal processing and machine learning

techniques in developing new and improved TI-based ventilation detection algorithms for the CC-CPR scenario. And second, through the design and evaluation of methodologies for automatic CPR quality assessment in heterogeneous datasets. Together, these contributions could facilitate the analysis of existing and future data in the study of ventilation, airway management and CPR strategies.

2 BACKGROUND

This second chapter provides context for the most important contributions of the thesis. Sections 2.1 and 2.2 discuss the evolution and current state of automatic CPR analysis. Covering chest compressions and ventilations respectively, these two sections introduce the most widely studied CPR quality metrics, and summarize the algorithms and solutions proposed in the literature to derive such metrics from typically available signal sources. A brief overview of the capabilities and limitations of commercial software is also included. Sections 2.3 and 2.4 cover, respectively, adaptive filtering approaches used in the suppression of chest compression artifacts, and machine learning algorithms for classification. Some of the techniques introduced in these two sections serve as fundamental building blocks in the design of the ventilation detection algorithms proposed in the thesis.

2.1 CPR quality analysis: chest compressions

2.1.1 Chest compression quality metrics

High-quality CPR, as determined by quantitative measures of chest compressions and ventilations, has been associated with improved resuscitation outcomes [49, 50, 52, 53, 59, 60, 61]. Resuscitation guidelines include evidence-based recommendations for many of these quality metrics [17, 36, 44]; however, they do not provide specific instructions on how to calculate and report them. The Utstein template for homogeneous reporting of OHCA and resuscitation data also fails to cover CPR quality [5]. To address this gap, Kramer-Johansen et al. [58] introduced in 2007 the most representative guidelines for

uniform calculation and reporting of CPR quality metrics. The most relevant metrics related to chest compressions include:

- **Pauses in chest compressions:** Chest compressions are essential to maintain circulatory flow during cardiac arrest, so a critical aspect of CPR quality is the time of resuscitation with no compressions delivered, known as *no flow time* (NFT). Kramer-Johansen et al. recommend reporting the total NFT over the full resuscitation episode as well as the fractional NFT relative to the episode duration. The latter is more commonly expressed as its complement, the fraction of time with compressions delivered, known as *chest compression fraction* (CCF).

  Uniform calculation of NFT/CCF requires a procedural definition of resuscitation time. Kramer-Johansen et al. suggest the use of the interval from the first therapeutic event (i.e., first chest compression, shock, or rhythm analysis) to the termination of resuscitation efforts (i.e., last chest compression or onset of sustained ROSC), excluding any intermediate time interval with spontaneous pulse (transient ROSC). Figure 2.1 illustrates the definition of resuscitation time and calculation of the CCF.

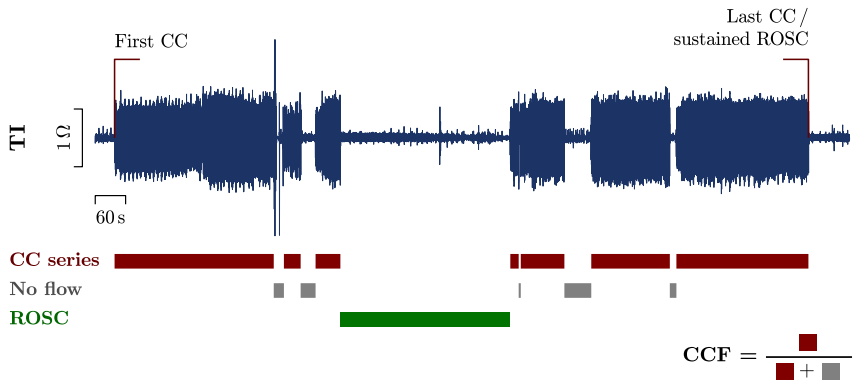


Figure 2.1. Example of CCF calculation. Chest compression (CC) series and no flow intervals are identified, respectively, as the presence or absence of thoracic impedance (TI) fluctuations. The CCF is then computed as the proportion of time corresponding to CC series. The analysis is restricted to resuscitation time, from first CC to last CC (also sustained return of spontaneous circulation, ROSC). Intermediate periods with spontaneous circulation (in green) are also excluded.

Similarly, a procedural definition is needed for the minimum no-flow interval, pause or interruption duration. Kramer-Johansen et al. defined the minimum pause as 1.5 s. However, other values such as 2 s or 3 s have often been used both in the literature and in commercial software [59, 127, 128].

Given their potential impact in outcomes [56, 57], Kramer-Johansen et al. also recommend reporting the median duration of pauses before and after defibrillation.

- **Compression depth (CD)** is defined as the maximum deflection of the sternum prior to chest recoil, and is typically measured using acceleration data from CPR assist pad devices.

  Kramer-Johansen el al. recommend reporting the average CD over the full resuscitation episode, as well as the fraction of minutes with average CD below the guideline minimum (5 cm, according to most recent guidelines [17, 44]). Average CD values on a minute-by-minute basis (or even in shorter segments) could also be considered.

- **Chest compression rate (CCR)** is defined as the frequency of chest compressions during compression series. For any given analysis interval, the CCR can be obtained by identifying the time points of maximal compression, and then averaging the reciprocals of the time differences Δt_i between contiguous chest compressions. Mathematically, this can be expressed as

$$\mathrm{CCR} = \overline{60 \,/\, \Delta t_i} \quad (\mathrm{min}^{-1}), \tag{1}$$

  where Δt_i are given in seconds. Values of Δt_i exceeding the procedural definition of minimum pause should not be considered part of compressions series and should therefore be excluded (see Figure 2.2).

  Alternatively to the CCR, the direct count of chest compressions in a given analysis interval may provide combined information of the interruptions and frequency of compressions.

  Kramer-Johansen et al. suggest reporting both the average CCR and the average count of compressions per minute over the full re-

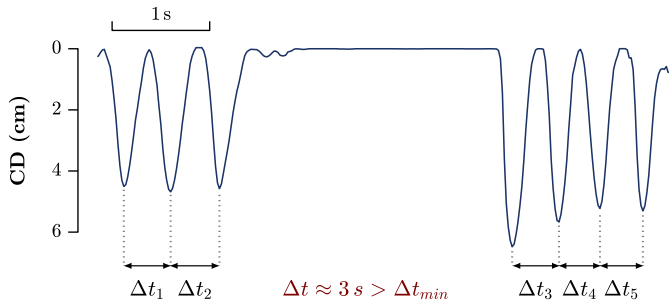


Figure 2.2. Example of a short segment of compression depth (CD) signal. Chest compression instants are assumed at negative peaks of the signal, and used to compute the times between compressions Δt_i. A pause of about 3 s is included. The Δt associated to this pause is not considered for CCR calculations, as it exceeds the procedural definition of minimum pause Δt_{min} (1.5 s, as per Kramer-Johansen et al. [58]).

suscitation episode. Given that such averaging could mask periods of poor quality CPR, standard deviations should be provided as a minimum. The fraction of minutes with CCR outside guideline recommendations ($100-120\,\text{min}^{-1}$ [17, 44]) may also be optionally reported.

Secondary chest compression quality metrics include *duty cycle*, defined as the fraction of time compressing the chest in a compression-decompression cycle, and *incomplete release of compressions*. Given the limited evidence on the impact of these variables, and the difficulties in measuring and/or actively intervening on them, Kramer-Johansen et al. recommend these metrics to be reported optionally, and only as average value over the full resuscitation episode.

2.1.2 Literature solutions for chest compression analysis

Acceleration, force and depth based solutions

CPR assist pads provide the most reliable signal sources to identify and characterize chest compressions. The primary signal acquired by modern CPR pads is chest acceleration, which can be integrated twice to estimate chest displacement or compression depth (CD) [103]. In practice, this is accomplished using numerical methods such as the *trapezoidal rule*, which can be implemented in digital signal processors and applied in real time to provide feedback to the rescuer. However, the integration process is also unstable, highly sensitive to non-zero

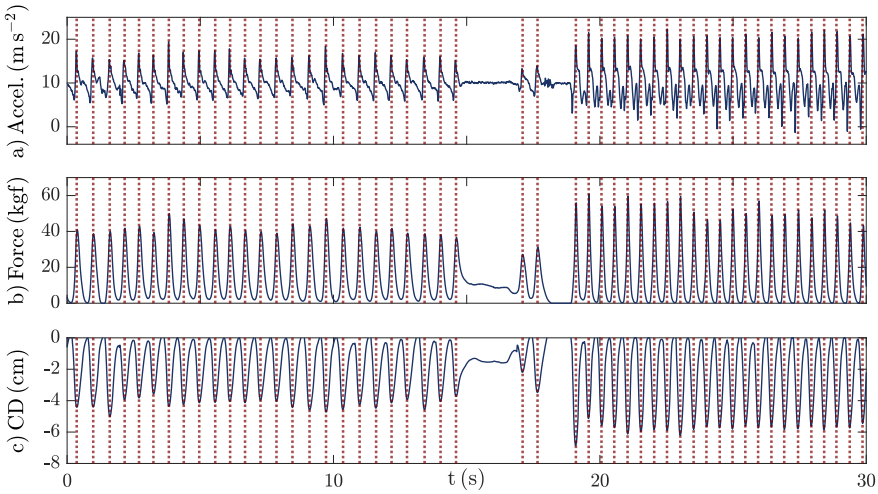


Figure 2.3. Example of the different signals acquired with a Philips-Laerdal Q-CPR assist pad, including: (a) chest acceleration, (b) chest force, and (c) compression depth (CD). Individual chest compressions, as given by the negative peaks of the CD signal, are shown as vertical red dotted lines.

offsets from gravity in acceleration data [103, 129]. Several solutions have been proposed to overcome this problem, either incorporating additional sensors and signal data, or using more sophisticated signal processing techniques to process the acceleration. Figure 2.3 shows an example of the signals acquired with a CPR assist pad.

Gruben et al. [129] proposed the first accelerometer-based system for CD measurement during CPR. The solution was accurate, with absolute errors in CD estimation within 0.5 mm. However, effective correction of the acceleration offset required considerable additional sensors and mechanisms, resulting in a cumbersome system not particularly suitable for out-of-hospital use. Aase et al. [103] proposed a simpler solution, which included, in addition to the accelerometer, a pressure switch to identify the start and end of individual compressions. Assuming the thorax returns to resting position after each compression, the latter could be used to set boundary conditions and avoid instabilities in the integration process. Under regular resuscitation conditions [1], this approach was able to estimate CD with 95% of

1 Regular conditions imply a stiff surface and a static medium. Effective estimation of CD in non-stiff surfaces (such as a mattress) or moving media (such as an ambulance) may require of additional secondary devices for acceleration reference [103, 130].

absolute errors within 1.6 mm. The solution was further developed by Myklebust et al. [131], who replaced the pressure switch with a force sensor. Although the relationship between force and CD is non-linear and patient dependent [132], Gohier et al. [133] showed that additional force data could also help estimate CD without assuming full chest recoil.

Regarding systems based solely on acceleration, with no additional sensors, Babbs et al. [134] proposed linear high-pass filtering before each integration step. González-Otero et al. [104] proposed the least-square reconstruction of the acceleration signal in terms of harmonics of the chest compression frequency, previously identified through spectral analysis. For both these approaches, 95% of absolute errors were found within 11.2 mm and 5.9 mm, respectively [126].

Inaccuracies in depth estimation aside, the signals acquired from CPR assist pads are also well suited for the identification of individual compressions, enabling the calculation of other quality metrics such as the CCF or CCR. Ayala et al. [100] showed that, given a CD signal constructed from acceleration and force readings, a negative peak detector with minimum amplitude of 15 mm was able to identify compressions with a sensitivity [2] (SE) of 98.4% and a positive predictive value [3] (PPV) of 99.8%. Considering one-minute episode segments, the CCF and CCR could be calculated with 95% confidence intervals (CI) of (-6.1, 5.3) % and (-1.4, 1.4) min^{-1}, respectively. For real-time applications, considering the last 8 chest compressions (or a window of about 4 – 5 s), the CCR could be calculated with a 95% CI of (-2.4, 2.4) min^{-1}. Also in the context of real-time feedback, Ruiz de Gauna et al. [126] showed that spectral analysis techniques could estimate CCR with a 95% CI of (-3.0, 3.2) min^{-1} using only acceleration data and in windows as short as 2 s.

2 Sensitivity (SE), also known as *True Positive Rate* or *recall*, is a performance metric associated to detection and classification problems which indicates the proportion of correctly identified ground truth positive class instances (in the above case, actual chest compressions).

3 Positive predictive value (PPV), also called *precision*, is a performance metric associated to detection and classification problems which indicates the proportion of instances detected or classified as positive which actually correspond to ground truth positive class instances.

Thoracic impedance based solutions

In the absence of CD or acceleration data, other highly available signals acquired through defibrillation pads can be used to identify individual compressions. Both the TI and ECG are sensitive to chest compressions, reflecting deformations of the thorax and changes in electrode-skin contact [113, 135]. However, these signals usually show more complex compression waveforms, with higher harmonic content [136], and may also present important confounders, such as electrical activity of the heart in the case of ECG, or ventilation [115] and circulation [137] components in the case of TI (see Figure 2.4). Although automatic solutions for CPR quality analysis based on the ECG have also been proposed [138], the ECG is considered overall inferior for compression monitoring due to major spectral overlap between some heart rhythms and chest compressions [139]. Therefore, most of the literature is focused on TI-based algorithms. It is important to note that the relationship between TI and CD is highly dependent on both the rescuer and the patient [140], so the analysis of compression quality through TI is limited to CCR and CCF.

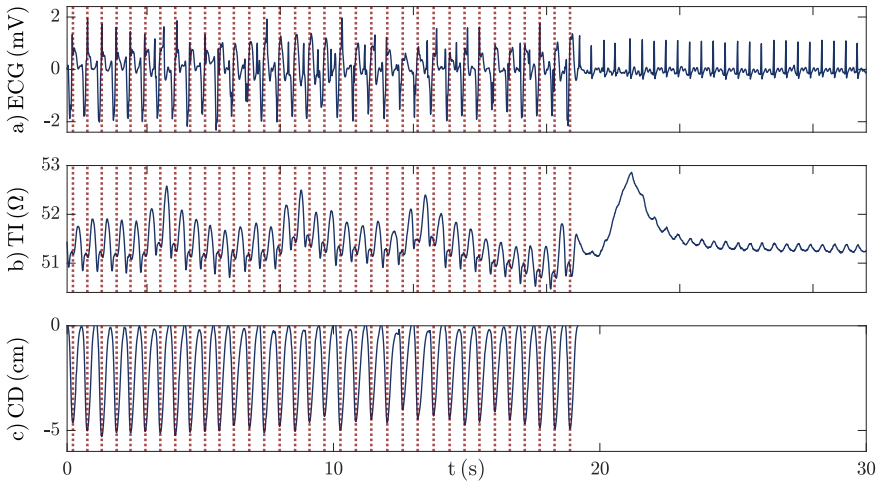


Figure 2.4. Example of the ECG, thoracic impedance (TI) and compression depth (CD) signals, corresponding to a Philips MRx monitor-defibrillator recording. Individual chest compressions, as given by negative peaks in the CD, are shown as vertical red dotted lines. Both ECG and TI show fluctuations corresponding to chest compressions. However, the ECG also shows an organized heart rhythm component. The TI shows slower fluctuations due to ventilations, and minor fast fluctuations due to circulation, correlated with the heart rhythm.

Stecher et al. [113] showed that, provided with both the ECG and TI signal traces, and evaluated against the count of compressions verbalized in the defibrillator record, expert clinicians could identify chest compressions with SE and PPV over 96% and 98%, respectively. They also validated the first commercial solution for automatic chest compression detection based on the TI. Evaluated on 122 clinician-annotated OHCA records, the solution scored a SE of 94% and a PPV of 90%. However, no information on the underlying algorithm was provided, and the impact of errors in compression quality metrics was not analyzed. Later, Aramendi et al. [141] introduced a method to calculate the instantaneous CCR from the TI, which comprised a 1 – 10 Hz band-pass filter and a positive peak detection algorithm to identify individual compressions. Considering episode segments of about 15 s, the CCR estimated from these compressions showed a high correlation ($r = 0.98$) with that computed from ground truth compressions marked in the CD. The main objective of this study was unrelated to CPR quality, so no CCF or detailed CCR analyses were reported either.

The first detailed algorithms for chest compression detection in TI, specifically evaluated in the derivation of CPR quality metrics, were proposed by Ayala et al. [100] and Alonso et al. [99] in 2014 and 2015, respectively. The general structure of the algorithm was similar in both cases, comprising: (a) preprocessing of the TI signal, based on static linear filters, to enhance chest compression components; (b) positive peak detection, to identify and characterize TI fluctuations potentially due to compressions; and (c) rule-based discrimination of chest compressions from other fluctuations. The most critical discrimination threshold Th_Z, corresponding to the minimum compression amplitude, was set dynamically based on the amplitudes Z_i of the last N detected compressions. So, for the potential i-th compression,

$$Th_Z^{(i)} = \sum_{k=1}^{N} w_k \, Z_{i-k} \tag{2}$$

where the weights w_k (uniform in Alonso et al., and monotonically decreasing in Ayala et al.) were optimized for maximum compression detection performance.

- Ayala et al. [100] analyzed 38 OHCA episodes, comprising more than 800 min of TI signal recordings and over 60,000 individual compressions. Evaluated against manual annotations based on the force and ECG signals, the solution identified compressions with a SE of 94.2% and a PPV of 97.4%. Considering complete episodes, CCR and CCF were estimated with 95% CI of (-3.4, 2.2) min^{-1} and (-4, 4) %, respectively. Considering one-minute episode segments, CCR and CCF were estimated with 95% CI of (-5.6, 4.7) min^{-1} and (-11, 11) %, respectively.
- Alonso et al. [99] analyzed 63 OHCA episodes, comprising over 2,500 min of TI recordings and 110,000 compressions. From these, 32 episodes were used for algorithm optimization and 31 for evaluation (using manual annotations based on the CD). In terms of compression detection, the solution scored a median per-episode SE and PPV of 97.2% and 97.7%, respectively. In terms of quality metrics, considering full episodes, the mean (SD) absolute errors in CCR and CCF were 1.8 (4.3) min^{-1} and 2 (2) %. Three episodes presented absolute CCF errors above 5%. A single episode presented an absolute CCR error above 10 min^{-1}.

Around the same time, González-Otero et al. [101] proposed an alternative solution for estimation of compression quality metrics based on the analysis of short TI windows: First, the TI signal was divided into non-overlapping 2 s segments, which were band-pass filtered in the 1 – 3.5 Hz range (or 60 – 210 min^{-1}, in terms of CCR). Then, the presence or absence of compressions was determined based on an adaptive signal power threshold. Finally, for those segments identified as presenting compressions, the CCR was estimated by analyzing the most prominent peaks in the power spectrum. The solution was evaluated using 180 episodes from three different OHCA datasets. In terms of identifying the presence of chest compressions, the mean per-episode SE and PPV were 96.3% and 97.0%, with 19 (10.6%) and 14 (7.8%) episodes, respectively, presenting SE and PPV values below 90%. The error in CCF estimation exceeded 10% in 20 (11.1%) episodes. Regarding the CCR, the 95% of analysis windows presented an absolute error below 5.8 – 13.8 min^{-1}, depending of the dataset. When considering full episodes, only one episode presented a relative CCR error exceeding the 10%.

More recently, more sophisticated statistical and machine learning models have been applied to assess the presence of chest compressions. Kwok et al. [102] introduced a *hidden Markov model* solution using consecutive one-second TI segments, each characterized by four emission features from the time and frequency signal domains. The mean dominant frequency of the segments was used to calculate the CCR. The solution was validated using 105 OHCA records, with ground truth compressions annotated based on the chest force signal, and scored a median (IQR) SE and specificity [4] (SP) of 99 (98 – 100) % and 98 (95 – 100) %, respectively. Considering one-minute episode intervals, the median (interdecile range, IDR) errors in CCR and CCF were 1.8 (-0.5, 5.0) min^{-1} and 0 (-3, 2) %. As similar solutions, Coult et al. [143] proposed a logistic regression model, using three time and frequency domain features, to identify the presence of compressions in variable length TI segments up to 5 s. Rueda et al. [144] proposed a *Random Forest* classifier [145], fed with 18 features from different signal domains, to identify chest compression pauses in TI. Despite potential improvement, however, none of the latter methods have been evaluated in estimating the CCF.

2.1.3 Chest compression analysis in commercial software

This section describes the capabilities of commercial software to automatically derive CPR quality information from defibrillator records. For practical reasons, the section only covers those commercial brands associated with the defibrillator files used in the development of the thesis: Philips, Stryker, and ZOLL. Some of the programs analyzed may not correspond to its latest, most developed version.

- Philips devices support the use of the Q-CPR assist pad (Laerdal, Stavanger, Norway; Philips Healthcare, Andover, MA, USA), which acquires both chest acceleration and force data. This information is then processed in real-time to provide CD and CCR feedback to the rescuer. The proprietary software *Event Review Pro* permits the

4 Specificity (SP) or *True Negative Rate* is a performance metric associated to binary classification problems, which indicates the proportion of correctly identified true negative class instances (in the case above, absence of chest compressions). Note that SP is not directly applicable to detection problems, as no negative class is defined [142].

analysis of Q-CPR data to generate detailed CPR quality reports. Quality metrics include: number of compressions, CD, compressions above and below recommended depth, absolute and relative NFT, CCR, incomplete releases, and duty cycle. Ayala et al. [100] proved the reliability of the CD signal in providing CCR and CCF estimates through simple peak detection. Default analysis intervals include the complete episode as well as 60 s or 30 s epochs; ad-hoc analysis intervals can also be manually specified. The latest versions of the software also enable compression analysis and limited quality reports based on TI recordings, but were not available over the course of this thesis.

- Stryker currently has no proprietary technology for CD estimation and real-time CPR feedback. Physio-Control used to have its own CPR assist pad technology *True CPR*, which measured distances by sending and receiving electromagnetic signals [146], but it is now discontinued. On the other hand, the *CODESTAT* analysis software incorporates proprietary algorithms to identify compressions in TI and generate detailed CPR quality reports. Back in 2008, Stecher et al. [113] reported a SE of 94% and a PPV of 90% in compression detection. Reported quality metrics include average and median CCR, and duration of chest compressions during analysis (which, in combination with the also reported analysis time, can be used to compute CCF or NFT). Default analysis intervals include the complete episode, one-minute epochs and CPR sections separated by pauses > 10 s; ad-hoc analysis intervals are also possible.
- ZOLL features the CPR-D-padz technology, which integrates an accelerometer in the defibrillation pads. This allows real-time CPR feedback to the rescuer as well as extensive CPR analysis with the *RescueNet CodeReview* software. Reported quality metrics include: number of compressions, mean (SD) CD, compressions above or below recommended depth, absolute NFT, CCF, mean (SD) CCR, and compressions with instantaneous CCR above or below recommendations. To compute this information, the software processes acceleration data and composes a CD-like signal, denoted as CPR waveform; no studies were identified evaluating the reliability of this signal for CPR analysis. Default analysis intervals include the complete episode and one-minute epochs. While some important

events such as the start of resuscitation or intervals with ROSC can be manually defined, fully ad-hoc analysis intervals are not supported. Recording of the TI signal is possible with adequate configuration of the defibrillator device, but *CodeReview* does not consider it for chest compression analysis.

An important limitation of most commercial software is the need for manual supervision to define the analysis interval. Defibrillator files often include signal recordings prior to and following the actual resuscitation efforts (especially if ROSC is achieved), which will be considered for analysis unless otherwise specified. While this is not critical for CCR or CD calculations, Gupta et al. [147] showed that average errors of about 15 – 20 % in CCF can be expected without manual definition of the resuscitation interval. Although the study by Gupta et al. was limited to the Stryker/Physio-Control *CODESTAT* program, similar issues were observed in other analysis software.

Note that this limitation may be extensible to many of the literature methods described in the previous section. Some of the studies explicitly mention the exclusion of ROSC [101, 113] and post-resuscitation intervals [100], and none describe any automatic procedure to deal with this type of problem. While automatic definition of the analysis interval using the first and last detected compressions may be a viable strategy and could avoid many of the errors described by Gupta et al., false positive detections outside the resuscitation interval would still result in significant CCF errors. The low CCF errors reported, coupled with the lack of information on this matter, suggest that manual definition of the analysis interval may have been the norm, and that the errors reported by most of the studies could be limited to missed or misdetected compressions within specified intervals.

In commercial software, Gupta et al. [147] associated these missed and misdetected compressions with errors of about 5% in CCF. All software mentioned in this section allows manual modification of compression instants or CPR intervals; however, this kind of review may be exceptionally time-consuming, and not feasible when working with large OHCA registries. It should be noted that Gupta et al. analyzed automatic compression marks based on the TI; errors could be smaller for more reliable signal references such as the CD.

Another direct limitation of commercial software is the inability to process files from defibrillator brands other than their own. While this may not be a problem for small/local research, the most ambitious studies and clinical trials in resuscitation are multicenter, and often include episode recordings from multiple different defibrillator models. Besides potential differences in signal acquisition that may be impossible to solve (e.g., no CD measures will be ever available for current Stryker devices), different software may also introduce additional procedural disparities, such as different metrics reported or different analysis intervals considered. The harmonization of the analyses can be extremely time-consuming, and will still suffer from all the combined limitations of the software involved.

2.2 CPR QUALITY ANALYSIS: VENTILATIONS

2.2.1 VENTILATION QUALITY METRICS

Ventilation rate (VR) is the main CPR quality metric associated with ventilation. Defined as the number of ventilations provided over time, Kramer-Johansen et al. [58] recommend VR to be reported in a minute-by-minute basis, either through direct count of ventilations or frequency analysis. Given their potential impact in survival [54, 61], both the fraction of minutes presenting hyper-ventilation (defined as $VR > 15\,min^{-1}$) and the fraction of minutes with no ventilations delivered ($VR = 0$) should also be reported.

Kramer-Johansen et al. [58] mention a number of other parameters related to ventilation, including *tidal volume*, *inspiratory time*, and *peak inspiratory pressure*. However, there were at the time important technical limitations to their measurement with monitor-defibrillators, as well as little evidence to support their clinical relevance. Recent studies suggest that tidal (insufflated) volumes could be associated with survival from OHCA. Using a procedural definition for effective ventilations of "TI fluctuations with duration $> 1\,s$ and amplitude $> 0.5\,\Omega$ [5]", Chang et al. [79] analyzed 560 OHCA episodes treated

5 TI ventilation amplitudes of $0.5\,\Omega$ have been experimentally associated with tidal volumes of about 300 mL [109, 115], which may be considered sufficient to produce gas exchange and visible chest rise [148].

with 30:2 CPR and observed significantly increased survival in those patients who received such ventilations in the majority of compression pauses. Further research may thus be required to identify and characterize other potentially relevant ventilation quality metrics besides VR.

2.2.2 Literature solutions for ventilation analysis

Capnography based solutions

Resuscitation guidelines emphasize the use of end-tidal capnography as part of advanced care due to its many applications, including indirect assessment of CPR quality, early detection of ROSC, control of tracheal tube placement, and ventilation rate monitoring [17, 36]. Capnography is the continuous measurement of CO_2 concentration in exhaled gases, and the resulting graphical signal trace is referred to as capnogram. As shown in Figure 2.5, the typical capnogram describes a characteristic four-phase pattern which enables the identification of ventilations. However, during resuscitation, chest compressions introduce artifacts in the capnogram due to pressure changes and in-

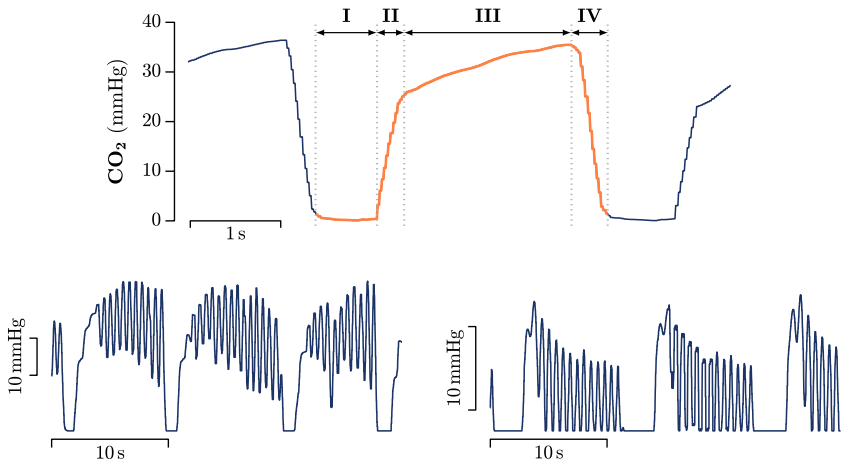


Figure 2.5. At the top, the typical capnogram waveform, with four distinguishable phases: (i) inspiratory plateau, corresponding to inspired air, generally free of CO_2; (ii) expiratory upstroke, which marks the onset of expiration and transition to alveolar gas; (iii) alveolar plateau, presenting the highest levels of CO_2; and (iv) inspiratory downstroke, which marks the onset of inspiration and transition to CO_2-free air. At the bottom, capnogram waveforms affected by chest compression artifacts.

termittent airway closure [149]. Several solutions have been proposed for the automatic detection of ventilations under such circumstances:

- Edelson et al. [106] proposed in 2010 the first automatic algorithm for ventilation rate monitoring using the capnogram. The solution was designed following a finite-state machine model. In summary, it identified inspiratory plateaus (from instant of inspiration t_{insp} to expiration t_{exp}) as sections of the capnogram below a baseline threshold $Th_{insp} = 2\,\text{mmHg}$, with duration $D_{insp,k} = t_{exp,k} - t_{insp,k}$ of 0.3 – 5.0 s, and minimum separation $D_{exp,k} = t_{insp,k+1} - t_{exp,k}$ of 0.4 s.

  The solution was evaluated using data from 37 in-hospital cardiac arrest (IHCA) episodes, all recorded with a Philips MRx monitor-defibrillator and treated with ETI. Ground truth ventilations were annotated by expert clinicians based on capnogram and TI signal data; the capnogram was deemed uninterpretable for annotation 9.7% of the time. Evaluated on a minute-by-minute basis, the solution scored a median SE and PPV of 82% and 91%, respectively. VR was underestimated by a mean (SD) of 1.6 (2.6) min^{-1}. Hyperventilation, defined as rates exceeding $10\,\text{min}^{-1}$, was identified with a median SE and SP of 86% and 100%, respectively.

- Aramendi et al. [107] proposed an alternative solution for ventilation detection, incorporating adaptive thresholding to exploit the similarity between contiguous ventilations. First, inspiration and expiration instants, t_{insp} and t_{exp}, were identified from positive and negative peaks in the first difference of the capnogram. Then, the resulting potential ventilations were discriminated based on several fixed and adaptive thresholds: A minimum inspiration duration D_{insp} of 0.3 s and a minimum distance between ventilations $D_{sep,k} = t_{exp,k+1} - t_{exp,k}$ of 1.5 s were imposed. Weighted averages from the previous N ventilation detections were computed to set minimum thresholds for: (a) the mean amplitude of the expiratory plateau A_{exp}, (b) the relative amplitude increase between inspiration and expiration Δ_r, and (c) the area under the signal during the first second of expiration S_{exp} (see Figure 2.6).

  The algorithm was developed using data from 62 OHCA and 21 IHCA episodes. Ground truth ventilations were annotated based

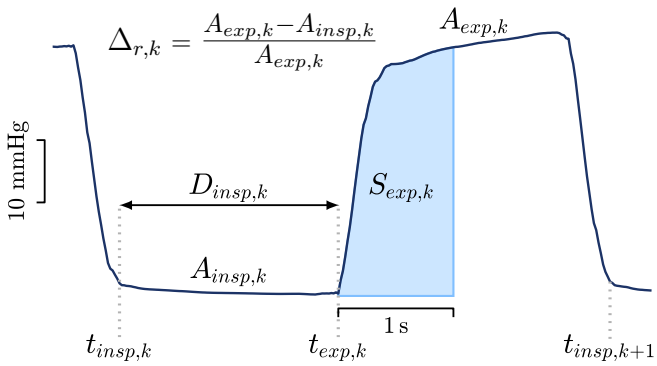


Figure 2.6. Feature calculation in the ventilation detection solution by Aramendi et al. [107]. Adapted from the original study.

on TI and capnogram recordings in the OHCA dataset, and based on air flow and volume recordings in the IHCA dataset (this is the only solution in the literature validated against a ground truth exported from gas measuring devices). A subset of 37 OHCA episodes were used to optimize the algorithm, and the rest for evaluation. Restricted to episode intervals with chest compressions delivered, ventilations were identified with median per-episode SE of 99.0% and PPV of 97.6% in the OHCA dataset, and median SE of 99.8% and PPV of 98.3% in the IHCA dataset. Considering one-minute episode segments, the 95% of absolute errors in VR estimation were within $1.8\,\text{min}^{-1}$ in OHCA data, and within $1.5\,\text{min}^{-1}$ in IHCA data. Hyperventilation, defined as one-minute segments with VR $> 15\,\text{min}^{-1}$, was identified with SE of 97.8% and PPV of 97.0% in the OHCA dataset, and SE of 99.9% and PPV of 99.7% in the IHCA dataset.

- Leturiondo et al. [108] proposed the most widely validated solution as of the start of this thesis, also following a finite-state machine model. First, the solution identified a potential expiration at $t_{exp,k}$ as an abrupt signal upstroke crossing a baseline threshold Th_{amp}. Then, potential candidates for the following inspiration at $t_{insp,k}$ and next expiration at $t_{exp,k+1}$, also crossing Th_{amp}, were iterated until both the inspiration duration $D_{insp,k} = t_{exp,k+1} - t_{insp,k}$ and expiration duration $D_{exp,k} = t_{insp,k} - t_{exp,k}$ exceeded, respectively, the minimum thresholds Th_{insp} and Th_{exp}.

The solution was developed using data from 232 OHCA episodes, in which about 50,000 ground truth ventilations were annotated based on the TI signal. The 42% (98) of episodes were labeled as distorted, presenting over one-minute of capnogram recordings affected by compression artifacts. A subset of 30 undistorted cases were used to optimize the algorithm, and the rest for evaluation. In a per-episode analysis, ventilations were identified with a median SE of 99.4% and PPV of 98.6%. In the subset of distorted episodes, the median SE and PPV were 97.4% and 95.6%, respectively. In the case of heavy artifacts (reaching the capnogram baseline, such as in Figure 2.5 bottom right), denoted as type-III, the median SE and PPV were 85.2% and 76.9%. VR was estimated with median relative error of -0.6%, $-6,1\%$ and -18.8% in clean, distorted and type-III one-minute segments, respectively. For the same groups, hyperventilation (VR $> 15\,\text{min}^{-1}$) was detected with SE of 99.5%, 95.7% and 90.9%, and PPV of 96.8%, 73.9% and 53.2%, respectively. Subsequent studies have shown that signal preprocessing techniques such as envelope detection [150] or adaptive filtering [151] may help enhance ventilation detection, achieving median SE and PPV scores over 95% even during type-III compression artifacts. Other quality metrics would also improve accordingly.

Table 1 summarizes the most relevant performance metrics for the different ventilation detection solutions described in this section.

Table 1. Performance metrics for the different capnogram-based ventilation detection solutions

Study	Ventilation detection *median (IQR)*		VR estimation *per-minute*	Hyper-ventilation minutes		
	SE (%)	PPV (%)		SE (%)	SP (%)	PPV (%)
Edelson et al. [106]	*per-minute* 82.0 (75.0 – 93.0)	91.0 (85.0 – 95.0)	*mean (SD)* -1.6 (2.6) min^{-1}	*patient median, > 10 min*$^{-1}$ 86.0	100.0	–
Aramendi et al. [107]	*per-episode* 99.1 (96.9–99.8)	97.0 (95.9 – 98.9)	*95% CI* $\pm$ 1.8 min^{-1}	*overall, > 15 min*$^{-1}$ 97.8	–	97.0
Leturiondo et al. [108]	*per-episode* 99.4 (97.8 – 100.0)	98.6 (96.4 – 99.5)	*median (IQR)** -0.6 (-1.9 – 0.0) %	*overall, > 15 min*$^{-1}$ 95.1	–	86.8

* Restricted to undistorted data. Numeric values not available for global data.

It should be noted that most capnography equipment employed in out-of-hospital resuscitation use a side-stream configuration, in which a small air sample is aspirated from the primary airway to the sensor. This introduces a time-delay of about 3.5 s between the capno-

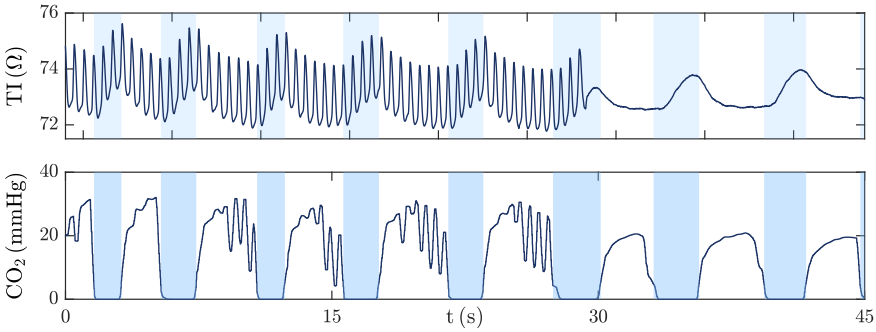


Figure 2.7. Example of ventilations in both the thoracic impedance (TI, top) and the capnogram (CO_2, bottom). The time-delay of the capnogram due to gas transport has been corrected for referencing purposes. As it can be observed, there is a close correspondence between the inspiration phases in the capnogram (shaded in blue) and the inflation phase of fluctuations in TI (associated with an increase in low-conductivity air in the lungs).

gram and other signals acquired by the monitor-defibrillator [106], which has to be corrected for reference and matching purposes. It is also important to note that the time-based capnography used in OHCA does not directly allow the measurement of air volumes [122]. Furthermore, in low blood-flow states, such as in cardiac arrest, pulmonary blood flow is the main determinant of capnogram levels [42], so indirect or surrogate measures are also mostly unfeasible.

Impedance based solutions

Ventilations are also reflected in the TI due to the poor electrical conductivity of the air blown into the lungs [113]. Impedance-based techniques to monitor respiration and estimate tidal volumes in the clinical practice have been studied for more than 50 years [152, 153]. However, it was not until 2002 that Pellis et al. [114] demonstrated that the principles of impedance pneumography — injection of a high-frequency current through the body, and continuous monitoring of voltage changes, directly proportional to impedance changes according to Ohm's Law — could be easily implemented in resuscitation using defibrillation pads. In an experiment involving five pigs, Pellis et al. observed two different components in the TI acquired through defibrillation pads: (a) a 1.4 – 20 Hz component, correlated with QRS complexes in the ECG and associated with cardiac activity (see Figure 2.4), and (b) a 0.1 – 2.0 Hz component, correlated with

capnography readings and associated with ventilation (see Figure 2.7). Later, Losert et al. [137] analyzed the reliability of defibrillator TI recordings for estimating tidal volumes in human data. A cohort of 73 patients was studied, including 20 patients in cardiac arrest for whom ventilations were measured during compression pauses. In line with previous pneumography studies [152, 154], they found a very high within-patient correlation between the amplitudes ΔZ of TI fluctuations due to ventilations and the insufflated volumes ΔV, but also a large between-patient variation of the correlation factor $\alpha = \Delta Z / \Delta V$. For a general volume estimation model applicable to all patients, the normalized standard error was 0.365, which represented errors $> 36.5\%$ for about the 30% of patients. As proposed by Valentinuzzi et al. [155], errors were smaller if accounting for patient weight, although still significant. It is unclear whether estimates of such precision would be adequate to provide real-time alarms during resuscitation [42, 137]; however, given enough data, they could still be used as part of large retrospective studies to assess the impact of ventilation [79].

Regardless of volume estimates, the detection of individual ventilations enables, in the same way as in capnography, the calculation of other quality metrics such as the VR or the occurrence of hypo- or hyper-ventilation. The detection of ventilations in TI may, however, be subject to a number of difficulties: First, TI fluctuations due to ventilation may adopt a wide range of shapes, amplitudes and durations [105]. And second, the TI signal itself is prone to artifacts due to patient and electrode motion [156, 157], which may be frequent during resuscitation. Roberts et al. [158] studied the performance of the ventilation detection algorithm internal to the MRx monitor-defibrillator. For a study cohort of 21 pediatric patients, the overall SE was of 90.5%, although it decreased to 80.0% for volumes $< 7\,\mathrm{mL/kg}$ (presumably associated with TI fluctuations of lower amplitude). A PPV of 100% was reported; however, some obvious artifacts were excluded during data preparation. The patients were also hemodynamically stable, not receiving CPR, which could have significantly limited the presence of motion noise. Recently, Aramendi et al. [109] proposed a solution to identify ventilations in compression pauses during 30:2 CPR. In summary, the method employed peak detection

and a fixed rule set to identify fluctuations with amplitude $\geq 0.5\,\Omega$ and duration $\geq 1\,s$, including pairs of overlapping fluctuations and fluctuations partially occurring during chest compressions. Validated on a study cohort of 550 OHCA patients, and tested against ground truth ventilations manually annotated by two different reviewers, the solution identified pauses with one or more ventilations with a SE of 90.8/92.4 %, similar to that reported by Roberts et al. [158], but a much a lower PPV of 85.3/82.3 %, which could be a better indicative of the motion noise present during resuscitation.

Nevertheless, the most common ventilation scenario in OHCA may be that of CC-CPR, which is the standard CPR mode after advanced airway insertion, and also considered in the guidelines [17, 44] and increasingly frequent in early phases of resuscitation. In addition to all the aforementioned problems for ventilation detection, CC-CPR entails concurrent chest compressions, which are also reflected in the TI and may obscure the ventilation waveform. Several solutions have been proposed for this scenario:

- Risdal et al. [105] proposed in 2007 the first TI-based ventilation detection algorithm tailored to CC-CPR. The solution comprised three different stages: First, a series of filters were applied to the TI in order to enhance ventilation components, including linear static band-pass filtering in the $0.06-2.3\,Hz$ range, and adaptive filtering of compression artifacts using acceleration and force as reference channels. Secondly, potential onsets of inspiration and expiration, t_{OI} and t_{OE}, were identified in the filtered signal. For this purpose, each point of the signal was characterized in terms of its second- and fourth-order polynomial fit coefficients over a centered window of 1.4 s. These coefficients were then fed to an *artificial neural network* (ANN), which returned the probabilities p_{OI} and p_{OE} of each point corresponding to a t_{OI} or t_{OE}, respectively (see Figure 2.8). Potential t_{OI} and t_{OE} were selected at the most prominent peaks of the continuous p_{OI} and p_{OE} outputs which exceeded minimum probability thresholds Th_{OI} and Th_{OE}. Finally, the discrimination of ventilations was performed according to a fixed rule set, with valid ventilation inflation times and amplitudes in the ranges $\Delta t_{min} - \Delta t_{max}$ and $\Delta z_{min} - \Delta z_{max}$, respectively. Δt_{min}, Δt_{max} and Δz_{max} were set based on the minimum and maximum

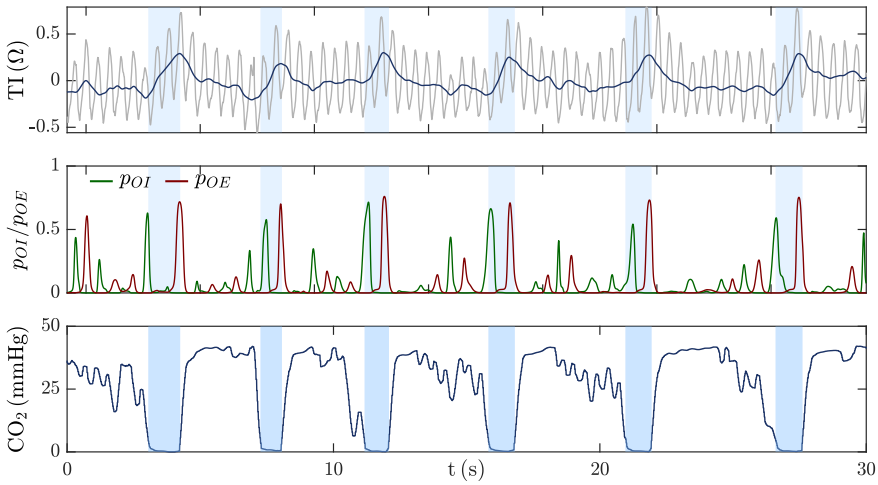


Figure 2.8. Example of the solution by Risdal et al [105]. At the top, the thoracic impedance (TI) after adaptive filtering of compression artifacts; the raw TI is also depicted in gray. At the middle, continuous probability estimates for the onsets of inspiration (p_{OI}, green) and expiration (p_{OE}, red). As it can be noted, there is a close match between the most prominent peaks in p_{OI} and p_{OE}, and the starts and peaks, respectively, of the fluctuations in TI. The time aligned capnogram (CO_2, bottom) is provided as ground truth reference; inspiration phases according to the capnogram appear shaded in blue.

values observed in the training dataset (0.2 s, 5.0 s and 8.0 Ω, approximately). Th_{OI}, Th_{OE} and $\Delta z_{min} = 0.25\,\Omega$ were optimized to maximize performance.

The solution was developed and evaluated using a leave-one-out cross-validation strategy. Recordings from a total of 30 OHCA and IHCA cases were included, and over 10,000 ground truth ventilations — both with and without concurrent compressions — were manually annotated according to the TI signal itself. Ventilation detection performance was optimized for maximum PPV at 90% SE. On a per-episode basis, the overall median (IQR) SE and PPV were 92.7 (10.9) % and 96.7 (5.4) %, respectively. Considering only episode intervals during CC-CPR, the median (IQR) SE and PPV were 90.6 (12.5) % and 97.4 (8.0) %. No analyses were conducted on VR or other ventilation quality metrics.

- Edelson et al. [106] proposed a lighter solution, in which the ANN segmentation stage was replaced by a finite-state machine model. The preprocessing stage was similar to that of Risdal et al. [105],

also including an adaptive filter for compression artifact removal using force and acceleration as reference channels. The state machine model identified waves in the filtered TI with the following specifications: amplitude of $0.25-4.5\,\Omega$, inflation time of $0.3-3.0\,s$, rise/fall rates of $0.2-5.0\,\Omega\,s^{-1}$, and peak-to-peak separation between ventilations of 1 s or more.

The solution shared study dataset with the previously described capnogram-based algorithm, comprising 37 IHCA episode records from MRx monitor-defibrillators. Ground truth ventilations were annotated based on the capnogram and TI signals; TI was deemed uninterpretable 19.5% of the time. On a minute-by-minute basis, ventilation detection was achieved with median SE and PPV of 78% and 87%, respectively. VR was underestimated by a mean (SD) of 1.9 (3.6) min^{-1}. Hyperventilation ($VR > 10\,min^{-1}$) was identified with inter-episode median (IQR) SE and SP of 81 (51 – 100)% and 100 (98 – 100) %, respectively.

- Alonso et al. [99] proposed an even lighter solution, in which adaptive filtering was also removed. Preprocessing of TI was limited to static low-pass filtering with a cut-off frequency of 0.6 Hz. Then, fluctuations potentially due to ventilation were identified through positive peak detection. Finally, discrimination of fluctuations was performed according to a fixed inflation-time threshold $D_{min} = 0.5\,s$, an adaptive amplitude threshold Z_{min} based on the amplitudes of the last N detected ventilations.

  The solution was developed and evaluated using 63 OHCA episode records (32 for optimization, 31 for validation), all acquired with a MRx monitor-defibrillator. About 17,500 ground truth ventilations were annotated based on capnogram and TI recordings. In terms of ventilation detection, the solution scored a per-episode median SE of 92.2% and a PPV of 81.0%. The mean (SD) absolute error in VR estimation was of 1.5 (1.4) min^{-1} and 3.3 (2.9) min^{-1} for full episodes and one-minute segments, respectively. In the 6.8% of minutes analyzed (inter-episode mean), hyperventilation ($VR > 15\,min^{-1}$) was either missed or falsely reported.

Table 2 summarizes the most relevant performance metrics for the ventilation detection solutions designed for CC-CPR. When disaggre-

gated data are available, results are limited to performance metrics during continuous chest compressions.

Table 2. Performance metrics for the different capnogram-based ventilation detection solutions

Study	Ventilation detection *median (IQR)*		VR estimation *per-minute*	Hyper-ventilation minutes *inter-patient mean/median*		
	SE (%)	PPV (%)		SE (%)	SP (%)	PPV (%)
Risdal et al. [105]	*per-episode* 90.6 (12.5)	97.4 (8.0)	–	–	–	–
Edelson et al. [106]	*per-minute* 78.0 (67.0 – 89.0)	87.0 (77.0 – 96.0)	*mean (SD)* -1.9 (3.6) min^{-1}	*median, VR* $> 10\,min^{-1}$ 81.0	100.0	–
Alonso et al. [99]	*per-episode* 92.2 (87.4 – 95.8)	81.0 (67.2 – 90.5)	*mean (SD)* 3.3 (2.9) min^{-1}	*mean, VR* $> 15\,min^{-1}$ * –	–	–

* Hyperventilation missed or falsely reported in the 6.8% of minutes. No SE, SP or PPV values available.

2.2.3 Ventilation analysis in commercial software

- Philips devices enable TI-based analysis of ventilations as part of the Q-CPR system. The integrated ventilation detection algorithm is presumed to be similar to that described by Edelson et al. [106], including adaptive filtering to remove compression artifacts and identify ventilations during CC-CPR. The accuracy of the algorithm has been demonstrated in absence of CPR, with SE scores above 90% for insufflated volumes $> 7\,\text{mL/kg}$ [158]. Edelson et al. reported median SE and PPV scores of 78% and 87% for IHCA data with active CPR. Defibrillator recordings may be loaded in *Event Review Pro* software to generate detailed ventilation reports, comprising: number of ventilations, VR, average inflation amplitude (in $\text{m}\Omega$), average inflation time, and number of ventilations too short in duration. As with chest compressions, possible analysis intervals include the full episode, time-epochs of 30 s or 60 s, and manually selectable ad-hoc intervals. MRx monitor-defibrillators also enable VR readings and alarms based on capnography data, meaning that some internal capnogram-based ventilation detection algorithm should be in place. However, the software does not appear to support the visualization and export of these ventilation marks, nor their use in composing CPR quality reports.
- Stryker/Physio-Control's *CODESTAT* software enables automatic annotation of ventilations based on capnography data. Although

no analogous TI-based functionality is included, the program still allows manual annotation of ventilations in the TI channel, and provides a filtered representation of the TI to facilitate this process during CC-CPR [113]. Reported metrics are limited to average and median VR over the analysis interval. Analysis interval options include the complete episode, one-minute epochs, CPR sections (separated by pauses > 10 s), and ad-hoc intervals. It is interesting to mention that, unlike with Philips or ZOLL devices/software, where raw TI data is available, *CODESTAT* exports a preprocessed, high-pass filtered version of the TI in which ventilation information may be compromised. External algorithms for TI-based ventilation analysis may be unsuited and/or require important modifications in order to work with Stryker/Physio-Control data.

- ZOLL's *RescueNet CodeReview* software enables automatic detection of ventilations using both capnography and TI data. Ventilation quality metrics are limited to the number of ventilations in each one-minute episode epoch. A fully detailed CPR analysis report is also available, which includes the timestamp and duration of each individual ventilation and could be processed externally to derive additional metrics. Ventilations detected in the capnogram and TI are reported separately. It is important to note that the TI-based detection algorithm appears to be highly sensitive to compression artifacts, with ventilation counts often exceeding 100 per minute. The newer *CaseReview* software, not available during this thesis, does not support anymore automatic ventilation detection based on TI data. It should also be noted that ZOLL defibrillator devices do not record TI signal data by default and must be configured explicitly to do so. ZOLL defibrillator files missing TI recordings are thus fairly common in OHCA data registries. The new ZOLL *X-Series Advanced* monitor-defibrillators incorporate *Real BVM Help* technology, which employs a differential pressure-based *AccuVent* flow sensor to measure tidal volumes and provide real-time VR and volume feedback. Studies on this technology, however, seem still limited to mannequin data [120, 159, 160]; further research is needed to assess its reliability in real OHCA scenarios.

2.3 Suppression of chest compression artifacts

During CC-CPR, concurrent chest compressions cause an artifact in TI which can hinder the detection and characterization of ventilations. Alonso et al. [99] showed that compressions and ventilations present limited spectral overlap, and that ventilation waves can still be identified in TI after applying a 0.6 Hz low-pass filter to the signal. Such an approach, however, may also result in excessive smoothing, losing detail of ventilations and making them harder to distinguish from other spurious fluctuations. In contrast, other authors [105, 106] have proposed the use of adaptive filters, which take advantage of known or external information about compressions in order to perform a more selective filtering. These adaptive filters may better preserve the ventilation waveform, enabling richer feature extraction both for the potential clinical interest and to support more accurate detection and discrimination solutions. The literature on adaptive suppression of compressions artifact focuses on the ECG, where it can enable rhythm analysis during CPR and help minimize interruptions. Nevertheless, most of the proposed solutions can also be applied to the TI case.

In the most accepted model, the input ECG or TI signal $s(n)$ is assumed to be the sum of a desired component $s_d(n)$ (representing the heart activity in the ECG case, or ventilation in the TI case), and a chest compression component/artifact $s_{cc}(n)$. This is a classical problem of noise/interference cancellation, in which adaptive filtering solutions provide an estimate of this compression component $\hat{s}_{cc}(n)$. The desired signal can be then approximated as

$$\hat{s}_d(n) = s(n) - \hat{s}_{cc}(n) \tag{3}$$

Proposed solutions may differ in the algorithmic approach, and in the artifact model and the information or reference signals used to estimate $\hat{s}_{cc}$. Optimization algorithms include, among others, *least mean squares* (LMS) [136, 161, 162], *recursive least squares* (RLS) [162], Wiener [163] and Kalman [164, 165, 162] filters, and matching pursuit algorithms [105, 166]. The compression artifact models relevant to this thesis can be divided into *multichannel models* and *parameterized models*.

2.3.1 Multichannel models

In these solutions, the chest compression component is estimated as a linear combination of P reference signals s_p and their time-shifted versions, following:

$$\hat{s}_{cc}(n) = \sum_{p=1}^{P} \sum_{k=-K_p}^{K_p} h_p(k;n) s_p(n-k) \tag{4}$$

where $h_p(k;n)$ represent a filter of length $2K_p + 1$, whose coefficients vary slightly over time (in terms of n). The reference signals s_p are assumed to be correlated with the compression component (as in the case of acceleration or force), and not with the information of interest. The goal of the optimization algorithm is to first find and then keep updated the filter coefficients h_p that better match the compression component; this is done according to some predefined criteria, such as the mean square error. Note that Eq. 4 describes a noncausal filter, as it uses future samples of the reference signals; if necessary, this can be avoided by setting the lower limit of the second summation to zero.

The first studies on multichannel filtering of compression artifacts were conducted by Langhelle et al. [167] and Aase et al. [163], using mixture models which combined human ECGs in VF with mechanical CPR artifacts recorded from a pig experiment during asystole. The solution proposed by Aase et al. was based on a Wiener filter, and used as references the TI and CD signals recorded from the pig. Later, Husøy et al. [166] introduced the *Multi-Channel Recursive Adaptive Matching Pursuit* (MC-RAMP) solution, which used a *matching pursuit*-based coefficient optimization and four reference channels: TI, CD, acceleration and ECG common-mode. The MC-RAMP was evaluated using a mixture model (although in this case two pigs were included, and manual CPR was performed), and obtained results comparable to the Wiener filter for much less computational cost. The MC-RAMP was adapted and tested for VF detection in real human resuscitation data by Eilevstjønn et al. [168]. Regarding ventilation detection, the filter was also adapted by Risdal et al. [105] to identify ventilations in the TI during CC-CPR; in this case, acceleration and force were used

as reference signals. A similar approach was later used by Edelson et al. [106], and is presumed to be implemented in Philips defibrillator devices and software.

2.3.2 Parameterized models

After development of multichannel filters, efforts began to be made to reduce or eliminate the need for reference signals which might not always be available. Within the new approaches, the major line of research considered that the compression artifact could be modeled from a limited set of parameters, most notably the frequency of chest compressions. Early solutions [169, 164, 170] were based solely on the ECG, with the frequency and harmonic content of compressions estimated through spectral or time-frequency analysis. Performances, however, were inferior to those of multichannel solutions.

Later, Irusta et al. [161] proposed to construct the instantaneous compression frequency from the chest compression instants already identified using a CPR assist pad. While this did not eliminate the need for reference signals, it could be implemented with minimal hardware modification of the defibrillator device. The compression frequency was assumed constant during a given compression, and variable between them, so for chest compressions at sample-instants n_i, the instantaneous phase $\phi(n)$ could be given as

$$\phi(n) = \frac{2\pi}{(n_{i+1} - n_i)}(n - n_i) + 2\pi i \qquad n_i \leq n < n_{i+1} \tag{5}$$

which represents a linear phase-change of 2π between pairs of contiguous compressions. In addition, they proposed a quasi-periodic artifact model, based on a truncated Fourier series of N harmonics, of the form:

$$\hat{s}_{cc}(n) = A(n) \sum_{k=1}^{N} a_k(n) cos(k\phi(n)) + b_k(n) sin(k\phi(n)) \tag{6}$$

where $A(n)$ is an indicator of compressions, introduced for stability ($A(n) = 1$ during compressions and $A(n) = 0$ during pauses, with smooth transitions). The final solution was implemented using an

LMS optmizer, whose goal was to find and keep tracked the optimal in-phase and quadrature coefficients $a_k(n)$ and $b_k(n)$, and yielded performances comparable to those of the four-channel MC-RAMP. Ruiz et al. [165] achieved a similar performance using a Kalman filter configuration. Later, Aramendi et al. [141] demonstrated de accuracy of the model when compression instants were detected in the TI; this effectively eliminated the need for reference signals outside of those acquired through defibrillation pads.

Introduced by Aramendi et al. [136] and further studied by Isasi et al. [162], a very particular variant of the model by Irusta et al. [161] may be applied during mechanical chest compressions. Mechanical compression devices operate at a very stable, fixed known frequency f_{cc} (e.g., $f_{cc} = 101.7\,\text{min}^{-1}$, or 1.695 Hz, for the LUCAS-2 device), so the fundamental instantaneous phase for the in-phase and quadrature terms of the model can be given as

$$\phi(n) = \Omega n = 2\pi(f_{cc}/f_s)n \tag{7}$$

where f_s is the sampling frequency of the signal under study, and both f_{cc} and f_s are given in hertzs. The artifact model then takes the form

$$\hat{s}_{cc}(n) = A(n)\sum_{k=1}^{N} a_k(n)cos(k\Omega n) + b_k(n)sin(k\Omega n) \tag{8}$$

in which $A(n)$ is the only term including compression information other than the nominal frequency of the device. If the optimization algorithm converges fast enough, and/or if compression residuals around pause transitions are not critical (note that pauses should be minimized during mechanical compressions [36]), then the $A(n)$ term may also be dropped from the model, and adaptive filtering of compression artifacts becomes possible without any sort of reference signal or information.

2.4 Machine learning for ventilation detection

Machine learning is a branch of artificial intelligence which focuses on the development of models and algorithms that can find and learn patterns in data, and make predictions or decisions without explicit programming. Machine learning algorithms have been successfully applied to many OHCA related problems, including discrimination of shockable rhythms [171, 172, 173], shock outcome prediction [174, 175], detection of spontaneous pulse [124, 176, 177], or multi-class heart rhythm classification [123, 178, 179]. Their use for CPR analysis, however, has been minimal. Most chest compression and ventilation detection solutions in the literature have been based on prespecified rule sets, typically in the form of finite-state machines or fixed or adaptive thresholds. For ventilation detection, only the solution by Risdal et al. [105] incorporated a machine learning algorithm — an *artifical neural network* (ANN) —, but still based its final decision on fixed amplitude and duration thresholds; such an approach could have limited the predictive potential of machine learning. This thesis explores the use of machine learning algorithms as final elements within ventilation detection solutions.

2.4.1 The classification problem

A classification problem involves assigning an input data instance, characterized by a vector $\boldsymbol{x} \in \mathbb{R}^K$ of K features, to a predefined class or category, denoted by a label $y \in C$, with $C = \{c_1, c_2 ..., c_D\}$ the output space or set of D possible classes. Given a set of N labeled examples $\{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2) ..., (\boldsymbol{x}_N, y_N)\}$, known as training set, the main goal of classification is to learn a function $f(\boldsymbol{x}, \boldsymbol{\omega}) : \boldsymbol{x} \rightarrow C$ that maps input features to the corresponding classes. This is generally achieved by finding the classifier parameters $\boldsymbol{\omega}$ that minimize some loss or error function $\mathcal{L}$ for the training set. Once the parameters $\boldsymbol{\omega}$ are defined, the model can be used to classify new unlabeled data instances by applying the function f to their input features.

At its core, ventilation detection is a detection problem, in which there is only one defined class, the ventilation [142]. In practice, and as far as this thesis is concerned, the detection problem is artificially translated into a classification problem through the prior detection

of TI fluctuations potentially due to ventilation. The output space C for these fluctuations is limited to two possible classes: ventilation ($y = 1$) and not-a-ventilation ($y = 0$). Class labels were assigned based on the correspondence between TI fluctuations and ground truth ventilations annotated in the capnogram.

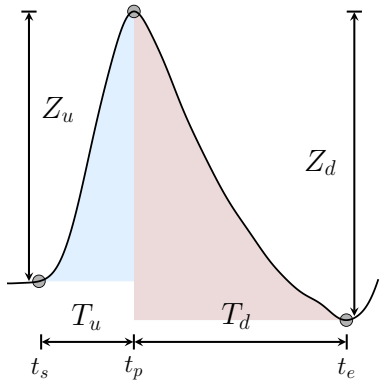


Figure 2.9. Example of typical ventilation features. Given a TI fluctuation due to ventilation (free of compression artifacts), with peak position at instant t_p, and start and end positions at instants t_s and t_e, respectively, the inspiration/inflation or upwards phase of the ventilation (shaded in blue) spans from t_s to t_p and can be characterized by its amplitude Z_u and duration T_u. Similarly, the expiration or downwards phase (shaded in red) spans from t_p to t_e and can be characterized by its amplitude Z_d and duration T_d.

One of the major challenges in classification problems is defining the set of features $\boldsymbol{x}$ to characterize input data. Expert knowledge of the problem at hand is required to design and select individual features x_i which may hold discriminative power. For widely studied problems, such as those involving ECG processing and heart rhythms, a plethora of different features have been proposed in the literature. In the case of ventilation detection, however, little is known besides the typical measures considered for thresholding, i.e., amplitudes and durations (see Figure 2.9). In addition to these, this thesis explores orthogonal polynomial approximations and central moments as potential features to characterize ventilation morphology.

2.4.2 Classification algorithms

Different classification algorithms employ different mapping functions f and different procedures to optimize model weights $\boldsymbol{\omega}$. The

most common example of a classifier is probably *logistic regression*, which is applied to binary classification problems ($C = \{0, 1\}$), and estimates the probability of an input data instance $\boldsymbol{x}$ corresponding to the positive class following

$$p(y = 1) = \frac{1}{1 + e^{-(\omega_0 + \sum w_k x_k)}} \tag{9}$$

The final class estimate $\hat{y}$ is then obtained by probability thresholding (e.g., $\hat{y} = 1$ if $p \geq 0.5$). The weights $\boldsymbol{\omega}$ of a logistic regression model are typically set so as they minimize a *binary cross-entropy* or *log-loss* function $\mathcal{L}$ of the form

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \Big[y_i \cdot log(p(\boldsymbol{x}_i)) + (1 - y_i) \cdot log(1 - p(\boldsymbol{x}_i)) \Big] \tag{10}$$

over the N training set instances $(\boldsymbol{x}_i, y_i)$. This is accomplished using *gradient descent* or other iterative optimization methods. Note that p, and thus $\mathcal{L}$, are functions of $\boldsymbol{w}$ according to (9).

Logistic regression is an example of a *linear classifier*, in the sense that it establishes a linear decision boundary in the input feature space (e.g., $p \geq 0.5 \longleftrightarrow \omega_0 + \sum w_k x_k \geq 0$). More complex decision boundaries can be obtained using *kernel methods* [180] — non-linear transformations of input data into higher-dimension spaces where the problem can be approached through linear algorithms — and *non-linear classifiers*. Examples of non-linear classifiers may include ANNs and *support vector machines* [181].

Of special interest to this thesis work is the Random Forest (RF) classifier [145]. The RF is an *ensemble learning* method — a method that combines the predictions of several base classifiers — consisting of multiple nearly uncorrelated decision trees. The term "random" may refer to two distinct aspects of its training procedure: First, individual trees are grown using *bootstrap aggregating* or *bagging* [182], that is, each tree is trained using only a random subset of N_b data instances, sampled with replacement from the whole training set of N. And second, each decision split at tree nodes is defined considering

only a random subset of input features (typically of $\sqrt{K}$, with K the dimension of the input feature space). All this helps decorrelate the trees, resulting in a final classifier that generalizes better to unseen data and is also robust against outliers and weak predictors/features. RF classifiers have been successfully applied in resuscitation-related problems such as spontaneous pulse detection [124] or rhythm classification [123, 172].

2.4.3 The sequence labeling problem

On occasion, temporal or other dependencies may exist between data instances which can play a critical role in predicting class labels. In the ventilation detection problem, for example, some degree of periodicity and similar morphologies are to be expected between neighboring ventilations. The odds of an average fluctuation being a ventilation may also not be the same if surrounded by other much larger ones than if not. Having access to information not only about a given data instance but also about its context may help the classifier make much more accurate predictions.

A sequence labeling problem is a particular type of time-series classification problem where a sequence of M input data instances $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2 ..., \boldsymbol{x}_M\} \in \mathbb{R}^{d \times M}$ has to be assigned to a sequence of same length M class labels $Y = \{y_1, y_2 ..., y_M\} \in C^M$. This assignment is not approached individually, but jointly through a mapping function $f(\boldsymbol{X}, \boldsymbol{\omega}) : \boldsymbol{X} \rightarrow C^M$ which may capitalize on context information. Algorithms particularly well-suited for this kind of problem include *hidden Markov models*, *conditional random fields*, and *recurrent neural networks* (RNN).

2.4.4 Model evaluation

In machine learning, model evaluation involves assessing the performance of a trained model on unseen data. That is, the data used to evaluate the model (the *test set*) should never mix with that used to train it (the *training set*). While this is general convention in machine learning, it is especially important for highly *complex* [6] models that

6 In a general sense, *model complexity* refers to the degrees of freedom of a trainable model, and is often measured as the number of adjustable parameters/weights.

could be able to capture specific details of individual data instances. In biomedical problems, this training/test separation is usually established at case/patient level, as different data instances from the same case/patient may present common particularities that may be learned by the model.

There are multiple ways to divide a dataset into training and test sets. The most common approach is to define a *hold-out* test set, a portion of the whole data which is kept away for evaluation purposes. In some cases, this approach may be particularly desirable, such as for evaluating how a given model generalizes to data from different EMS agencies or defibrillator devices. When important differences in the data exist (such as highly varying levels of noise), however, a randomly selected hold-out test set can also lead to overly arbitrary and difficult to interpret results. In such cases, it may be reasonable to generate multiple train-test data splits, and to report the statistical distribution of the observed performances. Another alternative is the use of *K-fold cross-validation* (CV), in which the data is divided into K subsets or folds and K train-test processes are carried out; on each of these processes, a distinct fold acts as test set, and the other $K-1$ as training set. Although computationally more expensive, CV enables the use of all data for testing purposes while avoiding data leakage. Notice that the generation of multiple K-fold partitions and statistical representation of results is also possible in CV. Given that the complete test set (the whole data) remains the same in all splits, performance variations are usually smaller than in the hold-out case, and more closely related to the *consistency*[7] of the model.

Regarding performance quantification, there are dozens of different performance metrics tailored to the many types of problem that can be addressed through machine learning. In detection problems, the main performance metrics are SE, which indicates the proportion of target instances detected, and PPV, which indicates the proportion of

---

Higher complexity implies that the model will be able to capture more complicated patterns in the data. However, if training data is limited, a high-complexity model may also end up learning particular details of the training samples. This leads to poor generalization and is usually referred to as *overfitting*.

7 *Consistency* refers to the ability of a machine learning model to reproduce an output for the same input across different trained instances of the model.

detections which actually corresponded to target instances. Overall performance is often reported using the F_1-*score*, computed as the harmonic mean of SE and PPV

$$F_1 = 2\,\frac{SE \cdot PPV}{SE + PPV} \tag{11}$$

Note that, even if disguised as binary classification, ventilation detection is in essence a detection problem and should be reported as such. Some common classification metrics such as the SP (the proportion of negative class or not-a-ventilation instances correctly identified) are of no practical interest except for their impact upon other metrics (in this case, the PPV). The final performance metrics should also account for any ventilation not reaching the classification stage (e.g., because the corresponding fluctuation was not identified).

2.4.5 Nested evaluation and feature selection

Development of machine learning solutions often involves design decisions for which a performance score of reference is needed. The most common approach to this scenario is nested evaluation, that is, the further division of training data into inner training and test (also known as *validation*) sets. In this way, the inner training set can be used to fit multiple potential models, whose expected performance can then be assessed using the validation set. Notice that using the final test set to take performance references would indirectly expose it to the model and should generally be avoided. If needed, nested evaluation schemas may include multiple layers or levels of nesting, each splitting the training set of the previous layer. Both hold-out and CV strategies are possible, although the latter is far more common.

Typical procedures which may require nested evaluation include *hyper-parameter tuning* and *feature selection*. *Hyper-parameters* denote model variables, other than trainable weights, that are configurable by the designer, such as the number of trees in a RF classifier. *Feature selection* implies the selection of an optimal reduced subset of input features; this may help reduce model complexity, improving generalization and preventing overfitting. It may also help reduce the computational cost of the solution, as less signal features have

to be calculated. Most feature selection techniques can be classified into three different categories:

- *Filter methods* rely on statistical measures between input features and outputs/other inputs in order to assign an importance score to each feature; the features with lowest scores are then removed from the model according to some predefined criteria. Filter methods are independent of the machine learning algorithm used, and can generally be applied without a nested validation architecture. A well-known example of filter method is the *Minimum Redundancy Maximum Relevance* algorithm [183].
- *Wrapper methods* rank combinations of features based on the performance observed in a machine learning model of reference, which does not necessarily have to match the algorithm used in the final problem. Wrapper methods tend to produce better results than filter ones, but are also computationally more expensive and may require of a nested evaluation schema (especially if the reference model is complex and prone to overfit). The most straightforward wrapper method, although usually unpractical, is *exhaustive search*, which ranks every possible combination of features. More efficient approaches reduce the feature space of search by applying different heuristics, which may be deterministic, such as in *recursive feature selection/elimination*, or random, such as in *genetic algorithms* [184].
- *Embedded methods* are selection methods which are inherently applied by some machine learning algorithms as part of their training process. This is the case of the Lasso-regularized logistic regression [185], or of decision trees, which employ measures such as the *Gini impurity* to select the feature to rule each decision split.

A particular case of feature ranking/selection can be observed for bagging methods such as the RF. Given that each individual tree is trained using only a subset N_b of the training set N, another subset N_b^C exists for each tree, referred to as *out-of-bag* (OOB) sample, which is unseen to the trained tree. Thus, the OOB instances can be used to evaluate the model and rank features without the need of a nested evaluation schema and without exposing the final test set. A feature importance score typically calculated in this scenario is the *permuted OOB error* [145], which measures the performance loss of the model

when the values of a given feature are randomly permuted over all OOB data instances. The resulting feature ranking can then be used to perform feature selection in many different ways, including direct removal of features with no significant impact in performance, or as reference within simplified recursive selection/elimination heuristics. Notice that the *permuted error* is not exclusive to bagging methods; the metric itself can be calculated for almost any machine learning algorithm, although nested evaluation may be necessary to assess performance loss.

2.4.6 Deep learning models

In the last decade, deep learning solutions have gained popularity and outperformed classic machine learning algorithms in many different tasks. In general, deep learning models eliminate the need for explicit feature engineering, and are instead fed with raw data (such as images or signal segments) from which they automatically learn and extract relevant features. In exchange, they usually involve an extremely large number of trainable weights, and may require much larger amounts of training data to prevent overfitting. Deep learning algorithms have been successfully applied in resuscitation for problems such as discrimination of shockable rhythms [125, 186] or pulse detection [187].

Of especial interest to this thesis work are *segmentation* networks. A one-dimensional segmentation problem can be seen as a limit case of sequence labeling, in which each time-sample of the input signal(s) is assigned a class. This is in part similar to the ANN configuration used by Risdal et al. [105] to locate inspiration and expiration onsets. However, segmentation networks are generally used not to identify individual points, but to label more extensive regions of the input, such as specific objects in an image, or the collections of samples that constitute ventilations in a TI signal segment. One typical example of deep learning segmentation network for biomedical imaging is the U-Net [188].

3 HYPOTHESIS AND OBJECTIVES

This thesis aims to cover several knowledge gaps in the automatic annotation and analysis of CPR, especially ventilation, in large OHCA datasets. Two main hypotheses were considered: First, that machine learning, and specifically context-aware algorithms, could contribute to design solutions for ventilation detection in TI. And second, that integrating different methods and clinical data could enable device-unconstrained CPR quality analysis. The following objectives were defined:

- **Objective 1: Development of algorithms for TI-based ventilation detection in the CC-CPR scenario.** Most analyses of ventilation in OHCA are currently performed using end-tidal capnography. However, capnography is usually not available until late phases of resuscitation. Moreover, recent studies of 30:2 CPR have highlighted the potential importance of tidal volumes in outcomes, a magnitude that can not be measured through capnography but can be estimated from TI. Accurate TI-based ventilation detection algorithms are needed to analyze the more common scenario of CC-CPR. Two algorithms were considered:

  – *Ventilation detection during mechanical CPR.* Mechanical compressions produce an artifact of fixed known frequency, enabling the use of dedicated narrow-band artifact suppression filters. Mechanical compression devices are also firmly secured to the patient, which could minimize movement artifacts. Therefore, minimally corrupted ventilation waveforms could be recovered,

enabling the extraction of detailed features both for clinical interest and high-performance machine learning discrimination.

  - *Ventilation detection during manual CPR.* Manual compressions imply irregular and highly variable artifacts which challenge the design of ventilation detection algorithms. Although several solutions have been proposed for this scenario, all of them rely on amplitude-based discrimination of individual ventilations. Context-aware machine learning algorithms could significantly improve ventilation detection in noisy and/or low amplitude scenarios.

- **Objective 2: Design and validation of methodologies for chest compression analysis in large and heterogeneous datasets.** While many accurate solutions for compression detection have been proposed, most have been validated on limited and uniform datasets, and do not automatically address important issues such as integration of different data sources, temporary signal loss, or definition of the analysis interval. The combination of algorithms, addition of new procedures, and integration of clinical data could represent a step forward towards a fully automated CPR analysis of large OHCA datasets.
- **Objective 3: Analysis of CPR quality in the PART dataset.** The PART trial aimed to compare the effectiveness of ETI and LT initial airway strategies, but did not incorporate CPR quality metrics. As a case study, the algorithms and methodologies developed in this thesis as part of the previous objectives could be applied to PART defibrillator and clinical data, with the following sub-objectives:
  - Comparison of chest compression quality metrics between ETI and LT strategies.
  - Comparison of ventilation rates and ventilation amplitudes in TI between ETI and LT airway devices.

4 DATA MATERIALS

This chapter describes the data materials used in the development of the thesis work. The opening section provides a brief overview of the Resuscitation Outcomes Consortium (ROC) [189], from which all data originated. Following sections describe the individual datasets within the ROC OHCA registry that were considered in the thesis: the D-FW database and the PART database.

4.1 The Resuscitation Outcomes Consortium (ROC)

Funded in 2004 by the National Heart Lung and Blood Institute, the Resuscitation Outcomes Consortium (ROC) is a North American research network focused on the study of the treatment and outcomes of OHCA and life-threatening trauma. The ROC network consists of 11 regional sites in the United States and Canada and one data-coordinating center (see Figure 4.1). More than 250 EMS agencies participate in the network, serving a total population of around 24 million [189].

The first major task within the ROC program was the design and creation of a unified epidemiological registry (*Epistry*), defining the data and variables to be collected for future studies and RCTs. Two different Epistries were considered for trauma [190] and OHCA [117]. The cardiac arrest Epistry was conceived to include all individuals (from all ages) experiencing OHCA and attended — although not necessarily treated — by EMS within the enrolled sites. The variables considered comprised information on the patient, event, EMS characteristics and hospital outcomes [116]. When possible, variables

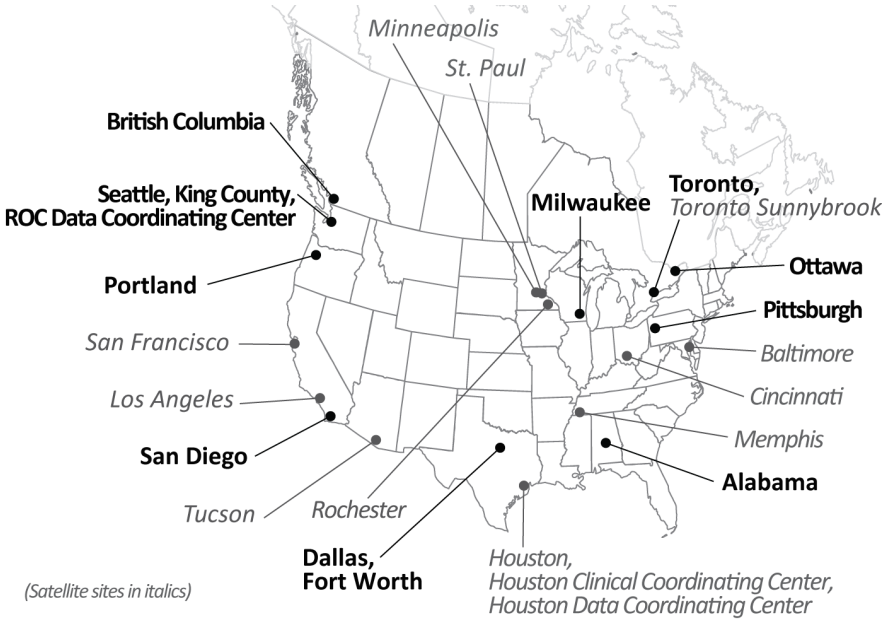


Figure 4.1. Regional sites within the Resuscitation Outcomes Consortium (ROC).

were defined following the previous Utstein [1] and National EMS Information System (NEMSIS) templates [117]. Unlike these prior templates, the Epistry also considered CPR quality variables, which could be derived from defibrillator recordings [117]; all participating sites demonstrated the ability to capture defibrillator recordings for more than 80% of treated patients who received CPR [43]. With an estimated incidence of 17,500 EMS-attended OHCA cases covered per year, the ROC Epistry constitutes one of the largest cardiac arrest registries worldwide [191].

Since the program started, ROC investigators have designed and coordinated four large OHCA-related multi-center RCTs: the *Prehospital Resuscitation Impedance Valve and Early Versus Delayed Analysis* trial (PRIMED) [192, 193]; the *Continuous Chest Compressions Versus Standard CPR* trial (CCC) [75]; the *Amiodarone, Lidocaine or Placebo Study* (ALPS) [194]; and the *Pragmatic Airway Resuscitation Trial* (PART) [92]. In addition, numerous epidemiological [116, 195] and CPR quality-related [53, 59] retrospective studies have been conducted supported by Epistry data.

4.2 The D-FW database

The D-FW database comprised the defibrillator electronic files and clinical Epistry information from 1,118 adult OHCA episodes (1,135 files) treated by EMS between July 2013 and December 2014 in the Dallas-Fort Worth area (Texas, USA). All defibrillator files were acquired using a Philips MRx monitor-defibrillator. Capnogram, CD and TI recordings were available in 1033 (1039), 1087 (1103), and 1092 (1093) cases (files), respectively. A distinctive feature of the D-FW database was the extensive use of mechanical compression devices (namely, LUCAS piston-driven devices), with 583 (593) cases (files) involving mechanical compressions according to Epistry data.

4.3 The PART database

The second and main database used in the thesis comprised the defibrillator files and corresponding clinical information from a subset of the cases enrolled in the ROC PART trial (NCT02419573) [92, 196]. The PART trial was a multi-center cluster-crossover RCT comparing the effectiveness of initial ETI and KING laryngeal tube advanced airway insertions. It enrolled a total of 3,004 adult non-traumatic OHCA patients, treated by EMS between December 2015 and November 2017, and with anticipated need of advanced airway management. The trial involved 27 EMS agencies from five different ROC regional sites: Alabama (AL), Dallas (DAL), Milwaukee (MLW), Pittsburgh (PGH) and Portland (PTL).

A subset of 2,472 cases (2,680 defibrillator files) were available for the thesis. The files corresponded mostly to monitor-defibrillators — 39 files corresponded to AEDs —, including devices from the three major commercial brands: Philips (Heartstart MRx), Stryker/Physio-Control (LP12, LP15) and ZOLL (X-Series, E-Series). Table 3 shows the number of cases (files) associated to each brand and ROC site. Table 4 shows the number of cases (files) of each brand including TI, CD and capnogram (CO_2) recordings. No mixed-brand cases were identified.

As per the specific RCT protocol, EMS personnel registered additional information on airway management not typically available in

Table 3. Cases (files) in the PART dataset by defibrillator brand and ROC regional site

Brand	Total	AL	DAL	MLW	PTL	PGH
Philips	976 (979)	–	812 (815)	–	164 (164)	–
Stryker	439 (439)	218 (218)	52 (52)	–	45 (45)	124 (124)
ZOLL	1057 (1262)	–	154 (154)	825 (1030)	78 (78)	–
Any	2472 (2680)	218 (218)	1018 (1021)	825 (1030)	287 (287)	124 (124)

Table 4. Cases (files) in the PART dataset by defibrillator brand and signal availability

Brand	Total	TI	CD*	CO_2
Philips	976 (979)	933 (935)	928 (930)	941 (942)
Stryker	439 (439)	410 (410)	–	75 (75)
ZOLL	1057 (1262)	11 (11)	1006 (1200)	962 (973)
Any	2472 (2680)	1354 (1356)	1934 (2130)	1978 (1990)

* Implies acceleration recordings for ZOLL, and both acceleration and force recordings for Philips.

Epistry data, including start/end timestamps, airway type, and outcome (success/failure) for every advanced airway insertion attempt. On top of regular Epistry information, these new variables were also available for the thesis, enabling CPR quality analyses restricted to specific airway groups and airway-related resuscitation intervals.

5 RESULTS

This chapter summarizes the studies conducted in relation to the objectives described in chapter 3. For the most part, the main focus is on the final studies published in JCR journals [197, 198, 199, 200, 201, 202]. Preliminary, side and summary results presented in different conferences are referenced in their respective sections and may be briefly discussed. All journal and key conference contributions [203, 204, 205] are included verbatim in Appendix A.

5.1 Results related to objective 1

Acquired through defibrillation pads, TI represents an alternative for ventilation analysis in resuscitation which could also help overcome some of the major limitations of capnography (i.e., typically late availability and inability to estimate tidal volumes). In this thesis work, TI-based ventilation detection solutions were developed for the CC-CPR scenario, where chest compression artifacts together with motion noise may severely hinder the identification of ventilations.

Two different solutions were considered: First, a solution for mechanical CPR, where the regularity of the compression artifact enables the use of dedicated filters, and where both detailed characterization and high-performance detection of ventilations could potentially be achieved. And second, a solution for the more general case of manual CPR, where the addition of context-aware classification could help identify ventilations under noise conditions and outperform previous solutions in the literature. Although with important differences in the individual stages, the general structure of the solution was

similar in both cases, comprising: (a) preprocessing of the TI signal, including adaptive filtering of compression artifacts, (b) detection and characterization of TI fluctuations potentially due to ventilation, and (c) machine learning discrimination of ventilations from other spurious fluctuations (see Figure 5.1).

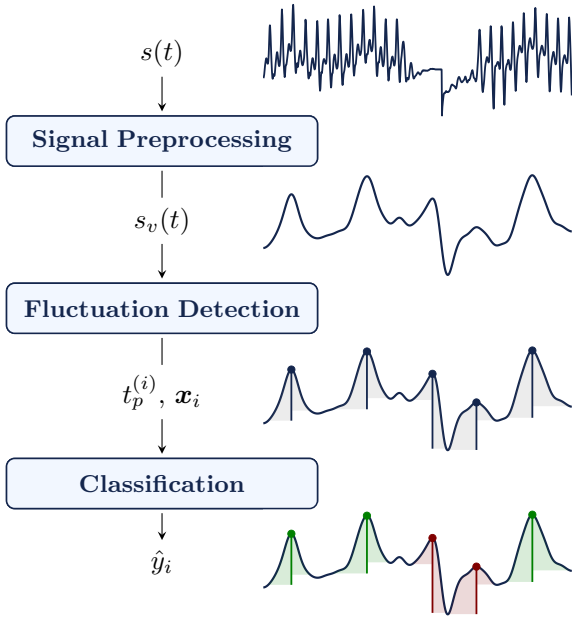


Figure 5.1. General block diagram for the ventilation detection solutions. First, the thoracic impedance $s(t)$ is preprocessed, including filtering of compression artifacts, to obtain the ventilation component $s_v(t)$. Then, fluctuations potentially due to ventilation (shaded in gray) with peak position at t_p are identified and characterized by a feature vector $\boldsymbol{x}$. Finally, this information is fed to a machine learning classifier to discriminate between actual ventilations (shaded in green) and other fluctuations (shaded in red).

Each of the solutions resulted in a JCR publication, namely $J1_1$ [197] and $J1_2$ [198], and are described in the following. Intermediate results were presented in international [203, 204, 206, 207] and national [208] conferences.

Finally, initial experiments were also conducted on the application of deep learning segmentation networks in the detection of ventilations during manual CC-CPR. The results, presented in a national conference [205], are summarized at the end of the section.

5.1.1 $J1_1$: Automatic detection of ventilations during mechanical cardiopulmonary resuscitation

For this study, a subset of 567 episodes from the D-FW database were initially considered; those including both TI and capnogram recordings, and in which a LUCAS mechanical compression device had been used according to clinical annotations. Episode intervals ≥ 60 s during mechanical compressions, with no pauses > 20 s, and with concurrent TI and interpretable capnogram signals were then selected. Mechanical CPR was identified by assessing a stable CCR of 100 min^{-1} [100]. Ventilations were annotated in the capnogram (first automatically [107], then manually reviewed), and used as ground truth for model training and evaluation; the time-delay of the capnogram (median of 3.3 s) was also manually corrected. The final dataset included data from 423 episodes, with a median (IQR) of 13 (8 – 19) analyzable minutes, and of 72 (43 – 108) ground truth ventilations. Figure 5.2 shows an example of all the signals involved in the development and execution of the ventilation detection solution.

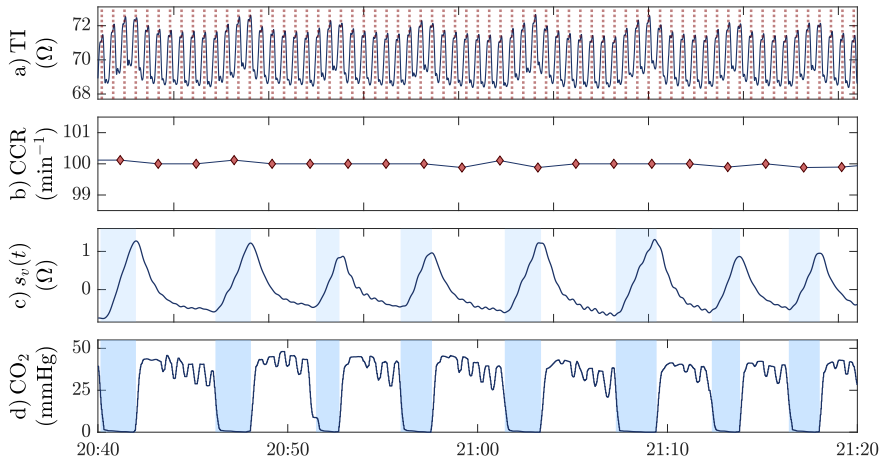


Figure 5.2. Example of the signals involved in the ventilation detection solution: (a) thoracic impedance (TI), with chest compression instants depicted in red; (b) chest compression rate (CCR), computed every 2 s over a moving window of 5 s, and used to identify mechanical chest compressions; (c) ventilation component $s_v(t)$ of the TI, obtained after the preprocessing stage; and (d) time-aligned capnogram (CO_2), used to annotate ground truth ventilations (shaded in blue) for training and evaluation.

The solution was based on the three-stage architecture shown in Figure 5.1, with the following implementation details:

- *Signal preprocessing:* The raw TI, originally acquired with sampling frequency of 200 Hz, was first resampled to $f_s = 50$ Hz and band-pass filtered in the 0.06 – 2.5 Hz range. Then, an adaptive filter was applied to suppress chest compression artifacts. Based on previous models for mechanical artifacts in the ECG [136, 162] (see Section 2.3.2), the compression artifact was estimated as

$$\hat{s}_{cc}(n) = \sum_{k=1}^{N} a_k(n)cos(k\Omega n) + b_k(n)sin(k\Omega n) \tag{12}$$

  with $\Omega = f_{cc}/f_s$, and $f_{cc} = 100\,\text{min}^{-1} = 1.667$ Hz the compression frequency of the LUCAS device. An LMS algorithm was used to optimize the coefficients a_k and b_k. A number of harmonics $N = 3$ and an LMS step-size $\mu = 0.15$ were considered.

- *Fluctuation detection:* First, the largest positive peaks in the filtered TI with minimum separation of 1.5 s were located a $t_{p,i}$. Then, a heuristic procedure was used to identify the start and end points of fluctuations, $t_{s,i}$ and $t_{e,i}$. Inspiration and expiration durations in the 0.45 – 5.5 s range were considered.

  Final fluctuations were characterized by a vector $\boldsymbol{x}_i$ of 14 features. The first four corresponded to the amplitudes (Z_u, Z_d), and durations (T_u, T_d) of the inspiration and expiration phases (see Section 2.4.2). In addition, central moments of first to fifth order m, $\mu_{u,m}$ and $\mu_{d,m}$, were computed for both phases: Let $\boldsymbol{p}_u$ be a vector of length L_u containing the signal samples of the inspiration phase, shifted in amplitude so the lowest sample (the start of inspiration) lies at zero, and then scaled so the sum of all samples equals one (as in a probability distribution), $\mu_{u,m}$ were computed as

$$\mu_{u,m} = \sum_{l=0}^{L_u-1} \boldsymbol{p}_u(l) \cdot (l/L_u)^m \quad \text{for} \quad m = 1, ..., 5 \tag{13}$$

  An analogous procedure was used to compute the moments of the expiration phase $\mu_{d,m}$.

- *Classification:* The feature vectors $\boldsymbol{x}_i$ were fed to a RF classifier to discriminate ventilations ($\hat{y}_i = 1$) from other spurious fluctuations ($\hat{y}_i = 0$). A number $B = 100$ of trees, an in-bag-fraction N_b/N of 0.5, and a default $\sqrt{K}$ number of predictors per split were selected as model hyper-parameters.

Both for training and evaluation, fluctuations were labeled as true ventilations ($y_i = 1$) when they matched a ground truth ventilation in the time-aligned capnogram, and as $y_i = 0$ otherwise. The solution was evaluated in terms of the typical detection performance metrics: SE, PPV and F_1. A patient-wise 10-fold CV strategy was followed, with fold assignment quasi-stratified so each fold included a similar number of ground truth ventilations. A recursive feature elimination procedure [1] was performed within each fold to reduce model complexity. The evaluation process was repeated using 20 different CV partitions, with results averaged, to minimize any data partition bias. During training, ventilation samples were weighted patient-wise to avoid biases due to imbalanced amounts of ventilations per patient.

The performance of the full solution was found very stable for ≥ 6 features, the six most important features (in approximate descending order) being: Z_u, $\mu_{d,1}/\mu_{d,2}$ (interchangeable), T_d, T_u, $\mu_{u,5}$, and Z_d (see Figure 5.3). For the six-feature model, the patient-averaged SE, PPV and F_1 were 96.3%, 96.4% and 96.3%, respectively. By contrast, the model based solely in Z_u (the main feature used to discriminate ventilations in the literature [99, 105, 106]) scored a much lower F_1 of 91.9%; this highlights the potential of machine learning classification for ventilation detection solutions. The partition-wise IDR was below 0.2% for all the performance metrics, showing the robustness of the solution to varying training data.

Broken down by patient, the F_1 was above 95% and 98% for the 77.1% and 49.4% of patients, respectively. Notice that these performances are similar to those reported in the literature for capnogram-

1 Starting from the full set of 14 features, the RF model was trained and the feature importance computed. Then, the least important feature was removed from the feature input space, and the process repeated until a single feature remained. The permuted OOB error of the RF was used to assess importance, so no nested CV was required.

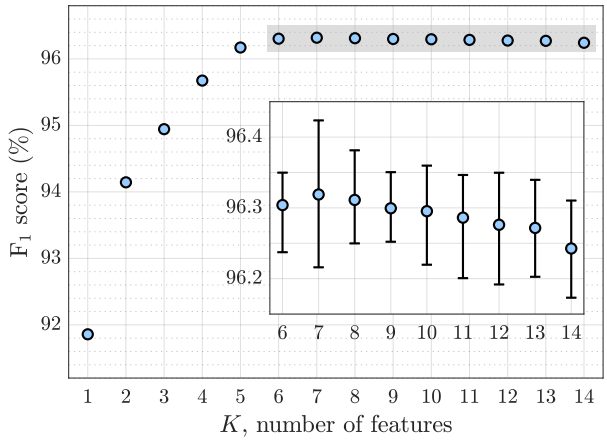


Figure 5.3. Patient-averaged F_1 scores of the ventilation detection solution for classifier models restricted to the most important K features. Scores for $K \geq 6$ are detailed in the inner axes; whiskers indicate inter-partition interdecile range.

based solutions [107, 108] (see Table 1). Only for 8 patients did the solution score an unreliable $F_1 < 75\%$; these were mainly associated with low ventilation amplitudes $< 0.2\,\Omega$, possibly because insufflated volumes were low. Figure 5.4 shows a more detailed distribution of the performance metrics by patient.

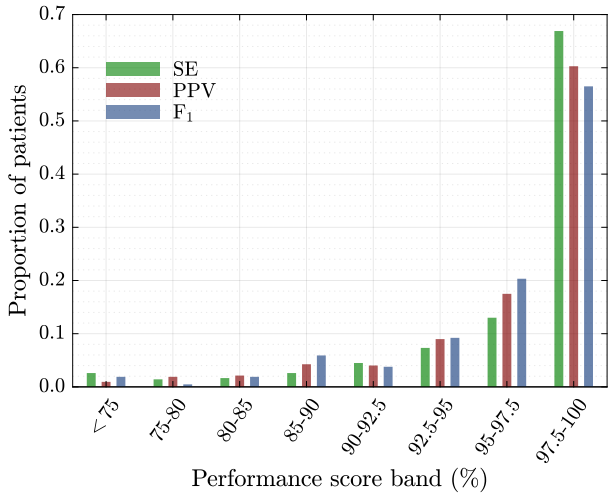


Figure 5.4. Distribution of patients for different performance metric score bands.

Finally, the performance of the ventilation detection solution in providing VR estimates was analyzed. The VR was computed as

$$\mathrm{VR} = 60 \,/\, \mathrm{med}\{\Delta t_{v,i}\} \quad (\mathrm{min}^{-1}), \tag{14}$$

with $\Delta t_{v,i}$ the time differences between contiguous ventilations, measured at the onsets of expiration. VR measures were acquired every 15 s over a centered one-minute window. The median ground truth VR was of 6.0 (4.5 – 8.0) min^{-1}. The global 90% levels of agreement (LoA) between the estimated and ground truth VRs were of (-0.82, 1.40) min^{-1}. A more detailed analysis of the errors in VR estimation is shown in Figure 5.5. Errors were larger for high ($> 10\,\mathrm{min}^{-1}$) and low ($< 6\,\mathrm{min}^{-1}$) rates, probably underrepresented in the study data.

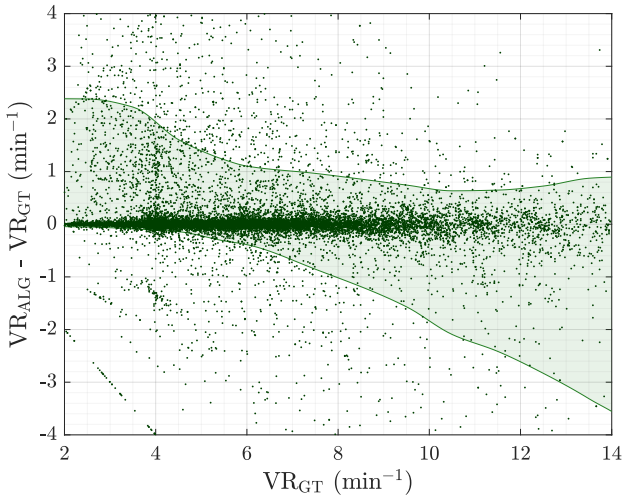


Figure 5.5. Bland-Altman plot, comparing estimated ($\mathrm{VR_{ALG}}$) and ground truth ($\mathrm{VR_{GT}}$) ventilation rates. Moving average 90% levels of agreement are depicted in light green.

5.1.2 $\mathrm{J1_2}$: Impedance-based ventilation detection and signal quality control during out-of-hospital cardiopulmonary resuscitation

The initial study dataset comprised the defibrillator files and clinical data from OHCA episodes in the PART database corresponding to the Dallas ROC site (Texas, USA) and acquired with a Philips MRx

monitor-defibrillator. Episodes including TI, CD (thus also force and acceleration) and capnogram recordings were considered. In a first step, episodes intervals ≥ 70 s with concurrence of all signals were selected. Pauses in chest compressions > 20 s and unusually heavy TI artifacts were excluded. Note that the Q-CPR pad used to acquire CD is not compatible with mechanical compressors, so manual CPR is warranted. Ground truth ventilation were automatically annotated in the capnogram [107] and then manually reviewed. The time-delay of the capnogram — mean (SD) of 3.5 (0.3) s — was also corrected, with a default delay of 3.5 s [106] applied when visual alignment was not possible. Uninterpretable capnogram sections were excluded from the study. Finally, the selected episode intervals were subdivided into non-overlapping one-minute segments, which constituted the basic analysis unit of the ventilation detection solution. A signal padding of 5 s was included to enable the characterization of partially contained ventilations. The final dataset comprised 2,551 one-minute segments from 367 different OHCA episodes, median (IQR) of 6 (3 – 10) minutes per episode. A total of 20,724 ventilations were annotated, median (IQR) of 8 (6 – 11) ventilations per minute and 45 (23 – 78) ventilations per episode. The 97.1% of ventilations were concurrent with chest compressions. Figure 5.6 shows an example of the signals involved in the development and functioning of the solution.

The ventilation detection solution was designed to analyze one-minute TI segments. The three-stage architecture described in Figure 5.1 was adopted; individual stages were implemented as follows:

- *Signal preprocessing:* The TI, acceleration and force signals were first resampled to $f_s = 50$ Hz and band-pass filtered in the 0.06 – 5 Hz range. Then, an adaptive filter was applied to suppress compression artifacts in the TI. Finally, the TI was smoothed using a finite impulse response low-pass filter with 1 Hz cutoff frequency.

  The suppression of compression artifacts considered a multichannel artifact model of the form

$$\hat{s}_{cc}(n) = \sum_{k=-M}^{M} a_k(n) s_a(n-k) + b_k(n) s_f(n-k), \qquad (15)$$

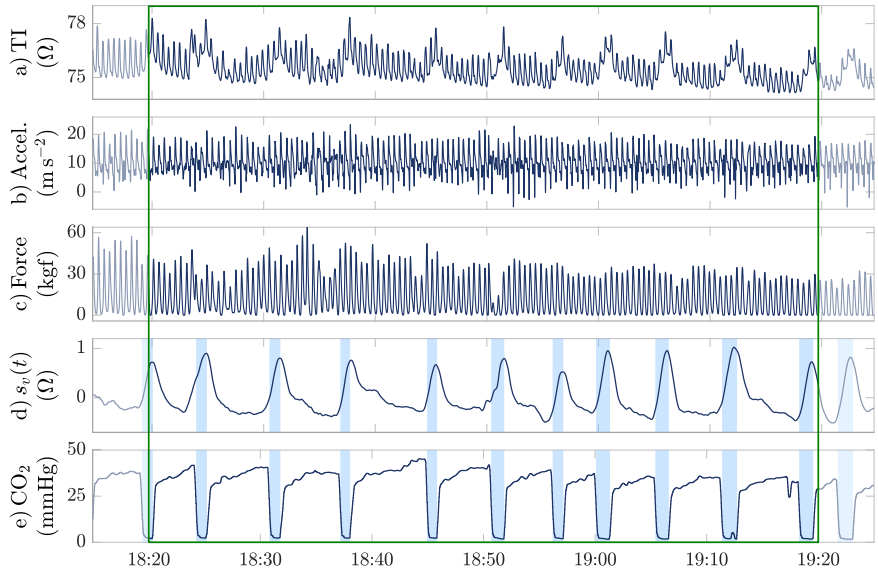


Figure 5.6. Example of a one-minute analysis segments (plus padding; faded, outside green box), including all the signals involved in the ventilation detection solution: (a) raw thoracic impedance (TI); (b) acceleration and (c) force, used as references for the suppression of chest compression artifacts; (d) ventilation component $s_v(t)$ of the TI, obtained after the preprocessing stage; and (e) time-aligned capnogram (CO_2), used to annotate ground truth ventilations (shaded in blue).

with $M = 10$, and s_a and s_f the processed acceleration and force signals, respectively. The time-varying coefficients a_k and b_k were assumed to follow Ornstein-Uhlenbeck processes and were estimated through Kalman smoothing [124, 209].

- *Fluctuation detection:* Potential ventilations were identified in the filtered TI as the largest fluctuations with peaks $t_{p,i}$ separated by a minimum of 1.5 s, and inspiration/expiration durations within 0.45 – 5.5 s. The updated heuristic to select the start and end points of fluctuations is fully described in the supplementary materials of the original article [198], and also included in Appendix A.

  The selected fluctuations were then characterized by a vector $\boldsymbol{x}_i$ of 14 features, including the amplitudes (Z_u, Z_d) and durations (T_u, T_d) of the inspiration and expiration phases (see Section 2.4.2), and the curve fit coefficients ($c_{u,m}$, $c_{d,m}$) of each phase in terms of order $m = \{0, ..., 4\}$ Legendre polynomials. Let $\boldsymbol{p}_u$ be a vector of length L_u containing the signal samples of the inspiration phase, and let

$\boldsymbol{g}_m$ be another vector resulting from the evaluation of the order m Legendre polynomial in a $[-1, 1]$ range of L_u equispaced samples, the coefficients $c_{u,m}$ were computed as

$$c_{u,m} = \frac{\boldsymbol{p}_u \cdot \boldsymbol{g}_m}{||\boldsymbol{g}_m||^2} \quad \text{for} \quad m = 0, ..., 4 \tag{16}$$

An analogous procedure was used to compute the coefficients $c_{d,m}$.

- *Classification:* The combined fluctuation information of the one-minute segment was used to compose a time series $\mathbf{X}'$ of 60 time-steps, each time-step representing one second. As shown in Figure 5.7, feature vectors $\boldsymbol{x}_i$ were assigned to time-steps based on their peak position $t_{p,i}$ within the segment. Similarly, a label time series $\mathbf{Y}'$ was composed for training and evaluation. The allocated time-steps took value $y_i = 1$ when the TI fluctuation corresponded to a ground truth ventilation (with a forward margin of 1 s), and $y_i = 0$ otherwise. Unallocated time-steps were assigned a $\mathbf{X}'(n) = \overline{\mathbf{0}}$ null feature vector and a $\mathrm{Y}'(n) = 0$ not-a-ventilation label.

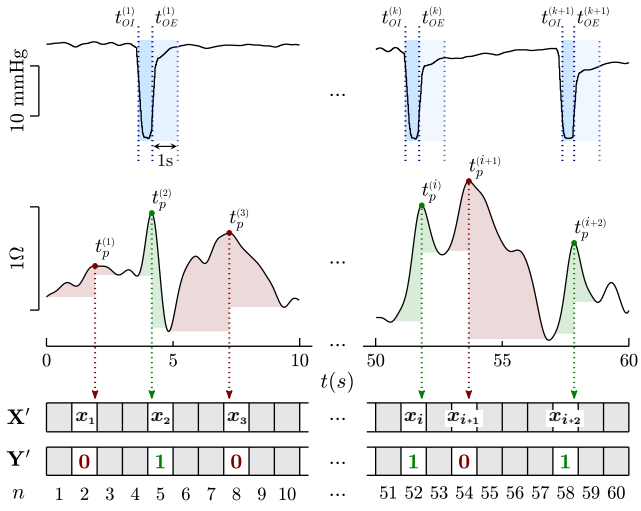


Figure 5.7. Labeling and time series composition of fluctuation data. Features $\boldsymbol{x}_i$ were assigned to time-steps of $\mathbf{X}'$ base on the fluctuation peak position $t_{p,i}$. The corresponding time-steps of $\mathbf{Y}'$ were labeled as $y_i = 1$ if the fluctuation matched a ground truth ventilation in the capnogram, and as $y_i = 0$ otherwise.

An RNN was used for classification, which could benefit from the enhanced context of the time-series representation. The network was composed of a single bidirectional recurrent layer of 20 *gated recurrent units* (GRU) [210] and a sigmoid output activation layer. A sequence-to-sequence configuration was used to produce a probability output p for each time-step. Outputs were recovered from the assigned time-steps and fluctuations classified as ventilations if $p_i \geq 0.5$. Given the predominance of the negative class (mostly due to unallocated time-steps), the network was optimized using the Dice loss function [211] (closely related to the F_1 metric).

In addition to the three fundamental stages of the solution, a fourth *signal quality control* stage was also designed in order to anticipate one-minute segments in which ventilation detection could be unreliable. This could be used to reduce data screening in retrospective studies, or to prevent erroneous feedback in real-time scenarios. The stage was designed as a linear regression model, in which the F_1 scores of the ventilation detection solution were the target variable. A logit link function was applied to ensure regression outputs between 0 and 1. Three amplitude-independent input features were considered, which were computed from the filtered TI resampled to 5 Hz: (1) the skewness of the sample distribution, (2) the amplitude of the first peak in the normalized autocorrelation [212], and (3) the SD12 variability score [213] for the first signal differences. Once trained, the model could be applied to new input segments to provide a quality score (QS), an estimation of the performance of the solution in the given segment.

The ventilation detection solution was evaluated using a patient-wise 5-fold CV strategy. Fold assignment was quasi-stratified, so each fold contained a similar number of patients, segments and ground truth ventilations. Performance was assessed for both segments and patients in terms of SE, PPV and F_1, and compared to that of previous solutions in the literature. A recursive feature elimination procedure based on the permuted feature importance (see Sections 2.4.5 and 5.1.1) was also carried out to reduce model complexity; in absence of OOB data, a nested 4-fold CV schema was used to measure the performance loss of the permuted features. The same nested CV was used to obtain the F_1 references required to train the quality control

model. The model was then applied to test data, and the performance of the ventilation detection solution reassessed for different test data inclusion rates (i.e., only the specified fraction of test segments with highest QS were considered for evaluation). The evaluation process was repeated and averaged over 20 different random CV partitions to minimize data partition bias.

Table 5 shows the median (IQR) performance metrics of the ventilation detection solution, evaluated on both one-minute segments and patients (aggregating all ventilation/segments from the same episodes). The proposed solution outperformed in about 15 points previous solutions in the literature [99, 105, 106], which were implemented and tested in the study dataset. The performances of these solutions were also much lower than originally reported, as shown in Table 2. This was probably due to lower ventilation amplitudes and higher levels of noise than in the original studies, resulting in a significant increase of false negatives and positives, respectively (note that the solutions are mainly based on amplitude thresholding). The addition of bidirectional filtering and context-aware classification also helped outperform a preliminary solution [204], presented in an international conference and which directly adapted the solution for mechanical CPR [197] (RF classifier, LMS filter, and parameterized model of compression artifacts [161]). Still, this preliminary solution also outperformed previous solutions in about 10 points, highlighting

Table 5. Median (IQR) performance metrics of TI-based ventilation detection solutions

Solution	F_1 (%)	SE (%)	PPV (%)
Per-segment evaluation			
Proposed solution [198]	89.1 (70.8 – 99.6)	93.3 (75.0 – 100.0)	90.0 (68.5 – 100.0)
LMS/RF solution [204]	85.2 (66.4 – 96.4)	87.5 (64.0 – 100.0)	87.9 (68.8 – 100.0)
Risdal et al. [105]	75.0 (53.3 – 90.0)	94.1 (75.0 – 100.0)	71.4 (46.7 – 91.0)
Edelson et al. [106]	66.7 (40.0 – 85.7)	83.5 (50.0 – 100.0)	71.4 (43.3 – 100.0)
Alonso et al. [99]	75.0 (55.2 – 92.3)	100.0 (85.7 – 100.0)	66.7 (42.9 – 90.0)
Per-patient evaluation			
Proposed solution [198]	84.1 (69.0 – 93.9)	86.5 (71.6 – 95.1)	85.4 (68.3 – 94.7)
LMS/RF solution [204]	80.3 (65.2 – 90.7)	80.0 (59.6 – 91.5)	83.9 (69.0 – 93.4)
Risdal et al. [105]	70.0 (56.3 – 82.8)	87.3 (71.7 – 95.8)	65.2 (48.1 – 82.1)
Edelson et al. [106]	62.1 (46.5 – 70.5)	77.3 (52.2 – 91.2)	64.3 (45.8 – 83.6)
Alonso et al. [99]	68.6 (56.1 – 85.5)	92.5 (84.6 – 97.1)	60.2 (43.4 – 79.5)

the importance of a more detailed characterization of fluctuations and the potential of machine learning classification.

The recursive feature elimination analysis showed that most features positively contributed to classification. Still, the RNN proved robust to less detailed information. A model including the four most important features (in descending order: Z_u, Z_d, $c_{u,0}$ and $c_{u,1}$) scored a median per-segment F_1 of 88.6%, only half point behind the fully featured solution. Interestingly, the median F_1 using only Z_u was of 87.1%. This was possible due to the context information managed by the RNN. By contrast, the application of a static amplitude threshold (found optimal at 0.25 Ω) resulted in a much lower F_1 of 78.3%.

Despite the comparatively good performance of the solution, there were still many segments and patients for which ventilation detection was inaccurate. Figure 5.8 shows the distribution of segments and patients for different performance bands. Per-patient metrics were less likely to show extreme values, as aggregated patient data may include both low- and high-performance segments. Low-performance segments were mostly characterized by low ventilation amplitudes and/or high levels of noise. Figure 5.9 shows one such example.

The quality control stage was designed to anticipate such segments, preventing erroneous feedback or reports. The proposed regression model proved effective at sorting segments by performance, with a

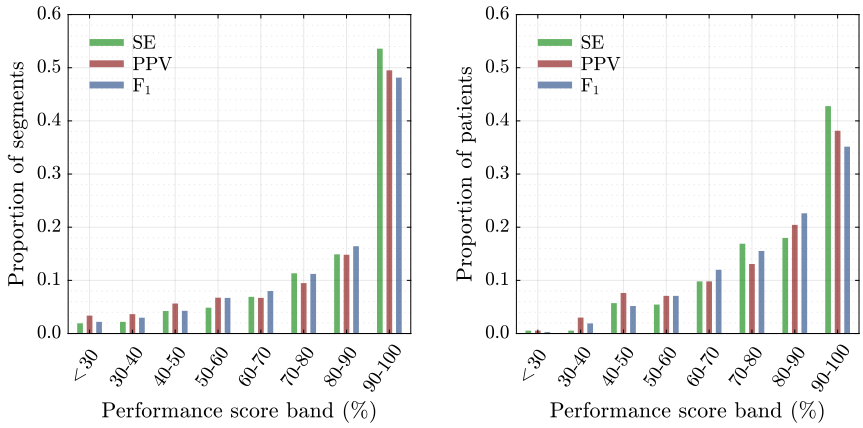


Figure 5.8. Distributions of one-minute segments (left) and patients (right) for different SE, PPV and F_1 score bands.

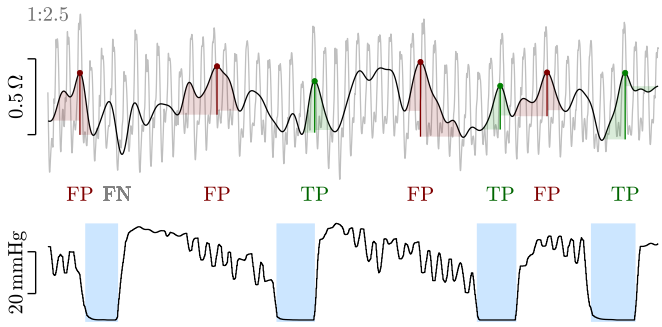


Figure 5.9. Example of an error-prone ventilation waveform. Some ventilations were correctly identified (true positive, TP), but low ventilation amplitudes and ventilation-level noise resulted in many false positives (FP). The fluctuation associated to the first ground truth ventilation was also missed (false negative, FN)

Spearman correlation between estimated QS and actual F_1 of $\rho = 0.7$. Figure 5.10 shows the per-segment and per-patient performances for different QS-based segment inclusion rates. The median F_1 grew monotonically as less segments were considered, and was close to that obtained with an ideal QS (the actual F_1).

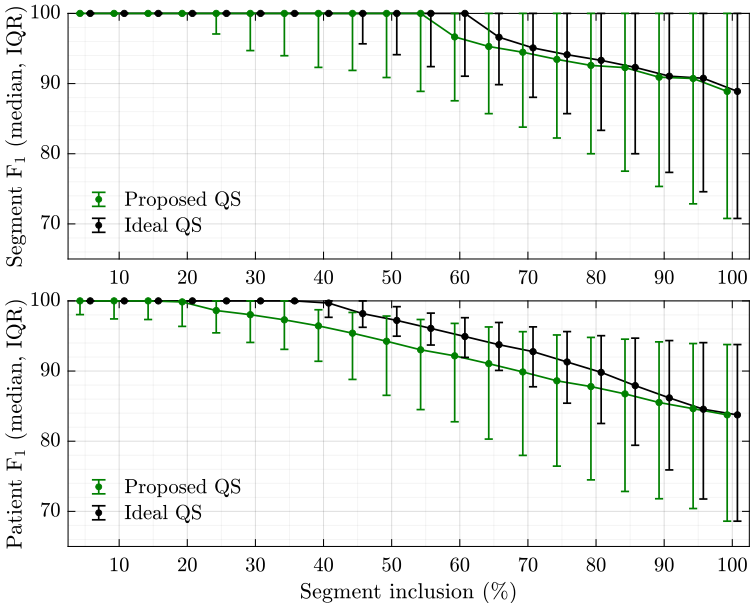


Figure 5.10. Median (IQR) per-segment (top) and per-patient (bottom) F_1 scores, with the evaluation restricted to different percentages of test segments with the highest estimated quality scores (QS). The ideal QS corresponds to the F_1 score itself.

Finally, the ventilation detection solution jointly with the quality control were evaluated in estimating VR. For each segment, the VR was computed as

$$\mathrm{VR} = \overline{60 \,/\, \Delta t_{v,i}} \quad (\mathrm{min}^{-1}), \tag{17}$$

with $\Delta t_{v,i}$ the time difference between contiguous ventilations, measured at expiration onsets. The median (IQR) ground truth VR was 8.4 (6.1 – 11.0) min^{-1}. Errors were large without quality control, with global 90% LoA between estimated and ground truth VRs of (−3.8, 6.8) min^{-1}. In contrast, the 90% LoA were (−3.4, 3.9) and (−2.0, 2.0), respectively, for 70% and 35% QS-based segment inclusion rates. Figure 5.11 shows a Bland-Altman plot with local LoA computed for different VR ranges. Overestimation was frequent for low to moderate rates, whereas underestimation was prevalent for high rates ($> 12\,\mathrm{min}^{-1}$). This was probably in part due to data imbalance, and could be alleviated with data augmentation techniques.

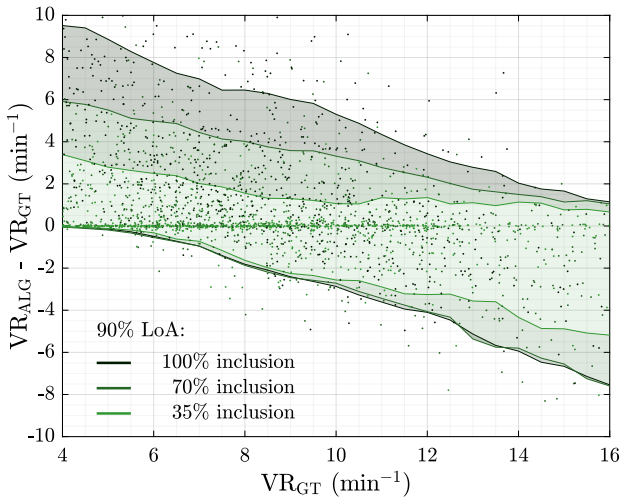


Figure 5.11. Bland-Altman plot, comparing the estimated ($\mathrm{VR}_{\mathrm{ALG}}$) and ground truth ($\mathrm{VR}_{\mathrm{GT}}$) ventilation rates. Local 90% levels of agreement (LoA) were calculated every 0.5 min^{-1} over a centered 3 min^{-1} window. Segment inclusion was based on the proposed quality control model.

5.1.3 Deep learning for impedance-based ventilation detection during continuous manual chest compressions

One of the major limitations of the proposed solution for manual CC-CPR [198] is the reliance on force and acceleration data to filter out chest compression artifacts. The removal of this artifact is critical for the solution, as it enables the detection and characterization of fluctuations. However, force and acceleration recordings require the use of a Q-CPR pad, which is only compatible with Philips devices (ZOLL devices acquire only acceleration, and Stryker does not have currently a technology for CD measurement) and may not always be available. Different adaptive filters based solely on TI and/or ECG (also acquired through defibrillation pads) data could enable a more universal solution, but could also affect performance. Another option would be the use of deep learning models. Since such models do not require handcrafted features, no preliminary fluctuation detection would be needed, and solutions could be fed directly with minimally filtered TI data. This section describes initial experiments and results on the application of deep learning segmentation architectures to the detection of ventilations in TI during manual CC-CPR [205].

The study was conducted using the same final dataset as in $J1_2$ [198] (see Section 5.1.2), comprising 2,551 one-minute CC-CPR segments from 367 different patients. The segments included TI, CD (thus force and acceleration) and capnogram signal recordings. The ventilations annotated in the capnogram were used as ground truth for training and evaluation. CD, force and acceleration data were not used by the deep learning solution, but enabled direct performance comparisons with other solutions.

A U-Net [188] architecture was implemented (see Figure 5.11). In summary, a U-Net is a particular type of *convolutional neural network* [2]

2 *Convolutional neural networks* are typically hierarchical deep learning architectures. They are fundamentally built on *convolutional layers*, which produce new representations of input data by applying different filters or *kernels*. Hierarchical levels are separated by *strided convolutions* or *pooling layers*, which downsample the data while retaining critical information. Thus, the first layers of the network learn to represent low-level features of the input (e.g, sharp edges), which are then progressively used by deeper layers to learn higher-level features of interest (e.g, QRS complexes, or even organized heart rhythms).

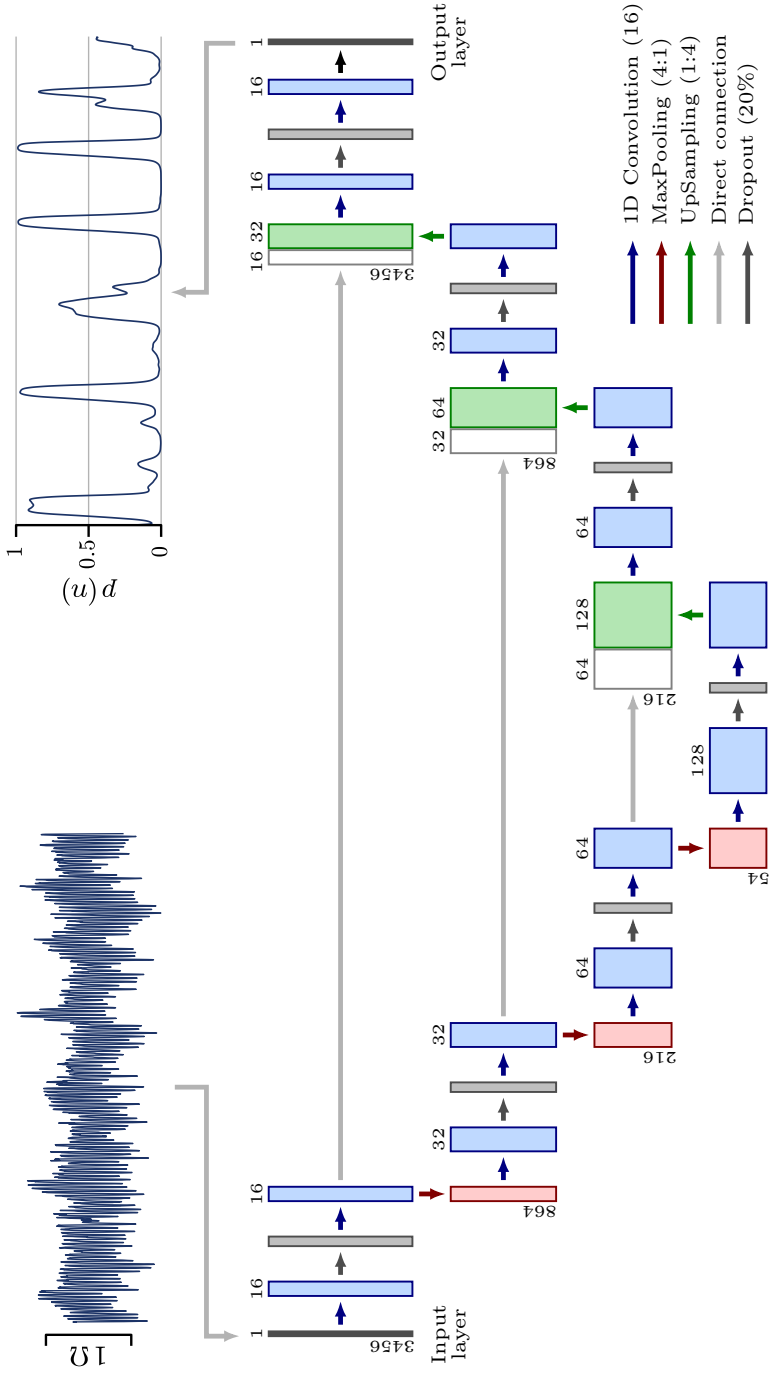


Figure 5.12. Proposed U-Net architecture for the ventilation detection solution. The number of channels is indicated at the top of the layers. Channel length (in samples) is indicated at the left side of the layers. The network was fed with one-minute thoracic impedance (TI) signal segments (plus ~4.5 s of signal padding), and returned a signal $p(n)$ with the predicted probability of each TI sample corresponding to a ventilation.

used in segmentation problems. It consists of a encoding path, which gradually downsamples the input signal while capturing high-level and context-related features, and a decoding path, which restores the original dimensions. Direct connections between encoding-decoding level pairs allow decoding layers to combine contextual information coming from deeper levels with lower-level or finer-detail features that were lost in the encoding process. The implemented architecture consisted of four hierarchical levels, with 4:1 *max-pooling* operations (downsampling operations which select the maximum value among the samples involved) between levels. So, for a TI sampling frequency of $f_s = 50\,\text{Hz}$ and a uniform convolution filter size of 16, the *receptive field* of the network went from 0.32 s at input level to 20.48 s at the fourth/deepest level. Given the size of the network and the limited training data, *separable convolutions*[3] were employed to reduce model complexity. *Dropout*[4] layers were also included to prevent overfitting.

The network was fed with minimally processed — resampling to $f_s = 50\,\text{Hz}$ and high-pass filtering at 0.06 Hz to remove the baseline component — one-minute TI segments. As in J1_2, a signal padding of 228 samples ($\sim 4.5\,\text{s}$) was included at both ends of the segment. To train the network, a target/label signal $y(n)$ of the same size as the input TI segments was composed; samples took value $y(n) = 1$ if corresponding to an inspiration phase in the time-aligned capnogram, and $y(n) = 0$ otherwise. Once trained, the network returned a signal $p(n)$ with the predicted probability of each input TI sample being part of an inspiration. Figure 5.2 shows an example of all the signals involved in the training and prediction processes of the network.

To enable comparisons between solutions, the network was evaluated in terms of the typical performance metrics used in discrete detection problems: SE, PPV and F_1. Predicted ventilation instants

3 A typical one-dimensional convolutional layer applies multichannel filters, which for N_i input and N_o output channels, and a filter size of L, implies $N_i \times N_o \times L + N_o$ trainable weights. *Separable convolutions* split the process into a *depthwise* convolution (single-channel filtering of input channels) and a *pointwise* convolution (combination of intermediate channels) for a generally much lower number of trainable weights.

4 During the training process, *dropout layers* randomly deactivate a given proportion of neurons, preventing the network from over-fixating on specific details and forcing it to generalize better.

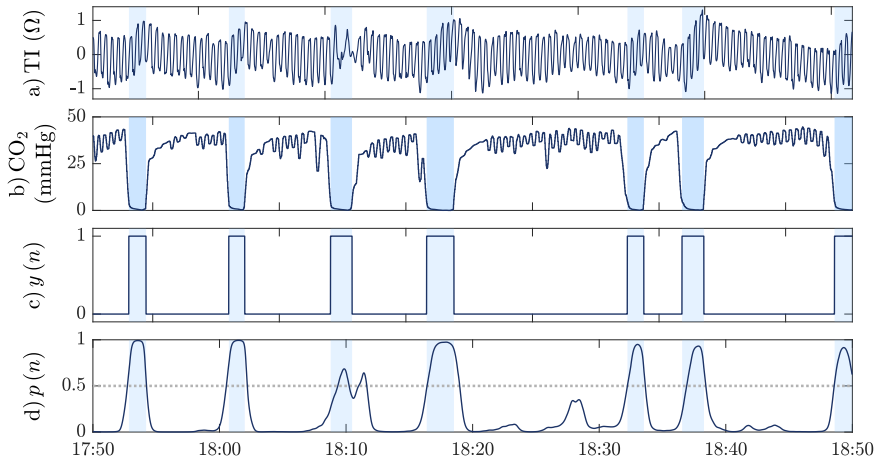


Figure 5.13. Example of the signals involved in the U-Net ventilation detection solution: (a) thoracic impedance (TI), high-pass filtered to remove baseline components; (b) time-aligned capnogram, used as ground truth to train and evaluate the solution; (c) target signal $y(n)$ to train the network, taking values $y = 1$ for samples within ground truth ventilations and $y = 0$ otherwise; and (d) output $p(n)$ of the trained network, trying to predict $y(n)$ from input TI samples.

were considered at the largest peaks of $p(n) \geq 0.5$, and were deemed correct if within an inspiration phase annotated in the capnogram (with a 0.5 s safety margin). A patient-wise 5-fold CV strategy was followed, with fold assignment quasi-stratified so each fold included a similar number of patients, segments and ventilations.

Table 6 shows the median (IQR) performance metrics of the U-Net solution. Despite the lack of reference signals, the U-Net was close in performance to the Kalman/RNN solution in $J1_2$ [198], and even outperformed a preliminary solution [204] (based on LMS filtering

Table 6. Median (IQR) performance metrics of proposed ventilation detection solutions

Solution	F_1 (%)	SE (%)	PPV (%)
Per-segment evaluation			
U-Net [205]	88.5 (66.7 – 100.0)	90.0 (69.6 – 100.0)	88.9 (71.4 – 100.0)
Kalman/RNN [198]	89.1 (70.8 – 99.6)	93.3 (75.0 – 100.0)	90.0 (68.5 – 100.0)
LMS/RF [204]	85.2 (66.4 – 96.4)	87.5 (64.0 – 100.0)	87.9 (68.8 – 100.0)
Per-patient evaluation			
U-Net [198]	82.2 (66.7 – 93.3)	83.1 (65.3 – 94.2)	84.9 (71.4 – 95.0)
Kalman/RNN [198]	84.1 (69.0 – 93.9)	86.5 (71.6 – 95.1)	85.4 (68.3 – 94.7)
LMS/RF [204]	80.3 (65.2 – 90.7)	80.0 (59.6 – 91.5)	83.9 (69.0 – 93.4)

and RF classification) which also relied on CD recordings to derive compression instants. Note that both these solutions already outperformed in about 15 and 10 points, respectively, previous solutions in the literature (see Table 5). Also, the small loss in performance of the U-Net could be easily outweighed by its potential applicability to a broader range of scenarios — further research is needed to assess its generalizability to different defibrillator devices. On the negative side, the solution does not explicitly characterize ventilations; additional methods would thus be needed to measure morphological features such as amplitudes.

5.2 Results related to objective 2

Multi-center studies and clinical trials are of great interest to resuscitation science, as they allow for larger study cohorts and help minimize EMS and population biases. However, different EMS agencies may use defibrillator devices from different vendors, each with its particular signal acquisition and CPR analysis tools, making it difficult to harmonize CPR quality analyses. Proprietary analysis software may also require considerable human input [147], potentially unfeasible for large study datasets. The many solutions in the literature to characterize chest compressions [99, 100, 101, 126, 102] offer an opportunity for ad-hoc automatic analyses. However, most of these solutions have been validated only in homogeneous datasets, and usually under controlled conditions (not considering signal loss, ROSC events, and other possible but frequent complications). As part of this thesis, a unified methodology for automatic CPR quality analysis in large multi-device datasets was developed and evaluated. This study resulted in a JCR publication ($J2_1$) [199].

5.2.1 $J2_1$: Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets

For this study, all defibrillator files available in the PART database were initially considered. Files from three major defibrillator commercial brands were included: Philips, Stryker/Physio-Control and ZOLL. As a first step of the proposed framework, defibrillator files

were converted to a common MATLAB (The MathWorks Inc., Natick, MA, USA) format to enable uniform processing (see Figure 5.14). This new format included general file information (original file name, device model, date and time, ...), all available biomedical signal waveforms, and timestamps for the most relevant device events (such as defibrillation shocks or rhythm analyses). The chest compression instants identified by the defibrillator device or associated software were also included. In addition, relevant clinical data (ROSC and airway insertion annotations, for this study in particular), available in spreadsheet format, were automatically linked and incorporated.

Episodes not suitable for the analysis of chest compression (less than 1 min of interpretable TI or CD recordings) were discarded in a preliminary data screening. Episodes involving several defibrillator files were also discarded or restricted to the most representative file when synchronization was unreliable. The final dataset included data from 2232 OHCA cases (2356 defibrillator files): 925 (926) Philips, 389 (389) Stryker, and 918 (1041) ZOLL.

All defibrillator files were subjected to a fully automatic annotation process, including the annotation of: (a) chest compression instants, (b) signal unavailability periods, (c) ROSC intervals, and (d) start and termination of resuscitation efforts. Specific details were as follows:

- *Chest compressions:* State-of-the-art algorithms were applied and adapted for the identification of chest compressions in the different signals available. Compression instants in the TI (Philips, Stryker) were annotated using the solution by Alonso et al. [99] (see Section 2.1.2). For Philips cases, compressions in the CD were identified as negative peaks < 1.5 cm with a minimum separation of 0.35 s, as proposed by Ayala et al [100]. For ZOLL cases, where the CD signal is zero-centered (see Figure 5.15), the solution by Alonso et al. was adapted; a $1-5$ Hz band-pass filter and a $1-2$ cm amplitude adaptive threshold were considered. When both TI and CD recordings were simultaneously available, compressions detected in the CD were prioritized. Compression series were defined as groups of 5 or more consecutive compressions without interruption (procedurally defined as > 3 s [127]); isolated compressions not part of of compression series were discarded.

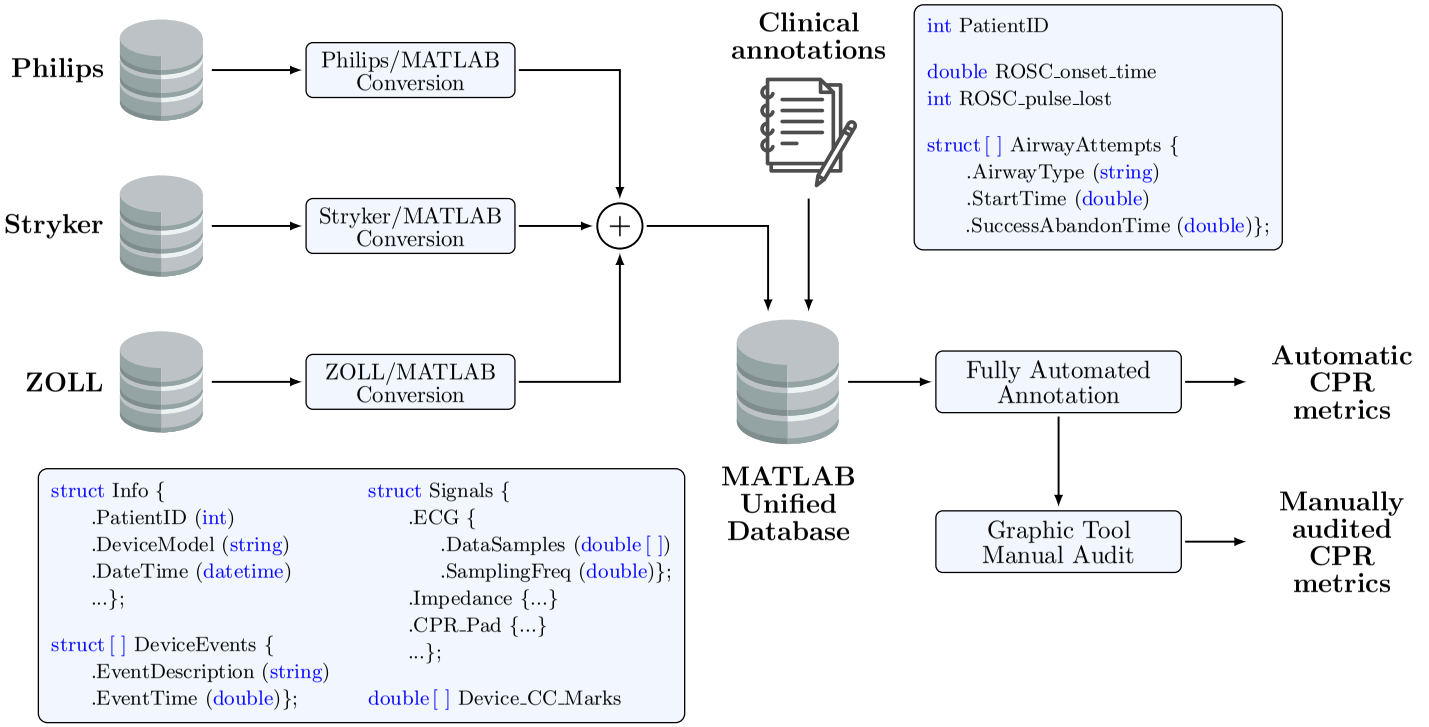


Figure 5.14. General workflow for CPR quality analysis in large multi-device datasets. First, all defibrillator files are converted to a common MATLAB format. Then, external clinical data is linked using some common variable (such as defibrillator file name or patient ID), and the information of interest is imported. Finally, the files are subjected to a series of annotation procedures to enable the computation of quality metrics. A dedicated graphical tool enables manual data screening if required.

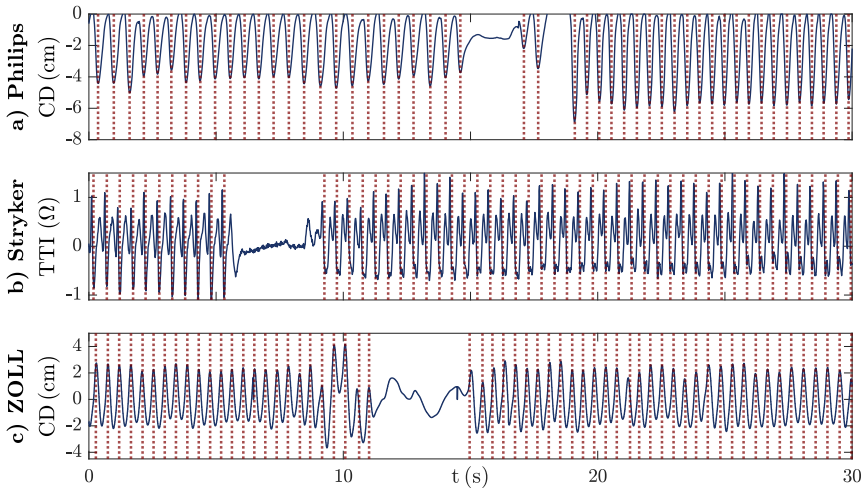


Figure 5.15. Example of signals from different monitor-defibrillators suitable for chest compression analysis: (a) compression depth (CD) from a Philips MRx, (b) thoracic impedance (TI) from a Stryker/Physio-Control LP12, and (c) CD from a ZOLL X-Series. Device chest compressions are depicted as red dashed lines.

- *Signal availability:* Episode intervals with unavailable signal sources for compression analysis were annotated for exclusion in the calculation of CPR quality metrics. For Stryker and ZOLL records, unavailability was derived directly from the device event log. For Philips records, TI and CD were deemed unavailable for TI values $< 30\,\Omega$ or $> 200\,\Omega$, and acceleration values $< 5\,\mathrm{m\,s^{-2}}$, respectively. For episodes involving multiple defibrillator files, blind intervals between different files were also annotated as unavailable.
- *ROSC:* Intervals with spontaneous circulation were identified using EMS ROSC annotations, which were automatically extracted from linked clinical data spreadsheets. Transient ROSC intervals were annotated covering the longest compression pause in the vicinity of an EMS ROSC timestamp (considering also the interval from last compression to end of file). For sustained ROSC events, the intervals were extended to the end of the file; any chest compression within a ROSC interval was discarded.
- *Resuscitation interval:* The span of resuscitation efforts was defined according to Kramer-Johansen et al. [58], i.e., from first compression, defibrillation shock or rhythm analysis, to last compression or onset of sustained ROSC. Sustained ROSC timestamps were de-

rived from EMS clinical annotations, whereas shocks and rhythm analyses were identified from the device event log. In the case of compressions, a simple heuristic was applied to avoid compression-like artifacts: given a compression series of length $L < 30\,\text{s}$ at the start or end of the episode, the series was considered an artifact and discarded if followed or preceded, respectively, by a pause $> L$ not corresponding to a shock, rhythm analysis or ROSC. A similar procedure was applied at ROSC boundaries.

The automatic annotations were manually audited using a purpose-built graphical interface. In addition, intervals with uninterpretable signals were annotated as noise to be excluded from quality metric calculations. Annotation of ROSC was especially relevant, as clinical data included a single timestamp per episode, corresponding to the first ROSC event. Secondary ROSC events were annotated for 132 cases, associated with pauses in chest compressions $> 1\,\text{min}$ during organized heart rhythms. ROSC intervals were also annotated in 87 cases with no clinically documented ROSC events. After data review, compression quality metrics were computed using both the automatic and the manually audited annotations. CCF, CCR and compression interruptions were considered as quality metrics (results related to interruptions are omitted in the following due to high redundancy with CCF ones). Two different analysis intervals were considered: the entire resuscitation period, and the airway insertion period (from the first advanced airway insertion attempt, to successful insertion or abandonment of insertion efforts).

Table 7 summarizes the errors of the automatic analysis procedure for the entire resuscitation episode. Median errors below 2% in CCF and $1\,\text{min}^{-1}$ in CCR were measured for all devices. The proportion of episodes with important errors ($> 10\%$ in CCF and $> 10\,\text{min}^{-1}$ in CCR) were below 10% in all cases. Median (IQR) errors in CCR and CCF, as well as important errors in CCR, were larger for Stryker devices, which only include TI recordings (less reliable, in general, than CD). Important errors in CCF were more similar between devices, and were related primarily to ROSC events. The 37.4% of episodes with secondary or undocumented ROSC events presented important errors in CCF. By contrast, only the 2.5% of episodes without ROSC events showed important errors.

Table 7. Median (IQR) analysis durations and unsigned errors in automatic quality metric calculation for the entire resuscitation period

		CCF		CCR	
Device (*N*)	duration (min)	error (%)	$>10\%$	error (min^{-1})	$>10\,\text{min}^{-1}$
Philips (925)	28.7 (19.3 – 34.9)	0.5 (0.1 – 1.9)	92 (9.9%)	0.3 (0.1 – 0.7)	1 (0.1%)
Stryker (389)	19.0 (12.5 – 27.1)	1.6 (0.6 – 3.9)	36 (9.3%)	0.8 (0.3 – 2.5)	12 (3.1%)
ZOLL (918)	21.8 (14.8 – 29.2)	0.7 (0.3 – 2.0)	68 (7.4%)	0.2 (0.1 – 0.6)	6 (0.7%)
TOTAL (2232)	23.6 (15.6 – 31.3)	1.2 (0.8 – 3.0)	196 (8.8%)	0.3 (0.1 – 0.9)	19 (0.9%)

Table 8 shows the errors of the automatic procedure in the analysis of the airway insertion period. This period was selected as a short-duration counterpart of the entire resuscitation analysis, but also to highlight another important aspect of the proposed methodology: the ability to automatically define and execute analyses in terms of clinical or other external information. The analysis was limited to the episodes for which the relevant airway insertion timestamps were available, and for which TI or CD recordings were present in $>50\%$ of the insertion period. Median (IQR) errors were lower than for the entire resuscitation, possibly because error sources concentrate in parts of the episode which may often not coincide with the airway insertion interval. By contrast, the proportion of episodes presenting important CCR errors was larger for all devices.

Table 8. Median (IQR) analysis durations and unsigned errors in automatic quality metric calculation for the airway insertion period

		CCF		CCR	
Device (*N*)	duration (min)	error (%)	$>10\%$	error (min^{-1})	$>10\,\text{min}^{-1}$
Philips (607)	1.3 (1.0 – 3.0)	0.0 (0.0 – 0.0)	21 (3.5%)	0.0 (0.0 – 0.3)	10 (1.6%)
Stryker (114)	1.3 (0.8 – 3.5)	0.3 (0.0 – 2.5)	8 (7.0%)	0.5 (0.1 – 1.7)	5 (4.4%)
ZOLL (418)	1.0 (0.7 – 3.0)	0.2 (0.1 – 1.0)	11 (2.6%)	0.1 (0.0 – 0.4)	7 (1.7%)
TOTAL (1139)	1.2 (0.8 – 3.0)	0.0 (0.0 – 0.7)	40 (3.5%)	0.0 (0.0 – 0.5)	22 (1.9%)

Finally, the performance of the compression detection algorithms used in the study was evaluated and compared to that of defibrillator devices and their associated software. The analysis intervals were defined using audited data (including the start/end of resuscitation and both ROSC and noise intervals), so potential errors were limited to those of the detection algorithms. As show in Table 9, errors were overall low both for the study algorithms and for vendor solutions.

TI-based compression detection was, in general, more error prone than CD-based one. In every case, the algorithm used in the study was at least as good as its commercial counterpart.

Table 9. Median (IQR) unsigned errors in CPR quality metrics due to malfunction of automatic chest compression detection algorithms.

		CCF		CCR	
Device	N	error (%)	$>10\%$	error (min^{-1})	$>10\,\text{min}^{-1}$
Chest compression instants from study algorithms					
Philips	925	0.2 (0.0 – 0.6)	10 (1.1%)	0.2 (0.1 – 0.7)	1 (0.1%)
- TI	894	0.8 (0.3 – 1.9)	33 (3.7%)	0.5 (0.2 – 1.0)	2 (0.2%)
- CD	531*	0.1 (0.0 – 0.3)	2 (0.4%)	0.3 (0.1 – 0.6)	0 (0.0%)
Stryker (TI)	389	0.9 (0.4 – 2.1)	5 (1.3%)	0.7 (0.3 – 2.2)	10 (2.6%)
ZOLL (CD)	918	0.4 (0.2 – 0.7)	4 (0.4%)	0.2 (0.1 – 0.4)	1 (0.1%)
Chest compression instants from device software					
Philips (CD)	531*	0.4 (0.2 – 1.0)	3 (0.6%)	0.3 (0.1 – 0.6)	0 (0.0%)
Stryker (TI)	389	1.8 (0.7 – 4.0)	27 (6.9%)	1.4 (0.4 – 3.4)	24 (6.2%)
ZOLL (CD)	918	0.4 (0.1 – 1.1)	20 (2.2%)	0.5 (0.2 – 1.4)	14 (1.5%)

* From 886 Philips files with CD, only 531 included compression annotations.

5.3 Results related to objective 3

The PART study [92] assessed for differences in clinical outcome between ETI and LT initial advanced airway strategies, but did not consider CPR quality variables. The algorithms and methodologies developed in the thesis resulted in an extensive characterization of CPR in the PART database, enabling comparisons of potential clinical relevance between airway groups. Three major aspects of CPR were considered: chest compression quality metrics, ventilation rates, and ventilation amplitudes in the TI. The results of these studies led to three JCR publication, namely $J3_1$ [200], $J3_2$ [201] and $J3_3$ [202], which are described in the following. Preliminary and summary results were presented in different international conferences [214, 215, 216, 217, 218].

5.3.1 $J3_1$: Airway strategy and chest compression quality in the Pragmatic Airway Resuscitation Trial

The initial study dataset comprised all the electronic defibrillator files and associated clinical data available in the PART database. Data

preparation was shared with $J2_1$ [199] (see Section 5.2.1), comprising: conversion of defibrillator files to a common MATLAB format, integration of clinical data, and automatic annotation and posterior manual review of chest compression instants, start and end of CPR, noise/signal unavailability sections and ROSC intervals. The following inclusion criteria were applied after data review: at least one ETI or LT advanced airway insertion attempt, minimum duration of CPR of 3 min, and minimum availability of TI or CD signals suitable for compression analysis of 50% of the CPR duration. The final study cohort included 1996 patients, 1001 and 995 assigned, respectively, to LT and ETI initial airway strategies.

Manually audited data were used to compute CPR quality metrics, including CCF, CCR, and individual and total interruptions in chest compressions; CD was not considered due to potential biases between different devices. A procedural definition of interruption of 3 s was used. The interval of resuscitation was defined according to Kramer-Johansen et al. [58], i.e., from first chest compression, shock or rhythm analysis, to last compression or onset of sustained ROSC. CPR quality metrics were calculated for the entire resuscitation episode, and for resuscitation periods after and before airway insertion (defined as successful airway insertion or abandonment of insertion efforts, as reported by EMS personnel). Other resuscitation periods (including the span of airway insertions and different 3 min epochs) were also considered, but did not add value to the results; the results of these analyses can be found in $J3_1$ [200] as supplementary material.

Table 10 shows the comparison of CPR quality metrics between LT and ETI patient groups. Group assignment was based on *intention-to-treat*, i.e, the initial airway strategy assigned according to the trial protocol. Quality metrics were compliant with resuscitation guidelines. For the 76.4% of patients, the CCR was in the recommended range of $100-120\,\mathrm{min}^{-1}$, and only for the 6.5% of patients deviated by more than $10\,\mathrm{min}^{-1}$. The CCF was above 80% and 60% for the 84.3% and 98.9% of patients, respectively. No significant differences in CCR or CCF were observed between airway groups for any of the periods considered. The mean duration of individual interruptions was also similar (12.6 LT vs 13.0 ETI, $p = 0.78$). The number and total duration of interruptions were significantly larger for ETI, but

Table 10. Mean (SD) chest compression variables for LT and ETI airway groups

Time-period	LT	ETI	p-value*
Analysis duration (min)			
Entire resuscitation	22.6 (10.8)	25.2 (11.3)	< 0.001
Before airway insertion	7.7 (4.9)	11.0 (5.8)	< 0.001
After airway insertion	16.2 (9.8)	15.9 (9.9)	0.45
CCF (%)			
Entire resuscitation	87.9 (8.4)	87.1 (8.7)	0.05
Before airway insertion	86.9 (12.4)	87.3 (10.6)	0.49
After airway insertion	88.8 (9.1)	88.2 (9.8)	0.17
CCR (min^{-1})			
Entire resuscitation	113.7 (9.1)	114.0 (10.5)	0.59
Before airway insertion	112.6 (11.2)	113.0 (11.2)	0.45
After airway insertion	113.6 (11.0)	113.8 (10.9)	0.72

* Calculated with t-test

mostly due to a longer resuscitation time (no differences in CCF). This concerned mainly the time before airway insertion, and could have been caused by unsuccessful or longer airway insertions in the ETI group. However, in contrast to older studies [81], it did not generally result in major CPR interruptions. Whether it could have affected outcomes due to impaired ventilation is unclear.

5.3.2 $J3_2$: Airway strategy and ventilation rates in the Pragmatic Airway Resuscitation Trial

The initial study cohort comprised all cases in the PART database with capnogram recordings, including defibrillator files from three major commercial brands: Philips, Stryker and ZOLL. To ensure uniform processing, the common format files derived from $J2_1$ [199] and $J3_1$ [200] were considered. Ventilations were first automatically annotated in the capnogram [107], and then manually reviewed. Intervals with uninterpretable capnogram were also manually annotated. After data screening, the following inclusion criteria were set: successful ETI or LT airway insertion (according to clinical Epistry data), and $> 50\%$ interpretable capnogram availability (with minimum duration of 3 min) from advanced airway insertion to termination of CPR. The final study cohort comprised 1010 cases (583 Philips, 36 Stryker, and 436 ZOLL). Successful airway insertion corresponded to LT and ETI in 714 and 296 cases, respectively.

Ventilation metrics were computed in the post-airway resuscitation interval (from airway insertion to end of CPR) using the reviewed ventilations annotated in the capnogram. TI was disregarded for ventilation analysis due to its lower reliability and the high availability of capnography after advanced airway insertion; the vast majority of ZOLL files also did not include TI recordings. The VR was calculated every 10 s as the direct count of ventilations in a one-minute window. Exposure to hypoventilation and hyperventilation were defined as $VR < 6\,min^{-1}$ and $VR > 12\,min^{-1}$, respectively.

Table 11. Median (IQR) ventilation variables for LT and ETI airway managed cases

	ETI	LT	p-value*
Analysis time (min)	16.4 (11.0 – 23.3)	16.7 (10.2 – 22.7)	0.78
VR (min^{-1})	8.0 (6.5 – 9.6)	7.9 (6.5 – 9.7)	0.94
Hypoventilation			
Total duration (min)	1.8 (0.0 – 5.6)	1.7 (0.0 – 6.1)	0.94
Time fraction (%)	10.5 (0.0 – 32.1)	11.5 (0.0 – 36.6)	0.60
Hyperventilation			
Total duration (min)	0.4 (0.0 – 2.4)	0.4 (0.0 – 2.2)	0.91
Time fraction (%)	2.1 (0.0 – 15.0)	1.9 (0.0 – 13.1)	0.99

* Calculated with Mann-Whitney U test

Table 11 shows the median (IQR) ventilation metrics in post-airway resuscitation for patients managed with LT and ETI. No significant differences were observed for any ventilation metric. No differences were observed either for other ranges of hyperventilation (mild: 12 – $16\,min^{-1}$, moderate: 16 – $20\,min^{-1}$, severe: $> 20\,min^{-1}$). Compared to the guideline recommendation of $10\,min^{-1}$ [17, 36], VRs were rather low. Exposure to hypoventilation was moderate, with the 33.2% of patients affected $> 25\%$ of time. Exposure to hyperventilation was low, with only the 17.2% of patients affected $> 25\%$ of time despite the low definition threshold of $12\,min^{-1}$. Only 16 patients (1.6%) received severe hyperventilation ($VR > 20\,min^{-1}$) $> 25\%$ of time.

The impact of exposure to hypo- and hyper-ventilation in OHCA outcomes was also analyzed. The analyses were based on *generalized estimating equation* models, with the durations of hypoventilation and all hyperventilation subgroups as independent variables. The models were adjusted for typical confounding demographic variables (age, sex, witnessed arrest, bystander CPR, ...) as well as airway treatment

group. Significant associations with outcomes were observed for hypoventilation (negative) and all hyperventilation ranges (positive). However, confidence intervals were large and results were not consistent for all outcomes considered (ROSC, survival at 72 h, survival to hospital discharge, and survival with good neurological status). Further research is needed to verify the veracity of these results.

As a secondary objective of this study, and in line with the validation of compression detection algorithms in $J2_1$ [199] (see Section 5.2.1), manually audited ventilations enabled the validation of ventilation detection solutions for potential application in large multi-device datasets. Table 12 shows the performance metrics observed for the solution by Aramendi et al. [107]. The analysis was restricted to interpretable capnogram intervals. Automatic ventilations were deemed correct if the onset of expiration was within 1 s of that of an audited ventilation. Median performance metrics were above 95% for every airway device and defibrillator brand considered. Overall, the 64.2% and 80.6% of patients presented F_1 scores above 95% and 90%, respectively. Figure 5.16 shows a more detailed distribution of patients for different performance score bands. There were 23 (3.2%) LT- and 20 (6.8%) ETI-managed patients for whom SE was below 50%. This was mostly due to capnogram sections below the 5 mmHg mark, which were undetectable by the automatic solution. The higher proportion of such cases in ETI patients could have been due to a worse patient status, consistent with the findings of the PART trial,

Table 12. Median (IQR) performance metrics of the ventilation detection solution by Aramendi et al. [107]

Device	*N*	F_1 (%)	SE (%)	PPV (%)
		LT advanced airway		
Philips	354	96.6 (91.3 – 98.7)	96.8 (88.0 – 99.2)	97.5 (93.6 – 99.2)
Stryker	21	97.9 (95.9 – 98.9)	98.2 (94.4 – 100.0)	98.2 (93.8 – 100.0)
ZOLL	339	97.6 (94.5 – 99.2)	98.7 (95.4 – 100.0)	97.7 (95.0 – 99.0)
TOTAL	714	97.2 (93.2 – 99.0)	97.8 (92.8 – 100.0)	97.6 (94.3 – 99.4)
		ETI advanced airway		
Philips	184	95.0 (86.2 – 98.5)	95.9 (86.1 – 98.9)	97.3 (91.0 – 99.2)
Stryker	15	98.0 (94.3 – 99.8)	98.3 (94.0 – 100.0)	99.3 (94.8 – 100.0)
ZOLL	97	97.1 (92.1 – 99.2)	98.1 (92.6 – 99.6)	98.4 (94.1 – 91.0)
TOTAL	296	95.9 (89.4 – 98.8)	97.0 (88.7 – 99.2)	97.8 (92.3 – 100.0)

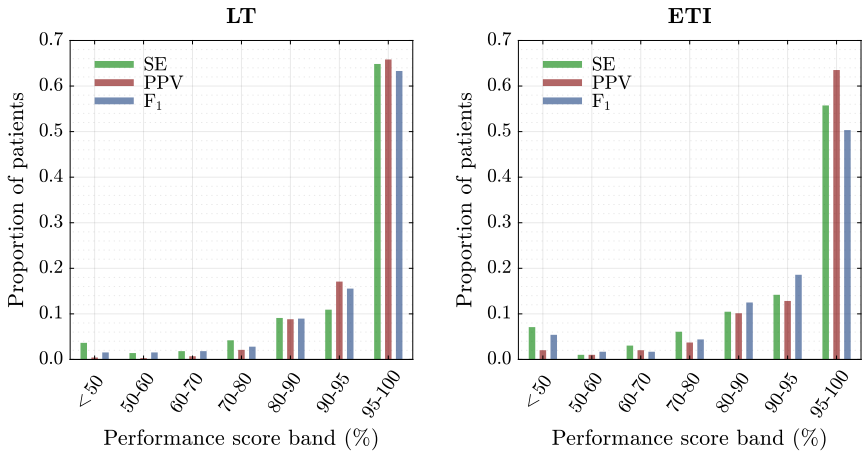


Figure 5.16. Distributions of LT- (left) and ETI-managed (right) patients for different performance metric score bands.

but also due to other possible complications, such as endobronchial intubation. Interestingly, performances were lower for the Philips MRx (the same device used in the original study); this was probably an annotation bias, as better capnogram resolution than other devices may had allowed interpretation of more complex waveforms. This validation analysis was not included in $J3_2$ [201], which had a more clinical focus, and is currently unpublished.

5.3.3 $J3_3$: Novel application of the thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the Pragmatic Airway Resuscitation Trial

For this study, Philips MRx defibrillator files in the PART database and associated to the Dallas ROC site were initially considered. The following inclusion criteria were then applied: successful LT or ETI airway insertion, at least 1 min of concurrent and good quality TI and capnogram recordings after airway insertion and during manual or mechanical CC-CPR, and a minimum of 10 ventilations within this interval. The quality of the capnogram was assessed visually. Ventilations in the capnogram were first automatically annotated [107] and then manually reviewed. The time-offset between capnogram and TI was also manually adjusted. The quality of the TI was indirectly assessed through the performance of automatic ventilation detection

solutions ($F_1 > 80\%$, evaluated against capnogram annotations): the $J1_1$ solution [197] in case of mechanical CPR (see Section 5.1.1), and a preliminary version [204] of the $J1_2$ solution in case of manual CPR. Mechanical CPR was identified by monitoring a CCR of 101.7 min^{-1}, corresponding to the LUCAS-2 device. Manual CPR intervals without concurrent CD recordings were discarded, as these were required by the ventilation detection solution. The final dataset included suitable manual CPR data from 209 cases (132 LT, 77 ETI) and mechanical CPR data from 94 cases (53 LT, 41 ETI)

The morphology of TI fluctuations due to ventilation was analyzed. The amplitudes (Z_u, Z_d) and durations (T_u, T_d) of the inspiration and expiration phases were considered (see Section 2.4.2). Median values among all available fluctuations were used to characterize each case; false positive ventilations that did not match capnogram annotations were discarded. The average VR was also calculated for each episode, following (17).

Table 13 shows a comparative between LT- and ETI-managed cases. Separate comparisons were performed for manual/mechanical CPR, as the detection solutions were based on different filters which could potentially affect the ventilation waveform. No significant differences were observed for inspiration/expiration durations. Consistent with the results in $J3_2$ (see Table 11), no differences were observed either for VR. Significant differences were observed for ventilation amplitudes,

Table 13. Median (IQR) ventilation variables for LT and ETI airway managed cases

Measure (units)	LT	ETI	p-value*
Manual CC-CPR			
Z_u (Ω)	0.46 (0.32 – 0.68)	0.71 (0.47 – 1.01)	< 0.01
Z_d (Ω)	0.45 (0.34 – 0.67)	0.70 (0.45 – 0.97)	< 0.01
T_u (s)	1.50 (1.24 – 1.70)	1.57 (1.36 – 1.86)	0.03
T_d (s)	2.32 (1.95 – 2.86)	2.33 (1.87 – 2.94)	0.78
VR (min^{-1})	8.5 (6.9 – 10.2)	8.3 (6.9 – 9.3)	0.35
Mechanical CC-CPR			
Z_u (Ω)	0.74 (0.44 – 1.39)	1.22 (0.78 – 1.74)	< 0.01
Z_d (Ω)	0.68 (0.43 – 1.27)	1.14 (0.70 – 1.53)	< 0.01
T_u (s)	1.52 (1.14 – 1.82)	1.50 (1.19 – 1.82)	0.96
T_d (s)	2.68 (2.21 – 3.37)	2.23 (1.66 – 3.29)	0.15
VR (min^{-1})	7.3 (5.3 – 8.7)	6.4 (5.2 – 8.0)	0.31

* Calculated with Mann-Whitney U test

50–70% larger in ETI-managed cases; this could be associated with an increased dead-space in LT devices, as part of the insufflated air must be used to fill the trachea and laryngopharynx. All the results were consistent between manual and mechanical CPR cases. The size of the dataset and the low number of survivors — only 9 (3.0%) patients discharged alive from hospital — were deemed insufficient for outcome analysis.

6 CONCLUSIONS

This chapter summarizes the main contributions of the thesis work. First, the most important results are highlighted, followed by a summary of all associated journal and conference contributions. Then, all research projects and funding sources that have supported the development of the thesis are noted. Finally, a brief section is dedicated to potential lines of future research in relation to the results of the thesis.

6.1 Major contributions of the thesis

The main goal of this thesis was to contribute to the automatic analysis of ventilation and CPR in large OHCA datasets, with particular focus on its application in the study of ventilation and airway management strategies. By the end of the thesis, the main contributions can be summarized as follows:

- *Automatic detection of ventilations during mechanical CC-CPR using the TI:* A solution incorporating dedicated compression frequency-based adaptive filtering and machine learning classification was developed [197]. Overall performance similar to capnogram-based solutions was achieved.
- *Automatic detection of ventilations during manual CC-CPR using the TI:* A solution incorporating bidirectional adaptive filtering and time series classification was designed [198], which outperformed in about 15 percentage points previous solutions in the literature. A signal quality control model was also designed, which effectively

anticipated the performance of the solution on new TI segments, and which could be used to minimize manual data screening or to prevent erroneous feedback in real-time applications. A second ventilation detection solution, based on deep-learning models, was also developed [205], which did not require reference signals and could potentially be applied to a broader range of scenarios.

- *Automatic analysis of CPR quality in large and heterogeneous OHCA datasets:* Prior state-of-the-art solutions and new procedures were combined, enabling the uniform and effective automatic extraction of compression quality metrics in multi-device datasets [199]. Definition of the resuscitation period, signal availability control, and ROSC annotation were all automatized. The different compression detection solutions were validated and proved overall superior to vendor-specific ones. A literature solution for capnogram-based ventilation detection was also validated in a multi-device dataset about 20-fold larger that the one in the original study.
- *Characterization of CPR in the PART database:* Novel algorithms proposed in this thesis and adaptation of old ones were combined to answer critical clinical questions in the PART study. Differences in chest compression quality metrics [200], ventilation rates [201], and ventilation amplitudes in TI [202] between LT- and ETI-managed cases were analyzed. Conclusions contributed to advance knowledge in resuscitation science.

6.2 Publications

The thesis work has materialized in several contributions to the scientific community — formally listed in Sections 6.2.1 and 6.2.2 —, including:

- Two long papers in journals indexed in the JCR Science Edition (A1 and A6), and six contributions to national and international conferences (C1, C2, C3, C6, C10 and C11), related to the first objective of the thesis.
- One long paper in a journal indexed in the JCR Science Edition (A3), related to the second objective of the thesis.

- Three long papers in journals indexed in the JCR Science Edition (A2, A4 and A5), and 5 contributions to international conferences (C4, C5, C7, C8 and C9), related to the third objective of the thesis.

6.2.1 Journals indexed in the JCR science edition

A1 *Automatic detection of ventilations during mechanical cardiopulmonary resuscitation*
Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Pamela C. Owens, Henry E. Wang, Ahamed H. Idris
IEEE Journal of Biomedical and Health Informatics 2020 (IF: 5.773, 4/30) [197]

A2 *Airway strategy and chest compression quality in the Pragmatic Airway Resuscitation Trial*
Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Jeffrey L. Jarvis, Jestin N. Carlson, Unai Irusta, Erik Alonso, Tom P. Aufderheide, Robert H. Schmicker, Matthew L. Hansen, Ryan M. Huebinger, M. Riccardo Colella, Richard Gordon, Robert Suchting, Ahamed H. Idris
Resuscitation 2021 (IF: 6.251, 3/32) [200]

A3 *Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets*
Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Erik Alonso, Tom P. Aufderheide, Robert H. Schmicker, Matthew L. Hansen, Robert Suchting, Jestin N. Carlson, Ahamed H. Idris, Henry E. Wang
Resuscitation 2021 (IF: 6.251, 3/32) [199]

A4 *Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the Pragmatic Airway Resuscitation Trial*
Michelle M.J. Nassal, Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Ashish R. Panchal, Henry E. Wang, Ahamed H. Idris
Resuscitation 2021 (IF: 6.251, 3/32) [202]

A5 *Airway strategy and ventilation rates in the Pragmatic Airway Resuscitation Trial*
Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Graham Nichol, Mohamud R. Daya, Matthew L. Hansen, Michelle M.J. Nassal,

Ashish R. Panchal, Dhimitri A. Nikolla, Erik Alonso, Jestin N. Carlson, Robert H. Schmicker, Shannon W. Stephens, Unai Irusta, Ahamed H. Idris
Resuscitation 2022 (IF: 6.500, 2/32) [201]

A6 *Impedance-based ventilation detection and signal quality control during out-of-hospital cardiopulmonary resuscitation*
Xabier Jaureguibeitia, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris
IEEE Journal of Biomedical and Health Informatics 2023 (IF: 7.700, 3/31) [198]

6.2.2 National and international conferences

C1 *Impedance based automatic detection of ventilations during mechanical cardiopulmonary resuscitation*
Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Erik Alonso, Pamela C. Owens, Henry E. Wang, Ahamed H. Idris
41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 [203]

C2 *Automatic detection of ventilations using the thoracic impedance signal during LUCAS chest compressions*
Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Pamela C. Owens, Henry E. Wang, Ahamed H. Idris
Resuscitation Science Symposium (AHA-ReSS) 2019 [206]

C3 *An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation*
Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Erik Alonso, Pamela C. Owens, Henry E. Wang, Ahamed H. Idris
Computing in Cardiology (CinC) 2020 [204]

C4 *Thoracic impedance reflects differences between endotracheal and laryngeal advanced airway during mechanical chest compressions*
Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Ahamed H. Idris, Henry E. Wang
Resuscitation Science Symposium (AHA-ReSS) 2020 [214]

C5 *Effect of airway strategy upon chest compression quality in the Pragmatic Airway Resuscitation Trial*

Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Graham Nichol, Mohamud R. Daya, Matthew L. Hansen, Michelle M.J. Nassal, Ashish R. Panchal, Dhimitri A. Nikolla, Erik Alonso, Jestin N. Carlson, Robert H. Schmicker, Shannon W. Stephens, Unai Irusta, Ahamed H. Idris
Resuscitation Science Symposium (AHA-ReSS) 2020 [215]

C6 *Algoritmo multietapa para la detección de ventilaciones en la impedancia torácica durante la resucitación cardiopulmonar*
Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris
XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB) 2020 [208]

C7 *Effect of airway strategy upon chest compression quality in the Pragmatic Airway Resuscitation Trial*
Henry E. Wang, Xabier Jaureguibeitia, Ahamed H. Idris, Unai Irusta, Erik Alonso, Tom P. Aufderheide, Matthew L. Hansen, Ryan M. Huebinger, Robert H. Schmicker, Jestin N. Carlson, M. Riccardo Colella, Richard Gordon, Robert Suchting, Elisabete Aramendi
Annual Meeting of the National Association of Emergency Medicine Service Physicians (NAEMSP) 2021 [216]

C8 *Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the Pragmatic Airway Resuscitation Trial*
Michell M.J. Nassal, Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Ashish R. Panchal, Henry E. Wang, Ahamed H. Idris
Resuscitation Science Symposium (AHA-ReSS) 2021 [217]

C9 *Effect of airway strategy upon ventilation rates in the Pragmatic Airway Resuscitation Trial*
Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Michelle M.J. Nassal, Ashish R. Panchal, Graham Nichol, Mohamud R. Daya, Matthew L. Hansen, Tom P. Aufderheide, Jestin N. Carlson, Dhimitri A. Nikolla, Robert H. Schmicker, Shannon W. Stephens, Unai Irusta, Erik Alonso, Ahamed H. Idris
Annual Meeting of the National Association of Emergency Medicine Service Physicians (NAEMSP) 2022 [218]

C10 *Efficacy of thoracic impedance for ventilation detection during continuous chest compressions*
Xabier Jaureguibeitia, Elisabete Aramendi, Ahamed H. Idris, Henry E. Wang
Resuscitation Science Symposium (AHA-ReSS) 2022 [207]

C11 *Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar*
Xabier Jaureguibeitia, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris
XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB) 2022 [205]

6.3 Financial support

This thesis has been primarily supported by a predoctoral grant (P1). All projects and funding sources that have supported financially the development of the thesis are acknowledged in the following:

P1 *Ayuda para la formación de personal investigador* (PRE-2019-1-0209, PRE-2020-2-0182, PRE-2021-2-0126, PRE-2022-2-0270). Basque Government Department of Education, Universities and Research. 2020-2023.

P2 *Pragmatic Airway Resuscitation Trial. CPR Process and Ventilation Ancillary Study, Phase I* (INT-UT Houston 19/01). University of Texas Health Science Center in Houston (TX, USA). January 2019 – December 2022 (22,000€).

P3 *Procesado multimodal de señal y aprendizaje automático para la mejora del tratamiento de la parada cardiorrespiratoria extrahospitalaria* (RTI2018-101475-BI00). Spanish Ministry of Economy and Competitiveness. February 2019 – September 2022 (96,000€).

P4 *BioRes (Biomedical Engineering and Resuscitation)* (IT1229-19). Basque Government Department of Education, Universities and Research. February 2019 – December 2021 (97,000€).

P5 *Multi-center observational study of the relationship of ventilation and outcomes from cardiac arrest using existing data* (R21HL1561969).

NIH – National Heart, Lung and Blood Institute, USA. February 2021 – January 2023 (22,252.96€).

P6 *Inteligencia artificial y nuevas tecnologías para el guiado de la terapia de resucitación en la parada cardiorrespiratoria extrahospitalaria* (PID2021-122727OB-I00). Spanish Ministry of Science, Research and Universities. September 2022 – August 2025 (151,250€).

P7 *BioRes (Biomedical Engineering and Resuscitation)* (IT1717-22). Basque Government Department of Education, Universities and Research. January 2022 – December 2025 (81,200€).

6.4 Future lines of research

The thesis has contributed to the characterization of ventilation and CPR in large OHCA datasets, and has helped answer some critical questions regarding ventilation and airway management strategies in resuscitation. As part of the process, new questions have arisen, and new opportunities have been identified to improve or complement the current solutions, all of which could be object of future research efforts.

- The proposed solution for TI-based ventilation detection during manual CC-CPR relied on force and acceleration signal data which may often not be available. Different filtering approaches should be explored to enable a broader applicability of the solution. Alternative approaches to time series classification, such as conditional random fields, transformers or attention mechanisms, could also be explored.
- Deep learning models showed also promising results for reference-free TI-based ventilation detection. Data augmentation techniques and more recent segmentation architectures [219, 220] could help close, or even turn around, the current performance gap with other solutions.
- The simple quality control model designed for TI-based ventilation detection proved effective, but could be potentially improved using more sophisticated techniques [212, 221]. The availability of two different ventilation detection solutions with similar performance — the Kalman/RNN ones [198] and the deep learning one [205] —

could allow the design of more universally applicable quality control models, and posterior refinement of the individual solutions.

- The proposed methodology for automatic CPR analysis in large multi-device datasets [199] yielded acceptable errors for the most part, but presented large errors in about 10% of the cases, mainly due to incomplete ROSC information. Automatic pulse detection solutions available in the literature [124, 222] could be integrated to potentially minimize these errors. The concept of signal quality control could also be extended to chest compression detection to prevent additional errors. A similar framework could be developed for ventilation analysis, considering both TI- and capnogram-based solutions.
- The analyses on CPR quality in the PART database did not reveal major differences apart from an increased ventilation TI amplitude and a delayed airway insertion for ETI cases. Outcome differences favoring LT may have been related to ventilation differences before and during airway insertion, which were not analyzed. Other analyses of ventilation are possible considering both the capnogram and the TI in order to highlight outcome related variables.

BIBLIOGRAPHY

[1] I. Jacobs, V. Nadkarni, J. Bahr, R. A. Berg, J. E. Billi, L. Bossaert *et al.*, "Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the utstein templates for resuscitation registries: a statement for healthcare professionals from a task force of the international liaison committee on resuscitation (american heart association, european resuscitation council, australian resuscitation council, new zealand resuscitation council, heart and stroke foundation of canada, interamerican heart foundation, resuscitation councils of southern africa)." *Circulation*, vol. 110, pp. 3385–3397, Nov. 2004.

[2] D. P. Zipes and H. J. Wellens, "Sudden cardiac death." *Circulation*, vol. 98, pp. 2334–2351, Nov. 1998.

[3] R. M. Merchant, A. A. Topjian, A. R. Panchal, A. Cheng, K. Aziz, K. M. Berg *et al.*, "Part 1: Executive summary: 2020 american heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care." *Circulation*, vol. 142, pp. S337–S357, Oct. 2020.

[4] F. Semeraro, R. Greif, B. W. Böttiger, R. Burkart, D. Cimpoesu, M. Georgiou *et al.*, "European resuscitation council guidelines 2021: Systems saving lives." *Resuscitation*, vol. 161, pp. 80–97, Apr. 2021.

[5] G. D. Perkins, I. G. Jacobs, V. M. Nadkarni, R. A. Berg, F. Bhanji, D. Biarent *et al.*, "Cardiac arrest and cardiopulmonary resuscitation outcome reports: Update of the utstein resuscitation registry templates for out-of-hospital cardiac arrest: A statement for healthcare professionals from a task force of the international liaison committee on resuscitation (american heart association, european resuscitation council, australian and new zealand council on resuscitation, heart and stroke foundation of canada, interamerican heart foundation, resuscitation council of southern africa, resuscitation council of asia); and the american heart association emergency cardiovascular care

committee and the council on cardiopulmonary, critical care, perioperative and resuscitation." *Resuscitation*, vol. 96, pp. 328–340, Nov. 2015.

[6] A. Claesson, T. Djarv, P. Nordberg, M. Ringh, J. Hollenberg, C. Axelsson *et al.*, "Medical versus non medical etiology in out-of-hospital cardiac arrest-changes in outcome in relation to the revised utstein template." *Resuscitation*, vol. 110, pp. 48–55, Jan. 2017.

[7] J.-T. Gräsner, J. Wnent, J. Herlitz, G. D. Perkins, R. Lefering, I. Tjelmeland *et al.*, "Survival after out-of-hospital cardiac arrest in europe - results of the eureca two study." *Resuscitation*, vol. 148, pp. 218–226, Mar. 2020.

[8] C. Hawkes, S. Booth, C. Ji, S. J. Brace-McDonnell, A. Whittington, J. Mapstone *et al.*, "Epidemiology and outcomes from out-of-hospital cardiac arrests in england." *Resuscitation*, vol. 110, pp. 133–140, Jan. 2017.

[9] M. Kuisma and A. Alaspää, "Out-of-hospital cardiac arrests of non-cardiac origin. epidemiology and outcome." *European heart journal*, vol. 18, pp. 1122–1128, Jul. 1997.

[10] M. Hayashi, W. Shimizu, and C. M. Albert, "The spectrum of epidemiology underlying sudden cardiac death." *Circulation research*, vol. 116, pp. 1887–1906, Jun. 2015.

[11] M. Rubart and D. P. Zipes, "Mechanisms of sudden cardiac death." *The Journal of clinical investigation*, vol. 115, pp. 2305–2315, Sep. 2005.

[12] M. L. Weisfeldt and L. B. Becker, "Resuscitation after cardiac arrest: a 3-phase time-sensitive model." *JAMA*, vol. 288, pp. 3035–3038, Dec. 2002.

[13] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt *et al.*, "Heart disease and stroke statistics-2022 update: A report from the american heart association." *Circulation*, vol. 145, pp. e153–e639, Feb. 2022.

[14] H. J. J. Wellens, P. J. Schwartz, F. W. Lindemans, A. E. Buxton, J. J. Goldberger, S. H. Hohnloser *et al.*, "Risk stratification for sudden cardiac death: current status and challenges for the future." *European heart journal*, vol. 35, pp. 1642–1651, Jul. 2014.

[15] J.-T. Gräsner, J. Herlitz, I. B. M. Tjelmeland, J. Wnent, S. Masterson, G. Lilja *et al.*, "European resuscitation council guidelines 2021: Epidemiology of cardiac arrest in europe." *Resuscitation*, vol. 161, pp. 61–79, Apr. 2021.

[16] S. Ballesteros-Peña, L. C. Abecia-Inchaurregui, and E. Echevarría-Orella, "Factors associated with mortality in out-of-hospital cardiac arrests attended in basic life support units in the basque country (spain)." *Revista espanola de cardiologia (English ed.)*, vol. 66, pp. 269–274, Apr. 2013.

[17] A. R. Panchal, J. A. Bartos, J. G. Cabañas, M. W. Donnino, I. R. Drennan, K. G. Hirsch *et al.*, "Part 3: Adult basic and advanced life support: 2020 american heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care." *Circulation*, vol. 142, pp. S366–S468, Oct. 2020.

[18] R. O. Cummins, J. P. Ornato, W. H. Thies, and P. E. Pepe, "Improving survival from sudden cardiac arrest: the "chain of survival" concept. a statement for health professionals from the advanced cardiac life support subcommittee and the emergency cardiac care committee, american heart association." *Circulation*, vol. 83, pp. 1832–1847, May 1991.

[19] A. Langhelle, J. Nolan, J. Herlitz, M. Castren, V. Wenzel, E. Soreide *et al.*, "Recommended guidelines for reviewing, reporting, and conducting research on post-resuscitation care: the utstein style." *Resuscitation*, vol. 66, pp. 271–283, Sep. 2005.

[20] J. Nolan, J. Soar, and H. Eikeland, "The chain of survival." *Resuscitation*, vol. 71, pp. 270–271, Dec. 2006.

[21] C. Sasson, M. A. M. Rogers, J. Dahl, and A. L. Kellermann, "Predictors of survival from out-of-hospital cardiac arrest: a systematic review and meta-analysis." *Circulation. Cardiovascular quality and outcomes*, vol. 3, pp. 63–81, Jan. 2010.

[22] J. Holmén, J. Herlitz, S.-E. Ricksten, A. Strömsöe, E. Hagberg, C. Axelsson *et al.*, "Shortening ambulance response time increases survival in out-of-hospital cardiac arrest." *Journal of the American Heart Association*, vol. 9, p. e017048, Nov. 2020.

[23] I. Hasselqvist-Ax, G. Riva, J. Herlitz, M. Rosenqvist, J. Hollenberg, P. Nordberg *et al.*, "Early cardiopulmonary resuscitation in out-of-hospital cardiac arrest." *The New England journal of medicine*, vol. 372, pp. 2307–2315, Jun. 2015.

[24] S. Rajan, M. Wissenberg, F. Folke, S. M. Hansen, T. A. Gerds, K. Kragholm *et al.*, "Association of bystander cardiopulmonary resuscitation and survival according to ambulance response times after out-of-hospital cardiac arrest." *Circulation*, vol. 134, pp. 2095–2104, Dec. 2016.

[25] T. D. Rea, M. S. Eisenberg, L. L. Culley, and L. Becker, "Dispatcher-assisted cardiopulmonary resuscitation and survival in cardiac arrest." *Circulation*, vol. 104, pp. 2513–2516, Nov. 2001.

[26] K. M. Berg, A. Cheng, A. R. Panchal, A. A. Topjian, K. Aziz, F. Bhanji *et al.*, "Part 7: Systems of care: 2020 american heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care." *Circulation*, vol. 142, pp. S580–S604, Oct. 2020.

[27] L. L. Bossaert, "Fibrillation and defibrillation of the heart." *British journal of anaesthesia*, vol. 79, pp. 203–213, Aug. 1997.

[28] T. D. Valenzuela, D. J. Roe, G. Nichol, L. L. Clark, D. W. Spaite, and R. G. Hardman, "Outcomes of rapid defibrillation by security officers after cardiac arrest in casinos." *The New England journal of medicine*, vol. 343, pp. 1206–1209, Oct. 2000.

[29] M. T. Blom, S. G. Beesems, P. C. M. Homma, J. A. Zijlstra, M. Hulleman, D. A. van Hoeijen *et al.*, "Improved survival after out-of-hospital cardiac arrest and use of automated external defibrillators." *Circulation*, vol. 130, pp. 1868–1875, Nov. 2014.

[30] M. L. Weisfeldt, R. E. Kerber, R. P. McGoldrick, A. J. Moss, G. Nichol, J. P. Ornato *et al.*, "Public access defibrillation. a statement for healthcare professionals from the american heart association task force on automatic external defibrillation." *Circulation*, vol. 92, p. 2763, Nov. 1995.

[31] "Guidelines 2000 for cardiopulmonary resuscitation and emergency cardiovascular care. part 4: the automated external defibrillator: key link in the chain of survival. the american heart association in collaboration with the international liaison committee on resuscitation." *Circulation*, vol. 102, pp. I60–I76, Aug. 2000.

[32] T. D. Valenzuela, D. J. Roe, S. Cretin, D. W. Spaite, and M. P. Larsen, "Estimating effectiveness of cardiac arrest interventions: a logistic regression survival model." *Circulation*, vol. 96, pp. 3308–3313, Nov. 1997.

[33] G. D. Perkins, A. J. Handley, R. W. Koster, M. Castrén, M. A. Smyth, T. Olasveengen *et al.*, "European resuscitation council guidelines for resuscitation 2015: Section 2. adult basic life support and automated external defibrillation." *Resuscitation*, vol. 95, pp. 81–99, Oct. 2015.

[34] M. P. Larsen, M. S. Eisenberg, R. O. Cummins, and A. P. Hallstrom, "Predicting survival from out-of-hospital cardiac arrest: a graphic model." *Annals of emergency medicine*, vol. 22, pp. 1652–1658, Nov.

1993.

[35] R. A. Waalewijn, R. de Vos, J. G. Tijssen, and R. W. Koster, "Survival models for out-of-hospital cardiopulmonary resuscitation from the perspectives of the bystander, the first responder, and the paramedic." *Resuscitation*, vol. 51, pp. 113–122, Nov. 2001.

[36] J. Soar, B. W. Böttiger, P. Carli, K. Couper, C. D. Deakin, T. Djärv *et al.*, "European resuscitation council guidelines 2021: Adult advanced life support." *Resuscitation*, vol. 161, pp. 115–151, Apr. 2021.

[37] J. P. Nolan, C. Sandroni, B. W. Böttiger, A. Cariou, T. Cronberg, H. Friberg *et al.*, "European resuscitation council and european society of intensive care medicine guidelines 2021: Post-resuscitation care." *Resuscitation*, vol. 161, pp. 220–269, Apr. 2021.

[38] G. D. Perkins, J.-T. Graesner, F. Semeraro, T. Olasveengen, J. Soar, C. Lott *et al.*, "European resuscitation council guidelines 2021: Executive summary." *Resuscitation*, vol. 161, pp. 1–60, Apr. 2021.

[39] R. W. Koster, M. A. Baubin, L. L. Bossaert, A. Caballero, P. Cassan, M. Castrén *et al.*, "European resuscitation council guidelines for resuscitation 2010 section 2. adult basic life support and use of automated external defibrillators." *Resuscitation*, vol. 81, pp. 1277–1292, Oct. 2010.

[40] P. SAFAR, T. C. BROWN, W. J. HOLTEY, and R. J. WILDER, "Ventilation and circulation with closed-chest cardiac massage in man." *JAMA*, vol. 176, pp. 574–576, May 1961.

[41] M. Georgiou, E. Papathanassoglou, and T. Xanthos, "Systematic review of the mechanisms driving effective blood flow during adult cpr." *Resuscitation*, vol. 85, pp. 1586–1593, Nov. 2014.

[42] M. R. Neth, A. Idris, J. McMullan, J. L. Benoit, and M. R. Daya, "A review of ventilation in adult out-of-hospital cardiac arrest." *Journal of the American College of Emergency Physicians open*, vol. 1, pp. 190–201, Jun. 2020.

[43] R. Fowler, M. P. Chang, and A. H. Idris, "Evolution and revolution in cardiopulmonary resuscitation." *Current opinion in critical care*, vol. 23, pp. 183–187, Jun. 2017.

[44] T. M. Olasveengen, F. Semeraro, G. Ristagno, M. Castren, A. Handley, A. Kuzovlev *et al.*, "European resuscitation council guidelines 2021: Basic life support." *Resuscitation*, vol. 161, pp. 98–114, Apr. 2021.

[45] D. Yannopoulos, S. McKnite, T. P. Aufderheide, G. Sigurdsson, R. G. Pirrallo, D. Benditt *et al.*, "Effects of incomplete chest wall decompression during cardiopulmonary resuscitation on coronary and cerebral

perfusion pressures in a porcine model of cardiac arrest." *Resuscitation*, vol. 64, pp. 363–372, Mar. 2005.

[46] N. T. Sugerman, D. P. Edelson, M. Leary, E. K. Weidman, D. L. Herzberg, T. L. Vanden Hoek *et al.*, "Rescuer fatigue during actual in-hospital cardiopulmonary resuscitation with audiovisual feedback: a prospective multicenter study." *Resuscitation*, vol. 80, pp. 981–984, Sep. 2009.

[47] P. L. Wang and S. C. Brooks, "Mechanical versus manual chest compressions for cardiac arrest." *The Cochrane database of systematic reviews*, vol. 8, p. CD007260, Aug. 2018.

[48] V. Dörges, H. Ocker, S. Hagelberg, V. Wenzel, and P. Schmucker, "Optimisation of tidal volumes given with self-inflatable bags without additional oxygen." *Resuscitation*, vol. 43, pp. 195–199, Feb. 2000.

[49] J. Kramer-Johansen, H. Myklebust, L. Wik, B. Fellows, L. Svensson, H. Sørebø *et al.*, "Quality of out-of-hospital cardiopulmonary resuscitation with real time automated feedback: a prospective interventional study." *Resuscitation*, vol. 71, pp. 283–292, Dec. 2006.

[50] I. G. Stiell, S. P. Brown, J. Christenson, S. Cheskes, G. Nichol, J. Powell *et al.*, "What is the role of chest compression depth during out-of-hospital cardiac arrest resuscitation?" *Critical care medicine*, vol. 40, pp. 1192–1198, Apr. 2012.

[51] H. Hellevuo, M. Sainio, R. Nevalainen, H. Huhtala, K. T. Olkkola, J. Tenhunen *et al.*, "Deeper chest compression - more complications for cardiac arrest patients?" *Resuscitation*, vol. 84, pp. 760–765, Jun. 2013.

[52] B. S. Abella, N. Sandbo, P. Vassilatos, J. P. Alvarado, N. O'Hearn, H. N. Wigder *et al.*, "Chest compression rates during cardiopulmonary resuscitation are suboptimal: a prospective study during in-hospital cardiac arrest." *Circulation*, vol. 111, pp. 428–434, Feb. 2005.

[53] A. H. Idris, D. Guffey, T. P. Aufderheide, S. Brown, L. J. Morrison, P. Nichols *et al.*, "Relationship between chest compression rates and outcomes from cardiac arrest." *Circulation*, vol. 125, pp. 3004–3012, Jun. 2012.

[54] K. G. Lurie, E. C. Nemergut, D. Yannopoulos, and M. Sweeney, "The physiology of cardiopulmonary resuscitation." *Anesthesia and analgesia*, vol. 122, pp. 767–783, Mar. 2016.

[55] K. G. Monsieurs, M. De Regge, K. Vansteelandt, J. De Smet, E. Annaert, S. Lemoyne *et al.*, "Excessive chest compression rate is asso-

ciated with insufficient compression depth in prehospital cardiac arrest." *Resuscitation*, vol. 83, pp. 1319–1323, Nov. 2012.

[56] T. Eftestøl, K. Sunde, and P. A. Steen, "Effects of interrupting precordial compressions on the calculated probability of defibrillation success during out-of-hospital cardiac arrest." *Circulation*, vol. 105, pp. 2270–2273, May 2002.

[57] S. Cheskes, R. H. Schmicker, J. Christenson, D. D. Salcido, T. Rea, J. Powell *et al.*, "Perishock pause: an independent predictor of survival from out-of-hospital shockable cardiac arrest." *Circulation*, vol. 124, pp. 58–66, Jul. 2011.

[58] J. Kramer-Johansen, D. P. Edelson, H. Losert, K. Köhler, and B. S. Abella, "Uniform reporting of measured quality of cardiopulmonary resuscitation (cpr)." *Resuscitation*, vol. 74, pp. 406–417, Sep. 2007.

[59] J. Christenson, D. Andrusiek, S. Everson-Stewart, P. Kudenchuk, D. Hostler, J. Powell *et al.*, "Chest compression fraction determines survival in patients with out-of-hospital ventricular fibrillation." *Circulation*, vol. 120, pp. 1241–1247, Sep. 2009.

[60] T. Rea, M. Olsufka, L. Yin, C. Maynard, and L. Cobb, "The relationship between chest compression fraction and outcome from ventricular fibrillation arrests in prolonged resuscitations." *Resuscitation*, vol. 85, pp. 879–884, Jul. 2014.

[61] T. P. Aufderheide, G. Sigurdsson, R. G. Pirrallo, D. Yannopoulos, S. McKnite, C. von Briesen *et al.*, "Hyperventilation-induced hypotension during cardiopulmonary resuscitation." *Circulation*, vol. 109, pp. 1960–1965, Apr. 2004.

[62] B. J. Bobrow, D. W. Spaite, R. A. Berg, U. Stolz, A. B. Sanders, K. B. Kern *et al.*, "Chest compression-only cpr by lay rescuers and survival from out-of-hospital cardiac arrest." *JAMA*, vol. 304, pp. 1447–1454, Oct. 2010.

[63] T. Iwami, T. Kitamura, K. Kiyohara, and T. Kawamura, "Dissemination of chest compression-only cardiopulmonary resuscitation and survival after out-of-hospital cardiac arrest." *Circulation*, vol. 132, pp. 415–422, Aug. 2015.

[64] B. J. Bobrow, L. L. Clark, G. A. Ewy, V. Chikani, A. B. Sanders, R. A. Berg *et al.*, "Minimally interrupted cardiac resuscitation by emergency medical services for out-of-hospital cardiac arrest." *JAMA*, vol. 299, pp. 1158–1165, Mar. 2008.

[65] N. C. Chandra, K. G. Gruben, J. E. Tsitlik, R. Brower, A. D. Guerci, H. H. Halperin *et al.*, "Observations of ventilation during resuscitation in a canine model." *Circulation*, vol. 90, pp. 3070–3075, Dec. 1994.

[66] C. D. Deakin, J. F. O'Neill, and T. Tabor, "Does compression-only cardiopulmonary resuscitation generate adequate passive ventilation during cardiac arrest?" *Resuscitation*, vol. 75, pp. 53–59, Oct. 2007.

[67] T. D. Rea, C. Fahrenbruch, L. Culley, R. T. Donohoe, C. Hambly, J. Innes *et al.*, "Cpr with chest compression alone or with rescue breathing." *The New England journal of medicine*, vol. 363, pp. 423–433, Jul. 2010.

[68] B. J. Bobrow, M. Zuercher, G. A. Ewy, L. Clark, V. Chikani, D. Donahue *et al.*, "Gasping during cardiac arrest in humans is frequent and associated with improved survival." *Circulation*, vol. 118, pp. 2550–2554, Dec. 2008.

[69] C. F. Babbs and K. B. Kern, "Optimum compression to ventilation ratios in cpr under realistic, practical conditions: a physiological and mathematical analysis," *Resuscitation*, vol. 54, no. 2, pp. 147–157, 2002.

[70] H. M. Ashoor, E. Lillie, W. Zarin, P. A. Khan, V. Nincic, F. Yazdi *et al.*, "Effectiveness of different compression-to-ventilation methods for cardiopulmonary resuscitation: a systematic review," *Resuscitation*, vol. 118, pp. 112–125, 2017.

[71] T. P. Aufderheide and K. G. Lurie, "Death by hyperventilation: a common and life-threatening problem during cardiopulmonary resuscitation." *Critical care medicine*, vol. 32, pp. S345–S351, Sep. 2004.

[72] K. G. Lurie, D. Yannopoulos, S. H. McKnite, M. L. Herman, A. H. Idris, V. M. Nadkarni *et al.*, "Comparison of a 10-breaths-per-minute versus a 2-breaths-per-minute strategy during cardiopulmonary resuscitation in a porcine model of cardiac arrest," *Respiratory care*, vol. 53, no. 7, pp. 862–870, 2008.

[73] G. Vissers, J. Soar, and K. G. Monsieurs, "Ventilation rate in adults with a tracheal tube during cardiopulmonary resuscitation: a systematic review," *Resuscitation*, vol. 119, pp. 5–12, 2017.

[74] G. Vissers, C. Duchatelet, S. Huybrechts, K. Wouters, S. Hachimi-Idrissi, and K. Monsieurs, "The effect of ventilation rate on outcome in adults receiving cardiopulmonary resuscitation," *Resuscitation*, vol. 138, pp. 243–249, 2019.

[75] G. Nichol, B. Leroux, H. Wang, C. W. Callaway, G. Sopko, M. Weisfeldt *et al.*, "Trial of continuous or interrupted chest compressions during

cpr," *New England Journal of Medicine*, vol. 373, no. 23, pp. 2203–2214, 2015.

[76] R. H. Schmicker, G. Nichol, P. Kudenchuk, J. Christenson, C. Vaillancourt, H. E. Wang *et al.*, "Cpr compression strategy 30:2 is difficult to adhere to, but has better survival than continuous chest compressions when done correctly." *Resuscitation*, vol. 165, pp. 31–37, Aug. 2021.

[77] E. M. Rottenberg, "Continuous or interrupted chest compressions during cpr." *The New England journal of medicine*, vol. 374, p. 1194, Mar. 2016.

[78] C. Duchatelet, A. F. Kalmar, K. G. Monsieurs, and S. Hachimi-Idrissi, "Chest compressions during ventilation in out-of-hospital cardiac arrest cause reversed airflow," *Resuscitation*, vol. 129, pp. 97–102, 2018.

[79] M. P. Chang, Y. Lu, B. Leroux, E. Aramendi Ecenarro, P. Owens, H. E. Wang *et al.*, "Association of ventilation with outcomes from out-of-hospital cardiac arrest." *Resuscitation*, vol. 141, pp. 174–181, Aug. 2019.

[80] J. L. Benoit, D. K. Prince, and H. E. Wang, "Mechanisms linking advanced airway management and cardiac arrest outcomes." *Resuscitation*, vol. 93, pp. 124–127, Aug. 2015.

[81] H. E. Wang, S. J. Simeone, M. D. Weaver, and C. W. Callaway, "Interruptions in cardiopulmonary resuscitation from paramedic endotracheal intubation." *Annals of emergency medicine*, vol. 54, pp. 645–652.e1, Nov. 2009.

[82] A. F. Jarman, C. L. Hopkins, J. N. Hansen, J. R. Brown, C. Burk, and S. T. Youngquist, "Advanced airway type and its association with chest compression interruptions during out-of-hospital cardiac arrest resuscitation attempts." *Prehospital emergency care*, vol. 21, pp. 628–635, 2017.

[83] K. Hasegawa, A. Hiraide, Y. Chang, and D. F. M. Brown, "Association of prehospital advanced airway management with neurologic outcome and survival in patients with out-of-hospital cardiac arrest." *JAMA*, vol. 309, pp. 257–266, Jan. 2013.

[84] J. McMullan, R. Gerecht, J. Bonomo, R. Robb, B. McNally, J. Donnelly *et al.*, "Airway management and out-of-hospital cardiac arrest outcome in the cares registry." *Resuscitation*, vol. 85, pp. 617–622, May 2014.

[85] C. Newell, S. Grier, and J. Soar, "Airway and ventilation management during cardiopulmonary resuscitation and after successful resuscitation." *Critical care (London, England)*, vol. 22, p. 190, Aug. 2018.

[86] C.-H. Wang, W.-J. Chen, W.-T. Chang, M.-S. Tsai, P.-H. Yu, Y.-W. Wu *et al.*, "The association between timing of tracheal intubation and outcomes of adult in-hospital cardiac arrest: A retrospective cohort study." *Resuscitation*, vol. 105, pp. 59–65, Aug. 2016.

[87] J. Izawa, S. Komukai, K. Gibo, M. Okubo, K. Kiyohara, C. Nishiyama *et al.*, "Pre-hospital advanced airway management for adults with out-of-hospital cardiac arrest: nationwide cohort study." *BMJ (Clinical research ed.)*, vol. 364, p. l430, Feb. 2019.

[88] P. Jabre, A. Penaloza, D. Pinero, F.-X. Duchateau, S. W. Borron, F. Javaudin *et al.*, "Effect of bag-mask ventilation vs endotracheal intubation during cardiopulmonary resuscitation on neurological outcome after out-of-hospital cardiorespiratory arrest: A randomized clinical trial." *JAMA*, vol. 319, pp. 779–787, Feb. 2018.

[89] J. L. Jarvis, D. Wampler, and H. E. Wang, "Association of patient age with first pass success in out-of-hospital advanced airway management." *Resuscitation*, vol. 141, pp. 136–143, Aug. 2019.

[90] J. L. Benoit, R. B. Gerecht, M. T. Steuerwald, and J. T. McMullan, "Endotracheal intubation versus supraglottic airway placement in out-of-hospital cardiac arrest: A meta-analysis." *Resuscitation*, vol. 93, pp. 20–26, Aug. 2015.

[91] J. R. Benger, K. Kirby, S. Black, S. J. Brett, M. Clout, M. J. Lazaroo *et al.*, "Effect of a strategy of a supraglottic airway device vs tracheal intubation during out-of-hospital cardiac arrest on functional outcome: The airways-2 randomized clinical trial." *JAMA*, vol. 320, pp. 779–791, Aug. 2018.

[92] H. E. Wang, R. H. Schmicker, M. R. Daya, S. W. Stephens, A. H. Idris, J. N. Carlson *et al.*, "Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: A randomized clinical trial." *JAMA*, vol. 320, pp. 769–778, Aug. 2018.

[93] L. Wik, J. Kramer-Johansen, H. Myklebust, H. Sørebø, L. Svensson, B. Fellows *et al.*, "Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest." *JAMA*, vol. 293, pp. 299–304, Jan. 2005.

[94] B. S. Abella, J. P. Alvarado, H. Myklebust, D. P. Edelson, A. Barry, N. O'Hearn *et al.*, "Quality of cardiopulmonary resuscitation during

in-hospital cardiac arrest." *JAMA*, vol. 293, pp. 305–310, Jan. 2005.

[95] J. F. O'Neill and C. D. Deakin, "Do we hyperventilate cardiac arrest patients?" *Resuscitation*, vol. 73, pp. 82–85, Apr. 2007.

[96] D. P. Edelson, B. Litzinger, V. Arora, D. Walsh, S. Kim, D. S. Lauderdale *et al.*, "Improving in-hospital cardiac arrest process and outcomes with performance debriefing." *Archives of internal medicine*, vol. 168, pp. 1063–1069, May 2008.

[97] T. M. Olasveengen, M. E. Mancini, G. D. Perkins, S. Avis, S. Brooks, M. Castrén *et al.*, "Adult basic life support: 2020 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations." *Circulation*, vol. 142, pp. S41–S91, Oct. 2020.

[98] D. Hostler, S. Everson-Stewart, T. D. Rea, I. G. Stiell, C. W. Callaway, P. J. Kudenchuk *et al.*, "Effect of real-time feedback during cardiopulmonary resuscitation outside hospital: prospective, cluster-randomised trial." *BMJ (Clinical research ed.)*, vol. 342, p. d512, Feb. 2011.

[99] E. Alonso, J. Ruiz, E. Aramendi, D. González-Otero, S. Ruiz de Gauna, U. Ayala *et al.*, "Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics." *Resuscitation*, vol. 88, pp. 28–34, Mar. 2015.

[100] U. Ayala, T. Eftestøl, E. Alonso, U. Irusta, E. Aramendi, S. Wali *et al.*, "Automatic detection of chest compressions for the assessment of cpr-quality parameters." *Resuscitation*, vol. 85, pp. 957–963, Jul. 2014.

[101] D. M. González-Otero, S. R. de Gauna, J. Ruiz, M. R. Daya, L. Wik, J. K. Russell *et al.*, "Chest compression rate feedback based on transthoracic impedance," *Resuscitation*, vol. 93, pp. 82–88, 2015.

[102] H. Kwok, J. Coult, C. Liu, J. Blackwood, P. J. Kudenchuk, T. D. Rea *et al.*, "An accurate method for real-time chest compression detection from the impedance signal," *Resuscitation*, vol. 105, pp. 22–28, 2016.

[103] S. O. Aase and H. Myklebust, "Compression depth estimation for cpr quality assessment using DSP on accelerometer signals," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 3, pp. 263–268, 2002.

[104] D. M. González-Otero, J. Ruiz, S. Ruiz de Gauna, U. Irusta, U. Ayala, and E. Alonso, "A new method for feedback on the quality of chest compressions during cardiopulmonary resuscitation," *BioMed research international*, vol. 2014, 2014.

[105] M. Risdal, S. O. Aase, M. Stavland, and T. Eftestøl, "Impedance-based ventilation detection during cardiopulmonary resuscitation." *IEEE transactions on bio-medical engineering*, vol. 54, pp. 2237–2245, Dec. 2007.

[106] D. P. Edelson, J. Eilevstjønn, E. K. Weidman, E. Retzer, T. L. V. Hoek, and B. S. Abella, "Capnography and chest-wall impedance algorithms for ventilation detection during cardiopulmonary resuscitation." *Resuscitation*, vol. 81, pp. 317–322, Mar. 2010.

[107] E. Aramendi, A. Elola, E. Alonso, U. Irusta, M. Daya, J. K. Russell *et al.*, "Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation." *Resuscitation*, vol. 110, pp. 162–168, Jan. 2017.

[108] M. Leturiondo, S. Ruiz de Gauna, J. M. Ruiz, J. Julio Gutiérrez, L. A. Leturiondo, D. M. González-Otero *et al.*, "Influence of chest compression artefact on capnogram-based ventilation detection during out-of-hospital cardiopulmonary resuscitation." *Resuscitation*, vol. 124, pp. 63–68, Mar. 2018.

[109] E. Aramendi, Y. Lu, M. P. Chang, A. Elola, U. Irusta, P. Owens *et al.*, "A novel technique to assess the quality of ventilation during prehospital cardiopulmonary resuscitation." *Resuscitation*, vol. 132, pp. 41–46, Nov. 2018.

[110] R. E. Kerber, L. B. Becker, J. D. Bourland, R. O. Cummins, A. P. Hallstrom, M. B. Michos *et al.*, "Automatic external defibrillators for public access defibrillation: recommendations for specifying and reporting arrhythmia analysis algorithm performance, incorporating new waveforms, and enhancing safety. a statement for health professionals from the american heart association task force on automatic external defibrillation, subcommittee on aed safety and efficacy." *Circulation*, vol. 95, pp. 1677–1682, Mar. 1997.

[111] M. T. Nguyen, T.-H. T. Nguyen, and H.-C. Le, "A review of progress and an advanced method for shock advice algorithms in automated external defibrillators." *Biomedical engineering online*, vol. 21, p. 22, Apr. 2022.

[112] W. Kubicek, D. Witsoe, R. Patterson *et al.*, "Development and evaluation of an impedance cardiographic system to measure cardiac output and other cardiac parameters: Final progress report (nasa 9-4500)," *National Aeronautics and Space Administration*, vol. 170, pp. 724–732, 1970.

[113] F. S. Stecher, J.-A. Olsen, R. E. Stickney, and L. Wik, "Transthoracic impedance used to evaluate performance of cardiopulmonary resuscitation during out of hospital cardiac arrest." *Resuscitation*, vol. 79, pp. 432–437, Dec. 2008.

[114] T. Pellis, J. Bisera, W. Tang, and M. H. Weil, "Expanding automatic external defibrillators to include automated detection of cardiac, respiratory, and cardiorespiratory arrest." *Critical care medicine*, vol. 30, pp. S176–S178, Apr. 2002.

[115] H. Losert, M. Risdal, F. Sterz, J. Nysæther, K. Köhler, T. Eftestøl *et al.*, "Thoracic impedance changes measured via defibrillator pads can monitor ventilation in critically ill patients and during cardiopulmonary resuscitation," *Critical care medicine*, vol. 34, no. 9, pp. 2399–2405, 2006.

[116] M. R. Daya, R. H. Schmicker, D. M. Zive, T. D. Rea, G. Nichol, J. E. Buick *et al.*, "Out-of-hospital cardiac arrest survival improving over time: Results from the resuscitation outcomes consortium (roc)." *Resuscitation*, vol. 91, pp. 108–115, Jun. 2015.

[117] L. J. Morrison, G. Nichol, T. D. Rea, J. Christenson, C. W. Callaway, S. Stephens *et al.*, "Rationale, development and implementation of the resuscitation outcomes consortium epistry-cardiac arrest." *Resuscitation*, vol. 78, pp. 161–169, Aug. 2008.

[118] B. McNally, A. Stokes, A. Crouch, A. L. Kellermann, and C. S. Group, "Cares: Cardiac arrest registry to enhance survival." *Annals of emergency medicine*, vol. 54, pp. 674–683.e2, Nov. 2009.

[119] J. T. Gräsner, J. Herlitz, R. W. Koster, F. Rosell-Ortiz, L. Stamatakis, and L. Bossaert, "Quality management in resuscitation–towards a european cardiac arrest registry (eureca)." *Resuscitation*, vol. 82, pp. 989–994, Aug. 2011.

[120] R. M. Lyngby, L. Clark, J. S. Kjoelbye, R. M. Oelrich, A. Silver, H. C. Christensen *et al.*, "Higher resuscitation guideline adherence in paramedics with use of real-time ventilation feedback during simulated out-of-hospital cardiac arrest: A randomised controlled trial." *Resuscitation plus*, vol. 5, p. 100082, Mar. 2021.

[121] B. Y. Yang, J. E. Blackwood, J. Shin, S. Guan, M. Gao, D. B. Jorgenson *et al.*, "A pilot evaluation of respiratory mechanics during prehospital manual ventilation." *Resuscitation*, vol. 177, pp. 55–62, Aug. 2022.

[122] S. Verscheure, P. B. Massion, F. Verschuren, P. Damas, and S. Magder, "Volumetric capnography: lessons from the past and current clinical

applications." *Critical care (London, England)*, vol. 20, p. 184, Jun. 2016.

[123] I. Isasi, U. Irusta, A. Bahrami Rad, E. Aramendi, M. Zabihi, T. Eftestøl *et al.*, "Automatic cardiac rhythm classification with concurrent manual chest compressions," *IEEE Access*, vol. 7, pp. 115 147–115 159, 2019.

[124] A. Elola, E. Aramendi, U. Irusta, P. O. Berve, and L. Wik, "Multimodal algorithms for the classification of circulation states during out-of-hospital cardiac arrest," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 6, pp. 1913–1922, 2021.

[125] A. Picon, U. Irusta, A. Álvarez Gila, E. Aramendi, F. Alonso-Atienza, C. Figuera *et al.*, "Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia." *PloS one*, vol. 14, p. e0216756, 2019.

[126] S. Ruiz de Gauna, D. M. González-Otero, J. Ruiz, and J. K. Russell, "Feedback on the rate and depth of chest compressions during cardiopulmonary resuscitation using only accelerometers," *PloS one*, vol. 11, no. 3, p. e0150139, 2016.

[127] M. Iyanaga, R. Gray, S. W. Stephens, O. Akinsanya, J. Rodgers, K. Smyrski *et al.*, "Comparison of methods for the determination of cardiopulmonary resuscitation chest compression fraction." *Resuscitation*, vol. 83, pp. 568–571, May 2012.

[128] M. C. Kurz, D. K. Prince, J. Christenson, J. Carlson, D. Stub, S. Cheskes *et al.*, "Association of advanced airway device with chest compression fraction during out-of-hospital cardiopulmonary arrest." *Resuscitation*, vol. 98, pp. 35–40, Jan. 2016.

[129] K. G. Gruben, J. Romlein, H. R. Halperin, and J. E. Tsitlik, "System for mechanical measurements during cardiopulmonary resuscitation in humans." *IEEE transactions on bio-medical engineering*, vol. 37, pp. 204–210, Feb. 1990.

[130] F. Gohier, K. Dellimore, and C. Scheffer, "Development of a smart backboard system for real-time feedback during cpr chest compression on a soft back support surface," in *Proc. 35th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 346–349.

[131] H. Myklebust and H. Fossan, "System for measuring and using parameters during chest compression for cardio-pulmonary resuscitation or a simulation thereof." *EP Patent 1057451A2*, Apr. 2009.

[132] A. Tomlinson, J. Nysaether, J. Kramer-Johansen, P. Steen, and E. Dorph, "Compression force–depth relationship during out-of-hospital cardiopulmonary resuscitation," *Resuscitation*, vol. 72, no. 3, pp. 364–370, 2007.

[133] F. Gohier, K. Dellimore, and C. Scheffer, "Development of a real-time feedback algorithm for chest compression during cpr without assuming full chest decompression," *Resuscitation*, vol. 85, no. 6, pp. 820–825, 2014.

[134] C. F. Babbs, A. E. Kemeny, W. Quan, and G. Freeman, "A new paradigm for human resuscitation research using intelligent devices," *Resuscitation*, vol. 77, no. 3, pp. 306–315, 2008.

[135] E. Fitzgibbon, R. Berger, J. Tsitlik, and H. R. Halperin, "Determination of the noise source in the electrocardiogram during cardiopulmonary resuscitation." *Critical care medicine*, vol. 30, pp. S148–S153, Apr. 2002.

[136] E. Aramendi, U. Irusta, U. Ayala, H. Naas, J. Kramer-Johansen, and T. Eftestøl, "Filtering mechanical chest compression artefacts from out-of-hospital cardiac arrest data." *Resuscitation*, vol. 98, pp. 41–47, Jan. 2016.

[137] H. Losert, M. Risdal, F. Sterz, J. Nysaether, K. Köhler, T. Eftestøl *et al.*, "Thoracic-impedance changes measured via defibrillator pads can monitor signs of circulation." *Resuscitation*, vol. 73, pp. 221–228, May 2007.

[138] L.-Y. Lin, M.-T. Lo, W.-C. Chiang, C. Lin, P. C.-I. Ko, K.-H. Hsiung *et al.*, "A new way to analyze resuscitation quality by reviewing automatic external defibrillator data." *Resuscitation*, vol. 83, pp. 171–176, Feb. 2012.

[139] S. Ruiz de Gauna, U. Irusta, J. Ruiz, U. Ayala, E. Aramendi, and T. Eftestøl, "Rhythm analysis during cardiopulmonary resuscitation: past, present, and future." *BioMed research international*, vol. 2014, p. 386010, 2014.

[140] E. Alonso, D. González-Otero, E. Aramendi, S. Ruiz de Gauna, J. Ruiz, U. Ayala *et al.*, "Can thoracic impedance monitor the depth of chest compressions during out-of-hospital cardiopulmonary resuscitation?" *Resuscitation*, vol. 85, pp. 637–643, May 2014.

[141] E. Aramendi, U. Ayala, U. Irusta, E. Alonso, T. Eftestøl, and J. Kramer-Johansen, "Suppression of the cardiopulmonary resuscitation artefacts using the instantaneous chest compression rate extracted from the thoracic impedance." *Resuscitation*, vol. 83, pp. 692–698, Jun. 2012.

[142] R. Padilla, S. L. Netto, and E. A. Da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 international conference on systems, signals and image processing (IWSSIP)*. IEEE, 2020, pp. 237–242.

[143] J. Coult, J. Blackwood, T. D. Rea, P. J. Kudenchuk, and H. Kwok, "A method to detect presence of chest compressions during resuscitation using transthoracic impedance," *IEEE journal of biomedical and health informatics*, vol. 24, no. 3, pp. 768–774, 2019.

[144] E. Rueda, E. Aramendi, U. Irusta, and A. H. Idris, "A method to detect pauses for ventilation during cardiopulmonary resuscitation using the thoracic impedance," in *Proc. Computing in Cardiology*, 2020, pp. 1–4.

[145] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[146] S. G. Beesems and R. W. Koster, "Accurate feedback of chest compression depth on a manikin on a soft surface with correction for total body displacement." *Resuscitation*, vol. 85, pp. 1439–1443, Nov. 2014.

[147] V. Gupta, R. H. Schmicker, P. Owens, A. E. Pierce, and A. H. Idris, "Software annotation of defibrillator files: Ready for prime time?" *Resuscitation*, vol. 160, pp. 7–13, Mar. 2021.

[148] P. Baskett, J. Nolan, and M. Parr, "Tidal volumes which are perceived to be adequate for resuscitation." *Resuscitation*, vol. 31, pp. 231–234, Jun. 1996.

[149] D. L. Grieco, L. J Brochard, A. Drouet, I. Telias, S. Delisle, G. Bronchti *et al.*, "Intrathoracic airway closure impacts co, javax.xml.bind.jaxbelement@7ca79c3, signal and delivered ventilation during cardiopulmonary resuscitation." *American journal of respiratory and critical care medicine*, vol. 199, pp. 728–737, Mar. 2019.

[150] S. Ruiz de Gauna, M. Leturiondo, J. J. Gutiérrez, J. M. Ruiz, D. M. González-Otero, J. K. Russell *et al.*, "Enhancement of capnogram waveform in the presence of chest compression artefact during cardiopulmonary resuscitation." *Resuscitation*, vol. 133, pp. 53–58, Dec. 2018.

[151] J. J. Gutiérrez, M. Leturiondo, S. R. de Gauna, J. M. Ruiz, L. A. Leturiondo, D. M. González-Otero *et al.*, "Enhancing ventilation detection during cardiopulmonary resuscitation by filtering chest compression artifact from the capnography waveform," *PLOS ONE*, vol. 13, no. 8, p. e0201565, aug 2018.

[152] A. Grenvik, S. Ballou, E. McGinley, J. E. Millen, W. L. Cooley, and P. Safar, "Impedance pneumography. comparison between chest impedance changes and respiratory volumines in 11 healthy volunteers." *Chest*, vol. 62, pp. 439–443, Oct. 1972.

[153] J. J. Freundlich and J. C. Erickson, "Electrical impedance pneumography for simple nonrestrictive continuous monitoring of respiratory rate, rhythm and tidal volume for surgical patients." *Chest*, vol. 65, pp. 181–184, Feb. 1974.

[154] L. E. Baker, L. A. Geddes, H. E. Hoff, and C. J. Chaput, "Physiological factors underlying transthoracic impedance variations in respiration." *Journal of applied physiology*, vol. 21, pp. 1491–1499, Sep. 1966.

[155] M. Valentinuzzi, L. Geddes, and L. Baker, "The law of impedance pneumography," *Medical and biological engineering*, vol. 9, pp. 157–163, 1971.

[156] A. V. Sahakian, W. J. Tompkins, and J. G. Webster, "Electrode motion artifacts in electrical impedance pneumography," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 6, pp. 448–451, 1985.

[157] S. Ansari, K. R. Ward, and K. Najarian, "Motion artifact suppression in impedance pneumography signal for portable monitoring of respiration: An adaptive approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 2, pp. 387–398, 2017.

[158] K. Roberts, V. Srinivasan, D. E. Niles, J. Eilevstjønn, L. Tyler, L. Boyle *et al.*, "Does change in thoracic impedance measured via defibrillator electrode pads accurately detect ventilation breaths in children?" *Resuscitation*, vol. 81, no. 11, pp. 1544–1549, 2010.

[159] J. R. Gould, L. Campana, D. Rabickow, R. Raymond, and R. Partridge, "Manual ventilation quality is improved with a real-time visual feedback system during simulated resuscitation." *International journal of emergency medicine*, vol. 13, p. 18, Apr. 2020.

[160] K. Charlton, G. McClelland, K. Millican, D. Haworth, P. Aitken-Fell, and M. Norton, "The impact of introducing real time feedback on ventilation rate and tidal volume by ambulance clinicians in the north east in cardiac arrest simulations." *Resuscitation plus*, vol. 6, p. 100130, Jun. 2021.

[161] U. I. *, J. Ruiz, S. R. de Gauna, T. EftestØl, and J. Kramer-Johansen, "A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 1052–1062,

2009.

[162] I. Isasi, U. Irusta, E. Aramendi, A. H. Idris, and L. Sörnmo, "Restoration of the electrocardiogram during mechanical cardiopulmonary resuscitation," *Physiological Measurement*, vol. 41, no. 10, p. 105006, nov 2020.

[163] S. O. Aase, T. Eftestøl, J. H. Husøy, K. Sunde, and P. A. Steen, "Cpr artifact removal from human ecg using optimal multichannel filtering." *IEEE transactions on bio-medical engineering*, vol. 47, pp. 1440–1449, Nov. 2000.

[164] S. Ruiz de Gauna, J. Ruiz, U. Irusta, E. Aramendi, T. Eftestøl, and J. Kramer-Johansen, "A method to remove cpr artefacts from human ecg using only the recorded ecg." *Resuscitation*, vol. 76, pp. 271–278, Feb. 2008.

[165] J. Ruiz, U. Irusta, S. Ruiz de Gauna, and T. Eftestøl, "Cardiopulmonary resuscitation artefact suppression using a kalman filter and the frequency of chest compressions as the reference signal." *Resuscitation*, vol. 81, pp. 1087–1094, Sep. 2010.

[166] J. H. Husoy, J. Eilevstjonn, T. Eftestol, S. O. Aase, H. Myklebust, and P. A. Steen, "Removal of cardiopulmonary resuscitation artifacts from human ECG using an efficient matching pursuit-like algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 11, pp. 1287–1298, 2002.

[167] A. Langhelle, T. Eftestøl, H. Myklebust, M. Eriksen, B. T. Holten, and P. A. Steen, "Reducing cpr artefacts in ventricular fibrillation in vitro." *Resuscitation*, vol. 48, pp. 279–291, Mar. 2001.

[168] J. Eilevstjønn, T. Eftestøl, S. O. Aase, H. Myklebust, J. H. Husøy, and P. A. Steen, "Feasibility of shock advice analysis during cpr through removal of cpr artefacts from the human ecg." *Resuscitation*, vol. 61, pp. 131–141, May 2004.

[169] E. Aramendi, S. R. de Gauna, U. Irusta, J. Ruiz, M. F. Arcocha, and J. M. Ormaetxe, "Detection of ventricular fibrillation in the presence of cardiopulmonary resuscitation artefacts." *Resuscitation*, vol. 72, pp. 115–123, Jan. 2007.

[170] A. Amann, A. Klotz, T. Niederklapfer, A. Kupferthaler, T. Werther, M. Granegger *et al.*, "Reduction of cpr artifacts in the ventricular fibrillation ecg by coherent line removal," *Biomedical engineering online*, vol. 9, no. 1, pp. 1–15, 2010.

[171] F. Alonso-Atienza, E. Morgado, L. Fernandez-Martinez, A. Garcia-Alberola, and J. L. Rojo-Alvarez, "Detection of life-threatening arrhythmias using feature selection and support vector machines," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 832–840, mar 2014.

[172] C. Figuera, U. Irusta, E. Morgado, E. Aramendi, U. Ayala, L. Wik *et al.*, "Machine learning techniques for the detection of shockable rhythms in automated external defibrillators," *PLOS ONE*, vol. 11, no. 7, p. e0159654, jul 2016.

[173] I. Isasi, U. Irusta, A. Elola, E. Aramendi, U. Ayala, E. Alonso *et al.*, "A machine learning shock decision algorithm for use during piston-driven chest compressions," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1752–1760, jun 2019.

[174] A. Neurauter, T. Eftestøl, J. Kramer-Johansen, B. S. Abella, K. Sunde, V. Wenzel *et al.*, "Prediction of countershock success using single features from multiple ventricular fibrillation frequency bands and feature combinations using neural networks," *Resuscitation*, vol. 73, no. 2, pp. 253–263, may 2007.

[175] J. Coult, J. Blackwood, L. Sherman, T. D. Rea, P. J. Kudenchuk, and H. Kwok, "Ventricular fibrillation waveform analysis during chest compressions to predict survival from cardiac arrest," *Circulation: Arrhythmia and Electrophysiology*, vol. 12, no. 1, jan 2019.

[176] M. Risdal, S. O. Aase, J. Kramer-Johansen, and T. Eftestol, "Automatic identification of return of spontaneous circulation during cardiopulmonary resuscitation," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 60–68, jan 2008.

[177] E. Alonso, E. Aramendi, M. Daya, U. Irusta, B. Chicote, J. K. Russell *et al.*, "Circulation detection using the electrocardiogram and the thoracic impedance acquired by defibrillation pads," *Resuscitation*, vol. 99, pp. 56–62, feb 2016.

[178] H. Kwok, J. Coult, M. Drton, T. D. Rea, and L. Sherman, "Adaptive rhythm sequencing: A method for dynamic rhythm classification during CPR," *Resuscitation*, vol. 91, pp. 26–31, jun 2015.

[179] A. B. Rad, T. Eftestøl, U. Irusta, J. T. Kvaløy, L. Wik, J. Kramer-Johansen *et al.*, "An automatic system for the comprehensive retrospective analysis of cardiac rhythms in resuscitation episodes," *Resuscitation*, vol. 122, pp. 6–12, jan 2018.

[180] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *The Annals of Statistics*, vol. 36, no. 3, jun 2008.

[181] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, sep 1995.

[182] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, aug 1996.

[183] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data." *Journal of bioinformatics and computational biology*, vol. 3, pp. 185–205, Apr. 2005.

[184] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, no. 2, jun 1994.

[185] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, jan 1996.

[186] I. Isasi, U. Irusta, E. Aramendi, T. Eftestøl, J. Kramer-Johansen, and L. Wik, "Rhythm analysis during cardiopulmonary resuscitation using convolutional neural networks," *Entropy*, vol. 22, no. 6, p. 595, may 2020.

[187] A. Elola, E. Aramendi, U. Irusta, A. Picón, E. Alonso, P. Owens *et al.*, "Deep neural networks for ECG-based pulse detection during out-of-hospital cardiac arrest," *Entropy*, vol. 21, no. 3, p. 305, mar 2019.

[188] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[189] D. P. Davis, L. A. Garberson, D. L. Andrusiek, D. Hostler, M. Daya, R. Pirrallo *et al.*, "A descriptive analysis of emergency medical service systems participating in the resuscitation outcomes consortium (ROC) network," *Prehospital Emergency Care*, vol. 11, no. 4, pp. 369–382, jan 2007.

[190] C. D. Newgard, G. K. Sears, T. D. Rea, D. P. Davis, R. G. Pirrallo, C. W. Callaway *et al.*, "The resuscitation outcomes consortium epistry-trauma: Design, development, and implementation of a north american epidemiologic prehospital trauma registry," *Resuscitation*, vol. 78, no. 2, pp. 170–178, aug 2008.

[191] P. Morley, "Steady as a ROC: The resuscitation outcomes consortium," *Resuscitation*, vol. 78, no. 2, pp. 105–106, aug 2008.

[192] I. G. Stiell, G. Nichol, B. G. Leroux, T. D. Rea, J. P. Ornato, J. Powell *et al.*, "Early versus later rhythm analysis in patients with out-of-hospital cardiac arrest," *New England Journal of Medicine*, vol. 365, no. 9, pp. 787–797, sep 2011.

[193] T. P. Aufderheide, G. Nichol, T. D. Rea, S. P. Brown, B. G. Leroux, P. E. Pepe *et al.*, "A trial of an impedance threshold device in out-of-hospital cardiac arrest," *New England Journal of Medicine*, vol. 365, no. 9, pp. 798–806, sep 2011.

[194] P. J. Kudenchuk, S. P. Brown, M. Daya, T. Rea, G. Nichol, L. J. Morrison *et al.*, "Amiodarone, lidocaine, or placebo in out-of-hospital cardiac arrest," *New England Journal of Medicine*, vol. 374, no. 18, pp. 1711–1722, may 2016.

[195] D. Zive, R. Schmicker, M. Daya, P. Kudenchuk, G. Nichol, J. Rittenberger *et al.*, "Survival and variability over time from out of hospital cardiac arrest across large geographically diverse communities participating in the resuscitation outcomes consortium," *Resuscitation*, vol. 131, pp. 74–82, oct 2018.

[196] H. E. Wang, D. K. Prince, S. W. Stephens, H. Herren, M. Daya, N. Richmond *et al.*, "Design and implementation of the resuscitation outcomes consortium pragmatic airway resuscitation trial (part)," *Resuscitation*, vol. 101, pp. 57–64, 2016.

[197] X. Jaureguibeitia, U. Irusta, E. Aramendi, P. C. Owens, H. E. Wang, and A. H. Idris, "Automatic detection of ventilations during mechanical cardiopulmonary resuscitation," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2580–2588, 2020.

[198] X. Jaureguibeitia, E. Aramendi, H. E. Wang, and A. H. Idris, "Impedance-based ventilation detection and signal quality control during out-of-hospital cardiopulmonary resuscitation," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 6, pp. 3026–3036, 2023.

[199] X. Jaureguibeitia, E. Aramendi, U. Irusta, E. Alonso, T. P. Aufderheide, R. H. Schmicker *et al.*, "Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets." *Resuscitation*, vol. 168, pp. 44–51, Nov. 2021.

[200] H. E. Wang, X. Jaureguibeitia, E. Aramendi, J. L. Jarvis, J. N. Carlson, U. Irusta *et al.*, "Airway strategy and chest compression quality in the pragmatic airway resuscitation trial." *Resuscitation*, vol. 162, pp. 93–98, May 2021.

[201] H. E. Wang, X. Jaureguibeitia, E. Aramendi, G. Nichol, T. Aufderheide, M. R. Daya *et al.*, "Airway strategy and ventilation rates in the pragmatic airway resuscitation trial." *Resuscitation*, vol. 176, pp. 80–87, Jul. 2022.

[202] M. M. J. Nassal, X. Jaureguibeitia, E. Aramendi, U. Irusta, A. R. Panchal, H. E. Wang *et al.*, "Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the pragmatic airway resuscitation trial." *Resuscitation*, vol. 168, pp. 58–64, Nov. 2021.

[203] X. Jaureguibeitia, U. Irusta, E. Aramendi, E. Alonso, P. Owens, H. Wang *et al.*, "Impedance based automatic detection of ventilations during mechanical cardiopulmonary resuscitation," in *Proc. 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 19–23.

[204] X. Jaureguibeitia, U. Irusta, E. Aramendi, H. Wang, and A. Idris, "An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation," in *Proc. Computing in Cardiology*, 2020, pp. 1–4.

[205] X. Jaureguibeitia, E. Aramendi, H. E. Wang, and A. H. Idris, "Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar," in *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (ISBN: 978-84-09-45972-8)*, 2022, pp. 175–178.

[206] X. Jaureguibeitia, U. Irusta, E. Aramendi, P. Owens, H. E. Wang, and A. H. Idris, "Abstract 264: Automatic detection of ventilations using the thoracic impedance signal during lucas chest compressions," *Circulation*, vol. 140, no. Suppl_2, nov 2019.

[207] X. Jaureguibeitia, E. ARAMENDI, A. H. Idris, and H. Wang, "Abstract 164: Efficacy of thoracic impedance for ventilation detection during continuous chest compressions," *Circulation*, vol. 146, no. Suppl_1, nov 2022.

[208] X. Jaureguibeitia, U. Irusta, E. Aramendi, H. E. Wang, and A. H. Idris, "Algoritmo multietapa para la detección de ventilaciones en la impedancia torácica durante la resucitación cardiopulmonar," in *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (ISBN: 978-84-09-25491-0)*, 2020, pp. 352–355.

[209] Z. Zhao, S. Särkkä, and A. B. Rad, "Kalman-based spectro-temporal ECG analysis using deep convolutional networks for atrial fibrillation detection," *Journal of Signal Processing Systems*, vol. 92, no. 7, pp.

621–636, apr 2020.

[210] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches," in *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*. Association for Computational Linguistics, 2014.

[211] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, oct 2016.

[212] A. Rozo, J. Moeyersons, J. Morales, R. G. van der Westen, L. Lijnen, C. Smeets *et al.*, "Data augmentation and transfer learning for data quality assessment in respiratory monitoring," *Frontiers in Bioengineering and Biotechnology*, vol. 10, feb 2022.

[213] M. Nardelli, A. Greco, J. Bolea, G. Valenza, E. P. Scilingo, and R. Bailon, "Reliability of lagged poincaré plot parameters in ultrashort heart rate variability series: Application on affective sounds," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 741–749, may 2018.

[214] X. Jaureguibeitia, E. Aramendi, U. Irusta, A. H. Idris, and H. E. Wang, "Abstract 203: Thoracic impedance reflects differences between endotracheal and laryngeal advanced airway during mechanical chest compressions," *Circulation*, vol. 142, no. Suppl_4, nov 2020.

[215] H. E. Wang, X. Jaureguibeitia, A. H. Idris, U. Iruste, E. Gonzalez, T. P. Aufderheide *et al.*, "Abstract 206: Effect of airway strategy upon chest compression quality in the pragmatic airway resuscitation trial," *Circulation*, vol. 142, no. Suppl_4, nov 2020.

[216] H. Wang, X. Jaurguibeitia, A. Idris, U. Irusta, E. Alonso, T. Aufderheide *et al.*, "Effect of airway strategy upon chest compression quality in the pragmatic airway resuscitation trial," in *Prehospital Emergency Care, Abstracts for the 2021 NAEMSP Scientific Assembly*, vol. 25, 2021, p. 125.

[217] M. Nassal, X. Jaureguibeitia, E. Aramendi, U. Irusta, A. R. Panchal, H. Wang *et al.*, "Abstract 13110: Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the pragmatic airway resuscitation trial," *Circulation*, vol. 144, no. Suppl_2, nov 2021.

[218] H. E. Wang, X. Jaureguibeitia, E. Aramendi, M. Nassal, A. Panchal, G. Nichol *et al.*, "Effect of airway strategy upon ventilation rates in

the pragmatic airway resuscitation trial," in *Prehospital Emergency Care, Research Abstracts for the 2022 NAEMSP Annual Meeting*, vol. 26, 2022, p. 108.

[219] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa *et al.*, "Attention u-net: Learning where to look for the pancreas," 2018.

[220] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang *et al.*, "Transunet: Transformers make strong encoders for medical image segmentation," 2021.

[221] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, P. J. Watkinson *et al.*, "An impedance pneumography signal quality index: Design, assessment and application to respiratory rate monitoring," *Biomedical Signal Processing and Control*, vol. 65, p. 102339, mar 2021.

[222] E. Alonso, U. Irusta, E. Aramendi, and M. R. Daya, "A machine learning framework for pulse detection during out-of-hospital cardiac arrest," *IEEE Access*, vol. 8, pp. 161 031–161 041, 2020.

A PUBLISHED STUDIES

A.1 Publications associated to objective 1

A.1.1 First journal paper

	Publication in international journal
Reference	Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Pamela C. Owens, Henry E. Wang, Ahamed H. Idris, "Automatic detection of ventilations during mechanical cardiopulmonary resuscitation", *IEEE Journal of Biomedical and Health Informatics 2020*, vol. 24, no. 9, pp. 2580-2588.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (4/30) based on Web of Science Rank 2020 • **Impact factor**: 5.773

Automatic Detection of Ventilations During Mechanical Cardiopulmonary Resuscitation

Xabier Jaureguibeitia, Unai Irusta, *Member, IEEE*, Elisabete Aramendi, *Member, IEEE*, Pamela C. Owens, Henry E. Wang, and Ahamed H. Idris

***Abstract*—Feedback on chest compressions and ventilations during cardiopulmonary resuscitation (CPR) is important to improve survival from out-of-hospital cardiac arrest (OHCA). The thoracic impedance signal acquired by monitor-defibrillators during treatment can be used to provide feedback on ventilations, but chest compression components prevent accurate detection of ventilations. This study introduces the first method for accurate ventilation detection using the impedance while chest compressions are concurrently delivered by a mechanical CPR device. A total of 423 OHCA patients treated with mechanical CPR were included, 761 analysis intervals were selected which in total comprised 5 884 minutes and contained 34 864 ventilations. Ground truth ventilations were determined using the expired CO_2 channel. The method uses adaptive signal processing to obtain the impedance ventilation waveform. Then, 14 features were calculated from the ventilation waveform and fed to a random forest (RF) classifier to discriminate false positive detections from actual ventilations. The RF feature importance was used to determine the best feature subset for the classifier. The method was trained and tested using stratified 10-fold cross validation (CV) partitions. The training/test process was repeated 20 times to statistically characterize the results. The best ventilation detector had a median (interdecile range, IDR) F_1-score of 96.32 (96.26–96.37). When used to provide feedback in 1-min intervals, the median (IDR) error and relative error in ventilation rate were 0.002 (−0.334–0.572) min^{-1} and 0.05 (−3.71–9.08)%, respectively. An accurate ventilation detector during mechanical CPR was demonstrated. The algorithm could be introduced in current equipment for feedback on ventilation rate and quality, and it could contribute to improve OHCA survival rates.**

***Index Terms*—Cardiopulmonary resucitation (CPR), ventilation, mechanical CPR, thoracic impedance, adaptive filter, random forest.**

I. Introduction

OUT of hospital cardiac arrest (OHCA) is an important public health problem. The annual incidence of treated OHCA in industrialized countries is between 35 and 60 cases per 100 000 persons, with survival rates below 10 % [1], [2]. High-quality cardiopulmonary resuscitation (CPR) maintains an artificial flow of oxygenated blood by means of chest compressions and ventilations, and is essential to improve OHCA survival [3]. Advanced life support resuscitation guidelines recommend both uninterrupted and high quality chest compressions, and concurrent ventilations with rates of approximately 10 breaths per minute (min^{-1}) after patient intubation [4]. Hyperventilation should be avoided because it increases intrathoracic pressure and may result in degraded hemodynamics [5]. However, hyperventilation during CPR is frequent with ventilation rates far exceeding the recommended values [6]–[8].

Chest compression detection systems are available on portable cardiac monitors. These technologies use accelerometers or changes in thoracic impedance to detect chest compressions [9]. During treatment, CPR feedback devices may help to improve rescuer compliance with treatment guidelines [10], [11]. After treatment, episode debriefing based on these recorded data may allow for retrospective performance assessment and quality improvement programs [12]. However, similar technologies for feedback on ventilation based on the impedance are not currently commercially available. The capnogram is the continuous measure of the partial pressure of expired CO_2 in respiratory gases. Capnography is the standard method for detecting ventilations during CPR [13], [14], but this signal is not available until the placement of an advanced airway. Thoracic impedance, which is recorded by most defibrillators to check pad placement and to adjust defibrillation energy, also provides detailed information on CPR activity [15]. Thoracic impedance varies with air volume changes in the lungs, and can therefore be used to identify ventilations. Even when capnography is available, impedance-based detection of ventilations is important, either to improve the accuracy of capnography-based ventilation detection algorithms [13], or to provide indirect evidence of tidal volume and peak positive ventilation pressures that cannot be inferred from the capnogram [16], [17].

Manuscript received October 18, 2019; revised December 13, 2019; accepted January 13, 2020. Date of publication January 17, 2020; date of current version September 3, 2020. This work was supported in part by the Spanish Ministerio de Ciencia, Innovación y Universidades under Grant RTI2018-101475-BI00, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), in part by the Basque Government under Grants IT-1229-19 and PRE_2019_1_0209, and in part by the NIH under Grant HL 077887. *(Corresponding author: Xabier Jaureguibeitia.)*

X. Jaureguibeitia, U. Irusta, and E. Aramendi are with the Department of Communications Engineering, University of the Basque Country UPV/EHU, 48013 Bilbao, Spain (e-mail: xabier.jaureguibeitia@ehu.eus; unai.irusta@ehu.eus; elisabete.aramendi@ehu.eus).

P. C. Owens and A. H. Idris are with the University of Texas Southwestern Medical Center, Dallas, TX 75390 USA (e-mail: pamela.owens@UTSouthwestern.edu; ahamed.idris@utsouthwestern.edu).

H. E. Wang is with the University of Texas Health Science Center, Houston, TX 77030 USA (e-mail: Henry.E.Wang@uth.tmc.edu).

This article has supplementary downloadable material available at https://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/JBHI.2020.2967643

Impedance ventilation waveforms are varied in shape, and the impedance signal is very sensitive to motion artifacts [18], [19]. Furthermore, during CPR the impedance presents a chest compression component with a spectrum that may overlap that of the ventilation waveforms [20]. Before intubation, if CPR is delivered in sequences of 30 compressions followed by 2 ventilations (standard 30:2 CPR), ventilations can be reliably identified during compression pauses [16]. After intubation, chest compressions and ventilations are given concurrently. Several ventilation detection algorithms have been proposed for this scenario [13], [20], [21], frequently using the adaptive filtering techniques originally conceived to remove chest compression artifacts from the electrocardiogram [22], [23]. Since manual chest compressions are variable in rate and depth, these adaptive filters often need accelerometer data to model the chest compression artifact [20].

Mechanical chest compression devices ensure high-quality chest compressions and have become popular in OHCA treatment. To date, there is no conclusive evidence of improved survival with mechanical CPR [24], [25], but the use of these devices has become widespread in scenarios like patient transport, invasive procedures, or prolonged resuscitation [26]. Mechanical chest compressions are stable in rate and depth, so there is no need for accelerometer data to model compressions [27], [28]. However, removing the mechanical chest compression component from the impedance is challenging because it has larger amplitudes and more spectral components than those observed during manual CPR [29].

The goal of this study was to determine whether an impedance-based algorithm can accurately detect ventilations during concurrent mechanical chest compressions. For this purpose, we implemented an adaptive filter to obtain the ventilation waveform from the raw impedance signal, designed features to characterize the impedance ventilation waveform, constructed an optimal model to identify true ventilations using a Random Forest (RF) with the best feature subset, and evaluated the performance of the model to detect ventilations and measure ventilation rate. A preliminary version of this work has been reported [30].

II. Materials

The study dataset was part of a large OHCA data repository collected by the Dallas-Fort Worth Center for Resuscitation Research, as part of the Resuscitation Outcomes Consortium [31]. A cohort of 567 patients treated between October 2012 and March 2016 were initially considered, those that contained concurrent impedance and capnography recordings as well as confirmed mechanical CPR according to the OHCA epistry data. Signals were acquired with the MRx monitor-defibrillator (Philips Medical Systems, Andover, MA, USA). The MRx measures impedance by applying a 32 kHz alternating current through the defibrillation pads and measuring the resulting voltage. The impedance signal was digitized with a sampling rate of 200 Hz and an amplitude resolution of 0.74 mΩ per least significant bit. The capnogram was acquired using Microstream (sidestream) technology, and the signal was sampled at 40 Hz with 0.004 mmHg resolution. Finally, mechanical chest compresions were given using the LUCAS-2 chest compression device (Physio-Control Inc/Jolife AB, Lund, Sweden), that delivers piston-driven compressions at a fixed rate of 100 min^{-1} and predefined depth between 1.5–2 inches. All signals from the MRx device were converted to an open format using custom Matlab (MathWorks Inc., Natick, MA) tools.

Signal intervals with confirmed LUCAS-2 use were extracted from the initial 567 patients. Chest compressions were automatically detected in the impedance signal using the algorithm proposed by Ayala *et al.* [32], and LUCAS-2 use was identified when the chest compression rate was fixed at 100 min^{-1} with small variability (see Fig. 1(b)). The inclusion criteria for the intervals was: minimum duration of 100 s with mechanical CPR, interpretable impedance and capnography signals, and no pauses in chest compressions longer than 20 s. In the dataset there were 8 917 min of confirmed mechanical CPR use with concurrent chest compressions, from which 5 884 min were used. Two were the main reasons to exclude 3 033 min. First, the lack of a proper gold standard to annotate ventilations because capnography was either unavailable (2 642 min) or strongly artefacted and thus uninterpretable (177 min). Second, low quality impedance or disconnections of the impedance channel (391 min). The latter give an estimate of how often impedance was unusable for ventilation detection during mechanical CPR (4.4% of the available minutes). So finally, 761 analysis intervals from 423 patients were included in the study, with a median (interquartile range, IQR) time of mechanical CPR per patient of 13 (8–19) minutes, and a median (IQR) duration of the analysis intervals of 5.4 (3.2–10.6) minutes. The median (IQR) proportion of time with concurrent compressions per patient was 98.6 (96.9–100)% in our data, so most of the time ventilations were provided concurrently with mechanical CPR.

The capnogram was used to annotate ground truth ventilations. First, the delay between the impedance and the capnogram caused by gas transport in the sampling tube (sidestream) was visually assessed and corrected. The delay in the capnogram line was different for each patient, with a median (IQR) value of 3.3 (3.1–3.5) s. Then, for each ventilation the insufflation (downfall) and expiration (uprise) onsets were automatically detected in the capnogram using the algorithm introduced by Aramendi *et al.* [14] (see Fig. 1(c)), and then manually inspected and revised. The revised annotations were considered the ground truth ventilations. The time interval between the onsets of inspiration and expiration marked the window for which ventilation detections in the impedance were considered correct (see Fig. 1(d)). As shown in the figure, the window for correct detections was prolonged by 1-sec after expiration onset to properly count those cases in which the impedance peak occurred shortly after expiration had started. In total, 34 864 true ventilations were annotated in the capnogram, with a median (IQR) of 72 (43–108) ventilations per patient.

III. Methods

The ventilation detection method is composed of the three stages shown in Fig. 2. First, the raw impedance signal is filtered

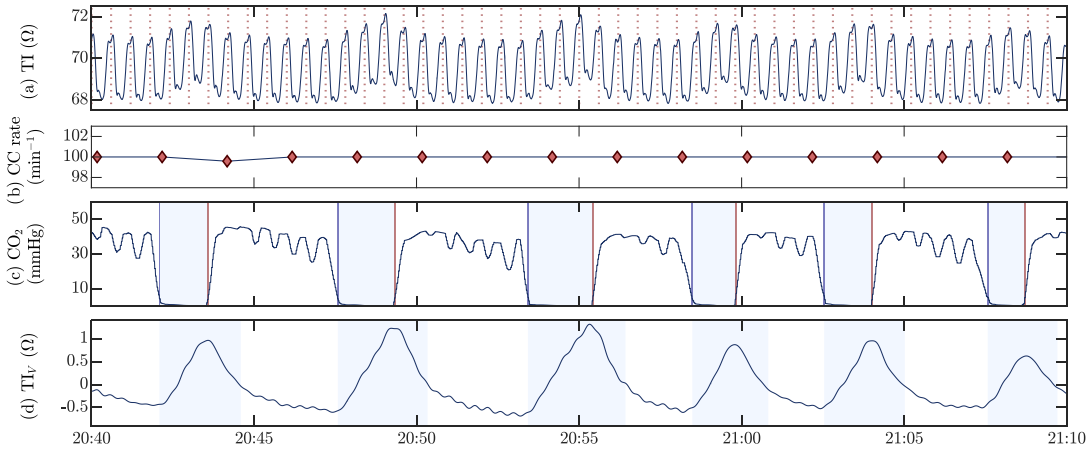


Fig. 1. An example of the signals in the study dataset with: (a) the raw thoracic impedance with chest compressions indicated by vertical dashed lines, (b) the compression rate computed every 2 s, (c) the capnogram to annotate the ground truth ventilations, and (d) the ventilation induced changes in the impedance obtained through signal processing from the raw impedance in (a) and used to detect ventilations. The compression rate in (b) was used to confirm the use of the LUCAS-2 device, and the shaded intervals in the capnogram (c) correspond to the true insufflation intervals. The shaded intervals in (d) are those in which a detected ventilation was considered a true positive detection, and correspond to the insufflation intervals extended by one second.

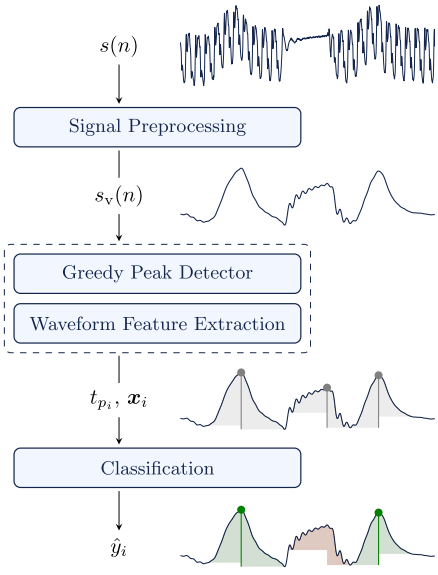


Fig. 2. Block diagram of the ventilation detection algorithm. The impedance signal $s(n)$ is filtered to obtain its ventilation component $s_v(n)$. Then a greedy peak detector detects the instants of the potential ventilations (t_{p_i}), and a waveform feature vector $\boldsymbol{x}_i$ is computed. The final classifier discriminates true ventilations (green) from false positive peak detections (red) using the waveform features.

to obtain the ventilation waveform component. Then, impedance fluctuations are detected and their peak times (t_{p_i}) identified using a greedy peak detector. The start and end of the fluctuation are calculated and its waveform is characterized by a vector of features $\boldsymbol{x}_i$. The greedy detector is designed to detect all candidate ventilations, with the tradeoff of producing many false positive detections. So the final stage is a machine learning classifier based on the waveform features to discriminate true ventilations (green) from false positives (red).

A. Signal Preprocessing

The raw impedance signal was first downsampled to $f_s = 50$ Hz to ease the design of the filters and reduce the computational load. In what follows n is the sample index so time is $t = n \cdot T_s$, where $T_s = 20$ ms is the sampling period. A high-pass filter with 0.05 Hz cut-off frequency was used to remove the DC component, and a low pass filter with a 2.5 Hz cut-off was used to remove high frequency residuals, including the high frequency components caused by chest compressions. Both filters were designed as 4-tap Butterworth filters, and zero-phase filtering was deployed. Finally, the most critical element was a Least Mean Squares (LMS) filter used to remove chest compression components.

The LMS algorithm was used in the classical configuration to cancel harmonic interferences [33], but adapted to mechanical chest compression components following the model introduced by Isasi *et al.* [28]. The block diagram of the filter is shown in Fig. 3, which follows the notation used in this subsection.The assumption is that after high pass and low pass filtering, the impedance signal contains two additive components:

$$s(n) \approx s_v(n) + s_c(n) \tag{1}$$

where $s_v(n)$ and $s_c(n)$ are the ventilation and compression components, respectively. These components are uncorrelated since they represent two independent treatments, compressions by the mechanical device and ventilations by the rescuer. The

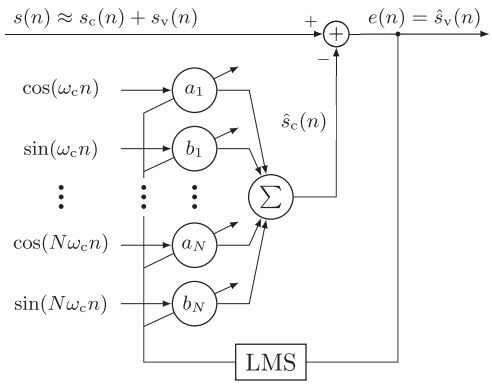


Fig. 3. Block diagram of the LMS filter used to remove the mechanical chest compression component from the impedance.

compression component is modeled as a quasi-periodic signal using a truncated Fourier-series representation with fundamental frequency $f_c = 1.667$ Hz (100 min^{-1}), and time-varying amplitudes $a_k(n)$ and $b_k(n)$ to adapt to changes in the impedance signal:

$$s_c(n) = \sum_{k=1}^{N} [a_k(n)\cos(k\omega_c n) + b_k(n)\sin(k\omega_c n)] \quad (2)$$

where $\omega_c = 2\pi f_c / f_s = 0.209$ is the discrete angular frequency of the LUCAS-2 chest compressions, and N is the number of harmonics in the model. In matrix notation the chest compression component is $s_c(n) = \mathbf{x}^\intercal(n)\mathbf{w}(n)$, where:

$$\mathbf{x}(n) = [\cos(\omega_c n), \sin(\omega_c n) \ldots \cos(N\omega_c n), \sin(N\omega_c n)]^\intercal, \quad (3)$$

$$\mathbf{w}(n) = [a_1(n), b_1(n) \ldots a_N(n), b_N(n)]^\intercal \quad (4)$$

are the reference signal (harmonics) and coefficient vectors, respectively. The LMS algorithm computes the $\mathbf{w}(n)$ coefficients to minimize the mean squared error $E\{|e(n)|^2\}$ between the desired signal $d(n) = s(n)$ and the estimated chest compression component $\hat{s}_c(n)$:

$$e(n) = s(n) - \mathbf{x}^\intercal(n)\mathbf{w}(n), \quad (5)$$

so the error signal is then the estimated ventilation component, $e(n) = s_v(n)$. The error is minimized using the steepest descent algorithm, and the gradient at time n of the squared error is:

$$\nabla_{\mathbf{w}} e^2 = \frac{\partial e^2}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}}(s - \mathbf{x}^\intercal \mathbf{w})^2 = -2e\mathbf{x}. \quad (6)$$

The filter coefficients are updated in the opposite direction, following:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{x}(n) \quad (7)$$

where the step-size parameter μ determines the adaptation speed and tracking capabilities of the filter. The values for the LMS filter were set to $\mu = 0.15$ and $N = 3$ after some preliminary tests.

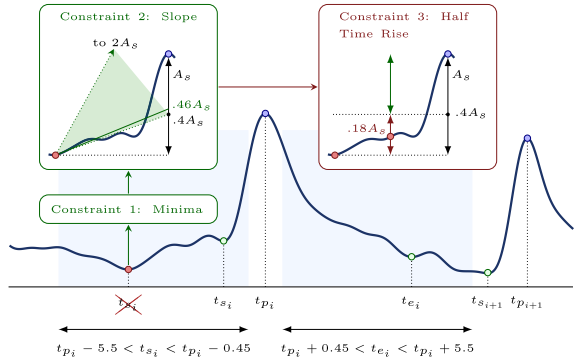


Fig. 4. Peak detection algorithm. The shaded intervals indicate the search intervals for the start/end of ventilations, and the constraints on how to determine these points. In the example for ventilation i the first global minimum in the search interval was discarded because it did not meet constraint 3.

B. Greedy Peak Detector

A greedy peak detector was designed to detect local maxima in the impedance ventilation component, $s_v(n)$. For each detected local maximum i three fiducial time points were calculated: the start of the ventilation (t_{s_i}, insufflation onset), the peak time (t_{p_i}, end of insufflation), and the end of the ventilation (t_{e_i}, end of expiration). As shown in Fig. 4 (shaded intervals) an interval of approximately 5 s was defined before and after each local maximum to search for t_{s_i} and t_{e_i}:

$$t_{p_i} - 5.5 < t_{s_i} < t_{p_i} - 0.45 \quad (8)$$

$$t_{p_i} + 0.45 < t_{e_i} < t_{p_i} + 5.5 \quad (9)$$

These thresholds were obtained after some preliminary tests, but are sensible values considering how ventilations should be provided. During CPR, ventilation breaths should be delivered over 1 s (insufflation) [4], so a minimum of 0.45 s is a conservative threshold to capture even quick ventilation events. Recommended ventilation rates are 10 min^{-1} [4], or about 6 s per ventilation, so the 5.5 s threshold for insufflation/exhalation includes even very slow ventilations. Finally, the minimum separation between detections was fixed at $\Delta T_m = 1.5$ s, which is sufficient for hyperventilation rates of up to 40 min^{-1}.

Three constraints were imposed to find t_{s_i} and t_{e_i} in the intervals defined in eqs (8) and (9). The constraints are graphically illustrated in the example in Fig. 4, and were applied in order to all potential timepoints in the search interval for the fiducial points. For the start of ventilation the constraints were:

1) t_{s_i} must correspond to the smallest impedance value in the interval (t_{s_i}, t_{p_i}).
2) The mean slope of the impedance in the interval (t_{s_i}, $t_{s_i} + 0.2$ s), m_s, and the total rise in impedance amplitude from t_{s_i} to t_{p_i}, A_s, were computed. The projection of that slope to the peak position had to be in the following range:

$$0.4 \cdot A_s \leq m_s \cdot (t_{p_i} - t_{s_i}) \leq 2 \cdot A_s \quad (10)$$

This is a starting slope constraint relative to ventilation amplitude. Low projection values are usually the result of

a slow baseline recovery towards zero, while high ones are caused by compression component residuals and signal distortions.

3) The rise in amplitude between t_{s_i} and the mid point to t_{p_i} (half-time rise) should be at least $0.4 \cdot A_s$, to prevent selecting t_{s_i} at some point far from the actual ventilation.

A similar procedure was followed for t_{e_i}, but with these conditions on constraint 2:

$$0.3 \cdot A_s \leq m_s \cdot (t_{e_i} - t_{p_i}) \leq 1.5 \cdot A_s \tag{11}$$

and a half-time fall of $0.2 \cdot A_s$ for constraint 3. The values are smaller for expiration because ventilation waveforms in the impedance tend to be concave for insufflation and convex for expiration. Finally, all peaks for which t_{s_i} or t_{e_i} could not be found were discarded, and for all the detected peaks the condition that two consecutive ventilations did not overlap was imposed ($t_{e_i} < t_{s_{i+1}}$).

C. Waveform Feature Extraction

Fourteen features were extracted for each detected peak to characterize the ventilation fluctuation waveform. The first four features were the duration of insufflation and expiration, T_{I_i} and T_{E_i}, respectively; and the amplitude change in impedance for the insufflation and expiration intervals, A_{I_i} and A_{E_i}, respectively. Ten waveform moments were also computed, five for each ventilation phase. Let us denote by $\boldsymbol{p}_{I_i}$ and $\boldsymbol{p}_{E_i}$ the vectors with the samples of either the insufflation or expiration phase of ventilation i, but normalized so that the total sum of the samples is unity. So for the insufflation phase we have $\boldsymbol{p}_{I_i}$ of length L_{I_i} and for the expiration phase $\boldsymbol{p}_{E_i}$ of length L_{E_i}. Let us also denote by $\boldsymbol{z}_i = [0, 1, 2, \ldots, L-1]/L$ a vector of length $L = L_{I_i}$ or $L = L_{E_i}$ depending on the case, with equispaced values between 0 and 1. Then $\boldsymbol{p}_{I_i}$ or $\boldsymbol{p}_{E_i}$ can be regarded as a probability density functions in the [0, 1] support interval, and we could compute their moments of order ℓ as:

$$\mu_{I_\ell} = \sum_{n=0}^{L_{I_i}-1} p_{I_i}(n) \cdot (n/L_{I_i})^\ell \quad \text{for} \quad \ell = 1, \ldots, 5 \tag{12}$$

$$\mu_{E_\ell} = \sum_{n=0}^{L_{E_i}-1} p_{E_i}(n) \cdot (n/L_{E_i})^\ell \quad \text{for} \quad \ell = 1, \ldots, 5 \tag{13}$$

The features $\mu_{I_1}, \ldots, \mu_{I_5}$ and $\mu_{E_1}, \ldots, \mu_{E_5}$, form the 10 waveform features used to parametrize the waveform during the insufflation and expiration phases, respectively.

D. Peak Classification

Potential ventilations from the greedy detector were compared to ground truth annotations, and labeled as true positives ($y_i = 1$, actual ventilation) or false positives ($y_i = 0$, no ventilation). When more than one peak detection fell in the interval for true positive detections (Fig. 1(d)), the one with t_{p_i} closest to the expiration onset in the capnogram was regarded as true positive, and the rest as false positives. After peak detection and feature extraction data was formatted as a set of instance-label pairs $\{(\boldsymbol{x}_i, y_i)\}_{i=1,\cdots N_p}$, where y_i are the true/false ventilation labels for the detected peaks, $\boldsymbol{x}_i \in \mathbb{R}^M$ contains the M features for peak t_{p_i}, and N_p is the number of detected peaks. The last step was to develop a Random Forest classifier to discriminate true from false ventilation detections.

A RF is an ensemble of nearly uncorrelated decision trees. Decision trees present some desirable characteristics like independence from the underlying data distribution, robustness to outliers, and protection from correlated and/or bad predictors. However, individual trees are poor classifiers, deep trees are prone to overfitting and shallow trees to underfitting. Aggregating the decisions of B uncorrelated decision trees boosts classifier performance [34]. To uncorrelate the $T_b(\boldsymbol{x})$ ($b = 1, \ldots B$) trees, these are trained with N_b bootstrap samples of the training data of size $N_b < N_p$, formed by randomly sampling the data with replacement. In addition, the RF algorithm randomizes the feature space by randomly selecting a subset of M_b features at each tree split ($M_b < M$). The final decision of the B trees for the sample $\boldsymbol{x}_i$ is obtained as the majority vote of the $\hat{y}_{i,b} = T_b(\boldsymbol{x}_i)$ for $b = 1, \ldots B$. We chose an in-bag fraction N_b/N_p of 0.5, the number of trees was fixed to $B = 100$, and the number of predictors per split to the default value of $M_b = \sqrt{M}$. Preliminary tests indicated that the choice of these RF parameters was not critical.

Data were partitioned using a 10-fold CV strategy to train and validate the classifier. At each iteration 9 folds were used as training data and the remaining fold as test data. The folds were partitioned patient wise and in a balanced way, so that each fold contained approximately 10% of the ground truth ventilation annotations [35]. All the calculations were weighted patient-wise to avoid biasing the results towards the patients with more ventilations in the dataset. Since the results may depend on the 10-fold CV partition used to train and validate the classifier, using a single 10-fold CV partition may overestimate or underestimate the accuracy of our method. To avoid biasing the results, the process was repeated 20 times with different random 10-fold CV partitions. And the accuracy metrics were statistically characterized using the 20 values obtained for each partition.

One of the salient characteristics of RF classifiers is a built-in feature ranking called feature importance. Importance was measured using the permuted out-of-bag (OOB) error. For a given tree, the subset of the training data left out in the bootstrap sample (out-of-bag samples) is used to evaluate the model's predictions. Then the values for that feature in the OOB sample instances are randomly shuffled, and the decrease in prediction accuracy is measured. The decrease is larger for more important features. The process is repeated for all trees and features, resulting in a ranking of the features from the most important to the least. Recursive feature elimination (RFE) based on feature importance was used to reduce the number of features [36], [37]. Starting from a full feature model ($M = 14$), at each step the classifier was trained and the individual importance of each feature was computed. Then, the least important feature was removed and the process repeated until a model with a single feature was obtained. In this way we had 14 different models, from $M = 14$ to $M = 1$. The feature elimination process was carried out 200 times, once per test fold.

E. Evaluation of the Detector

The performance of the ventilation detector was evaluated in terms of sensitivity (Se), positive predictive value (PPV) and F_1-score (F_1), computed as:

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}},\ \mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}},\ \mathrm{F}_1 = 2\frac{\mathrm{Se}\cdot\mathrm{PPV}}{\mathrm{Se}+\mathrm{PPV}} \quad (14)$$

where TP, FP and FN are the true positive, false positive and false negative detections, respectively. There is a large imbalance in the number of ventilations that each patient contributed to the dataset, which is associated to how much time the mechanical compressor was used on each patient. In order make the method applicable to as many patients as possible, we weighted the contribution of each patient equally. This was done by calculating the metrics in equations (14) individually for each patient, and then averaging those values for the final Se, PPV and F_1 for the complete set.

IV. Results

A. Classification Performance

Ventilation detection was evaluated in two stages, first for the greedy detector and then after adding the classifier. The greedy detector outputted 55 908 detections, from which 34 615 were actual ventilations and 21 223 were false positives. The greedy peak detector missed 249 ventilations (0.71% of the total amount), which were regarded as false negative detections for the complete algorithm in the subsequent performance evaluations. The patient-weighted Se, PPV and F_1 for the greedy detector were 99.36%, 62.04% and 76.27%, respectively.

The classifier corrected the false positive detections. The best compromise for simplicity and performance was obtained for a classifier with $M = 6$ features. After adding the classification block, the median (interdecile range, IDR) value of Se, PPV and F_1 for the complete algorithm were 96.26 (96.15–96.31)%, 96.37 (96.32–96.43)% and 96.32 (96.25–96.36)%, respectively. The effect of the number of features, M, in the performance of the algorithm is presented in Fig. 5. The figure shows that performance was very stable for $M \geq 6$.

The selection probability for each feature was estimated as the proportion of times they were selected, these probabilities are shown in Fig. 6. For models with more than six features all four amplitude-duration features were included, and the amplitude of the insufflation phase (A_I) was the best predictor. A model that used only A_I produced an F_1 score of 91.85 (91.78–91.94)%. No amplitude constraints were imposed on the greedy detector, so most false positives were caused by small impedance fluctuations, thus the importance of A_I. The model tends to select A_I over A_E because of the difficulty to accurately determine the end of exhalation, although both predictors are very correlated (pearson correlation coefficient of $\rho = 0.894$). Waveform moments were also very correlated within the insufflation and exhalation phases. The smallest correlation coefficients were found between moments 1 and 5, with values of 0.931 and 0.915 for insufflation and exhalation, respectively. Consequently, features μ_{E2} and μ_{E3} could be used interchangeably (see Fig. 6), and when only one was used it became the second most selected feature. The selection probabilities for the moments of the insufflation phase were more evenly distributed.

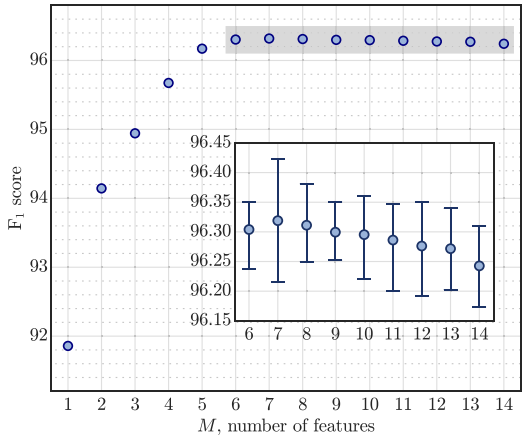


Fig. 5. Performance of the detector as a function of the number of features used in the detector. For $M \geq 6$ performance stabilizes, and the median (IDR) values are zoomed out ($F_1 \geq 96.15\,\%$) in the box.

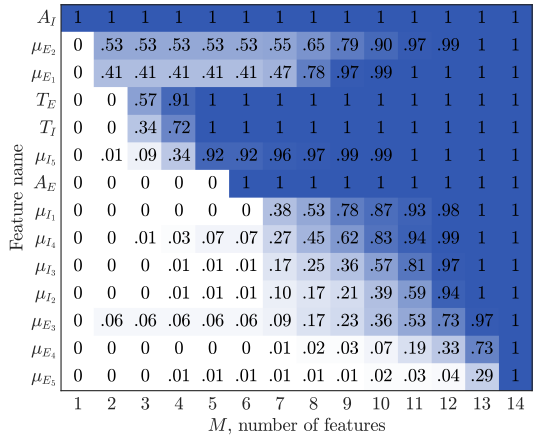
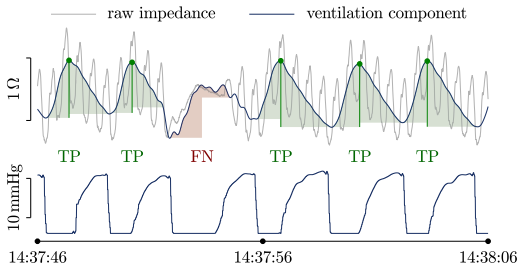


Fig. 6. Probability of selecting a feature ordered by the number of times features were selected.

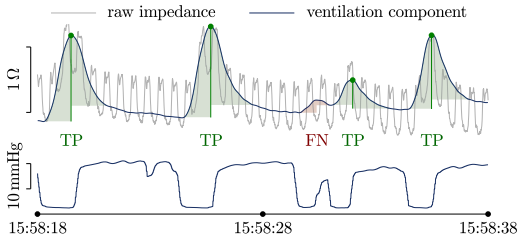
Finally, Fig. 7 shows some typical examples of the errors made by the ventilation detection algorithm. Most missed detections were caused by abrupt changes in impedance, mostly caused by pauses in chest compressions, that produce transient effects when using the LMS filter, or by very shallow ventilations with short durations and possibly low insufflated volumes. Most false positives were caused by low frequency components in the impedance, such as motion artifacts caused by rescuers during treatment.

B. Analysis Per Patient

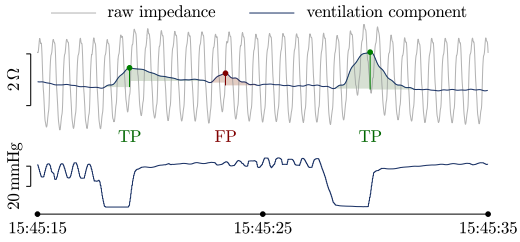
A relevant sub-analysis is to evaluate how the method performs for each patient, and to evaluate in what proportion of



(a) False negative that the classifier missed to validate as a true ventilation. Cause: a pause in chest compressions.



(b) False negative that the classifier missed to validate as a true ventilation. Cause: a low amplitude/duration ventilation.



(c) False positive detection validated by the classifier. Cause: a low frequency fluctuation in the impedance.

Fig. 7. Examples of incorrect ventilation detections. Each example shows the impedance in grey with the ventilation component superposed in blue, and the capnogram with the ground truth ventilations below. The ventilations output by the detector are indicated by dots and are shaded in green (true positive, TP) or red (false positive, FP). The missed ventilations (false negative, FN) are shaded in red.

patients feedback on ventilations could be accurately provided. Fig. 8 shows the distributions for the performance metrics per patient. In the boxplot each sample represents a patient, and for each patient the median value for the metric over the 20-CV partitions is represented. The proportion of patients with very low performance metrics (under 90%) is depicted in the right panel. As shown in the figure, accurate ventilation detection was possible in a large proportion of patients. The F_1-score was above 95% for 77.1% of patients, and above 98 % for 49.4% of patients. For a few patients accurate ventilation detection was not possible with F_1 scores under 75% ($n = 8$). For these patients the amplitude of the impedance ventilation component was small ($< 0.2\,\Omega$), probably because the insufflated volume was low [16].

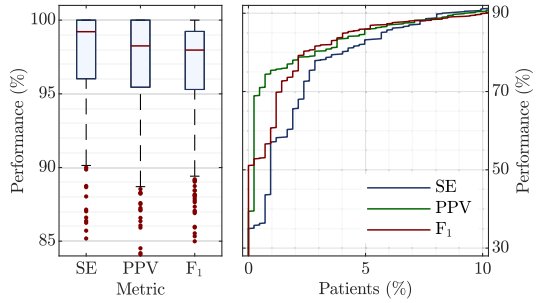


Fig. 8. Per patient performance metrics for all n = 423 patients as boxplots (left), and for the 10% (n = 43) of patients with lowest detection accuracy (right). The rightmost graph only shows the proportion of patients in the low accuracy range ($< 90\%$).

C. Feedback on Ventilation Rate

The most important application of a ventilation detector during OHCA treatment is to provide feedback on ventilation rates. For this purpose the algorithm was implemented in the way it would be incorporated to a monitor-defibrillator. The detector was programmed to analyze 1-min signal intervals, and to give feedback on that minute with no information on future impedance values. The ventilation rate was calculated every 15-s, that is, with a 75% overlap between the 1-min windows. For each window, ventilation time instants (t_{p_i}) were calculated using the process outlined in Fig. 2. The ventilation rate for the interval was calculated as:

$$\mathrm{VR} = \frac{60}{\mathrm{median}\{\Delta t_{p_i}\}}\ (\mathrm{min}^{-1}) \tag{15}$$

These values were compared to those obtained from the capnogram's ground truth annotations, in which ventilation instants were annotated in the exhalation onset (rise in CO_2). The analysis included all patients, but a separate analysis was done excluding the patients ($n = 43$) with low accuracy ($F_1 < 90\%$). For those patients, the impedance had either long intervals of lower quality signal, and/or very low amplitude ventilation components which could be associated to low insufflated volumes [16]. A separate sub-analysis was done excluding those patients, because in those cases the actual problem is not with ventilation rate but with the quality of ventilations (volumes) or with the quality of the signal used to give feedback. The Bland-Altman plot for feedback on ventilation rate is shown in Fig. 9. The global 90% levels of agreement (LoA) were (−0.82, 1.40) min^{-1} for all patients, and (−0.51, 1.10) min^{-1} when the low F_1 patients were excluded. The moving average LoAs for different VR intervals are shown in the figure, in red when all patients were included and in green after excluding low F_1 patients. Rate was overestimated at rates under 6 min^{-1} and underestimated at rates above 10 min^{-1}, although errors were small in all cases. The median (IDR) error and relative error in ventilation rate for all patients were 0.002 (−0.334–0.572) min^{-1} and 0.05 (−3.71–9.08)%, respectively. Excluding the low F_1 patients the error and relative error were 0.002 (−0.204–0.351) min^{-1} and 0.06 (−2.45–5.11)%, respectively. So ventilation feedback

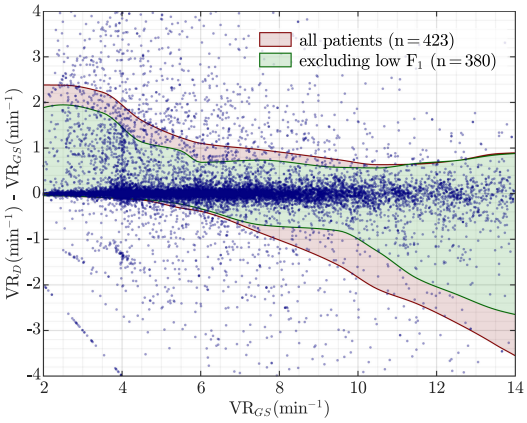


Fig. 9. Bland-Altman plot for feedback on ventilation rate (VR). Ventilation rates were computed using 1-min impedance signal intervals and compared to the ground truth VR obtained from the capnogram. Moving average levels of agreement (LoA) are shown in red for all patients, and in green when the patients for which the ventilation detector's F_1-score was under 90% were excluded ($n = 43$).

TABLE I
COMPARISON OF THE MEDIAN (IQR) SE AND PPV PER PATIENT OF OUR ALGORITHM (FROM FIG. 8, LEFT) DURING MECHANICAL CPR WITH METHODS TO DETECT VENTILATIONS DURING MANUAL CPR

		Per-patient performance	
Studies		Se (%)	PPV (%)
Impedance (manual CPR)			
Risdal *et al.*	[20]	90.6 (12.5)	97.4 (8.0)
Alonso *et al.*	[21]	92.2 (87.4 – 95.8)	81.0 (67.2 – 90.5)
Edelson *et al.*	[13]	78 (67 – 89)	87 (77 – 96)
Capnogram (manual CPR)			
Edelson *et al.*	[13]	82 (75 – 93)	91 (85 – 95)
Aramendi *et al.*	[16]	99.0 (95.7 – 100)	97.6 (94.8 – 100)
Our study (mechanical CPR)		**99.2 (96.0 – 100)**	**98.3 (95.4 – 100)**

could be provided with errors under 9% for all patients, and under 5% for the patients with better quality impedance, which in our dataset amounted for over 90% of patients.

V. DISCUSSION AND LIMITATIONS

This paper presents a new approach to impedance-based ventilation detection during mechanical CPR that combines adaptive signal processing and machine learning techniques. As shown by our results, accurate ventilation detection is possible with median (IQR) Se and PPV per patient of 99.2 (96.0–100)% and 98.3 (95.4–100)%, respectively (see Fig. 8, left). Previous studies have addressed the detection of ventilations during manual CPR using the impedance and capnogram signals, a comparative assessment of our method to those methods is shown in Table I. Our results are comparable to those obtained using state-of-the art algorithms based on the capnogram [14], and were better than those obtained for impedance based methods during manual CPR [13], [20], [21]. Two reasons could explain why results were better for mechanical than for manual CPR in impedance ventilation detection. First, chest compression components in the impedance are much more stable during mechanical CPR because the piston is at a fixed position in the patient's chest, and compressions are always delivered in the same way by the machine [29]. Moreover, since the patient is fixed to the mechanical compressors, other movement artifacts and low quality signal intervals are less frequent. Second, our approach combined adaptive signal processing and machine learning, while the methods presented for manual CPR either relied on overly complex adaptive filters [20], or were based on rule-based detection of ventilations [13], [21]. In the future, new approaches similar to the one presented in this study could be demonstrated during manual CPR to exploit the potential of machine learning algorithms [27], and thus provide a better estimate of how accurate impedance based ventilation detection could be during manual CPR.

A key application of the ventilation detector is ventilation rate feedback during CPR to ensure compliance with the recommended rate of 10 min^{-1}. In our data, ventilation rates were abnormally low, the median (IQR) ventilation rate per patient was 6.0 (4.5–8.0) min^{-1}, and rates only exceeded the recommended values in 12.5% of our patients (for a detailed account see supplementary materials). These low ventilation rates were associated with some distinct ventilation patterns (see figures in supplementary materials). In many cases the patients were not ventilated for intervals of up to one minute, ventilation rates were very low, or the ventilation pattern followed the one observed during 30:2 CPR. Interestingly, the ventilation rates in our data are similar to the 7 min^{-1} ventilation rate observed during 30:2 CPR in a recent study [38], or to the 8 min^{-1} reported for the early stages of treatment during ALS [39]. Our data demonstrates the need for tools like the one presented in this study, both for feedback during treatment but also as a tool for retrospective analysis of large OHCA datasets that could shed light into how patients are being ventilated in the different phases of a resuscitation episode.

Finally, this study has some limitations. First, the algorithm was trained and tested using a 10-fold CV architecture and should be further validated in an independent dataset. Second, the algorithm was tailored to a piston driven mechanical CPR device (LUCAS-2), but there are other mechanical CPR technologies based on load distribution bands (the Autopulse system by Zoll) in which chest compression components may be different. Third, the algorithm can only be used during mechanical CPR and given the cost of mechanical CPR devices many EMS agencies still rely on manual chest compressions, although there is an increased trend towards the use of mechanical devices. And fourth, data was obtained from the Philips MRx device, so the algorithm may need to be readjusted to be used in other monitor-defibrillators with different impedance acquisition circuitry.

VI. CONCLUSION

This study demonstrates the feasibility of an accurate impedance-based ventilation detection during concurrent mechanical CPR. The method efficiently combines adaptive signal

processing techniques to obtain and detect ventilation waveforms, with a machine learning algorithm to identify true ventilations. This ventilation detection algorithm could be used before advanced airway placement and capnography are available during resuscitation, but also to obtain additional information on ventilation such as insufflated volumes that are not available from the capnogram. Its use would broaden both the time feedback on ventilation is available, but also the available information on the quality of ventilations. Moreover, the method could also be used to retrospectively assess the effects of ventilations during CPR in OHCA outcomes, by applying the detector to large datasets of resuscitation episodes.

REFERENCES

[1] J. Berdowski, R. A. Berg, J. G. P. Tijssen, and R. W. Koster, "Global incidences of out-of-hospital Cardiac Arrest and survival rates: Systematic review of 67 prospective studies," *Resuscitation*, vol. 81, pp. 1479–1487, Nov. 2010.

[2] A. Myat, K.-J. Song, and T. Rea, "Out-of-hospital Cardiac Arrest: Current concepts," *Lancet*, vol. 391, pp. 970–979, Mar. 2018.

[3] M. E. Kleinman *et al.*, "Part 5: Adult basic life support and Cardiopulmonary resuscitation quality: 2015 American Heart Association guidelines update for Cardiopulmonary resuscitation and emergency Cardiovascular care," *Circulation*, vol. 132, pp. S414–S435, Nov. 2015.

[4] M. S. Link *et al.*, "Part 7: Adult advanced cardiovascular life support: 2015 American Heart Association guidelines update for Cardiopulmonary resuscitation and emergency Cardiovascular care," *Circulation*, vol. 132, pp. S444–S464, Nov. 2015.

[5] T. Henlin, P. Michalek, T. Tyll, J. D. Hinds, and M. Dobias, "Oxygenation, ventilation, and airway management in out-of-hospital Cardiac Arrest: A review," *BioMed Res. Int.*, vol. 2014, 2014, Art. no. 376871.

[6] T. P. Aufderheide and K. G. Lurie, "Death by hyperventilation: A common and life-threatening problem during Cardiopulmonary resuscitation," *Crit. Care Medicine*, vol. 32, pp. S345–S351, Sep. 2004.

[7] J. F. O'Neill and C. D. Deakin, "Do we hyperventilate Cardiac Arrest patients?" *Resuscitation*, vol. 73, pp. 82–85, Apr. 2007.

[8] G. Vissers, C. Duchatelet, S. A. Huybrechts, K. Wouters, S. Hachimi-Idrissi, and K. G. Monsieurs, "The effect of ventilation rate on outcome in adults receiving Cardiopulmonary resuscitation," *Resuscitation*, vol. 138, pp. 243–249, May 2019.

[9] J. Coult, J. Blackwood, T. D. Rea, P. J. Kudenchuk, and H. Kwok, "A method to detect presence of chest compressions during resuscitation using Transthoracic impedance," *IEEE J. Biomed. Health Inf.*, to be published, doi: 10.1109/JBHI.2019.2918790.

[10] J. Kramer-Johansen, D. P. Edelson, H. Losert, K. Köhler, and B. S. Abella, "Uniform reporting of measured quality of Cardiopulmonary resuscitation (CPR)," *Resuscitation*, vol. 74, pp. 406–417, Sep. 2007.

[11] D. Hostle *et al.*, "Effect of real-time feedback during Cardiopulmonary resuscitation outside hospital: Prospective, cluster-randomised trial," *Brit. Med. J.*, vol. 342, pp. 1–10, Feb. 2011.

[12] D. P. Edelson *et al.*, "Improving in-hospital cardiac arrest process and outcomes with performance debriefing," *Arch. Internal Medicine*, vol. 168, pp. 1063–1069, May 2008.

[13] D. P. Edelson, J. Eilevstjønn, E. K. Weidman, E. Retzer, T. L. V. Hoek, and B. S. Abella, "Capnography and chest-wall impedance algorithms for ventilation detection during Cardiopulmonary resuscitation," *Resuscitation*, vol. 81, pp. 317–322, Mar. 2010.

[14] E. Aramendi *et al.*, "Feasibility of the Capnogram to monitor ventilation rate during Cardiopulmonary resuscitation," *Resuscitation*, vol. 110, pp. 162–168, 2017.

[15] F. S. Stecher, J.-A. Olsen, R. E. Stickney, and L. Wik, "Transthoracic impedance used to evaluate performance of Cardiopulmonary resuscitation during out of hospital Cardiac Arrest," *Resuscitation*, vol. 79, pp. 432–437, Dec. 2008.

[16] E. Aramendi *et al.*, "A novel technique to assess the quality of ventilation during pre-hospital Cardiopulmonary resuscitation," *Resuscitation*, vol. 132, pp. 41–46, Nov. 2018.

[17] P. O. Berve *et al.*, "Transthoracic impedance measured with defibrillator pads-new interpretations of signal change induced by ventilations," *J. Clin. Medicine*, vol. 8, pp. 1–12, May 2019.

[18] S. Ansari, K. Ward, and K. Najarian, "Epsilon-tube filtering: Reduction of high-amplitude motion artifacts from impedance Plethysmography signal," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 2, pp. 406–417, Mar. 2015.

[19] S. Ansari, K. R. Ward, and K. Najarian, "Motion artifact suppression in impedance Pneumography signal for portable monitoring of respiration: An adaptive approach," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 2, pp. 387–398, Mar. 2017.

[20] M. Risdal, S. O. Aase, M. Stavland, and T. Eftestøl, "Impedance-based ventilation detection during Cardiopulmonary resuscitation," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2237–2245, Dec. 2007.

[21] E. Alonso *et al.*, "Reliability and accuracy of the thoracic impedance signal for measuring Cardiopulmonary resuscitation quality metrics," *Resuscitation*, vol. 88, pp. 28–34, Mar. 2015.

[22] J. H. Husøy, J. Eilevstjønn, T. Eftestøl, S. O. Aase, H. Myklebust, and P. A. Steen, "Removal of Cardiopulmonary resuscitation artifacts from human ECG using an efficient matching pursuit-like algorithm," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1287–1298, Nov. 2002.

[23] U. Irusta, J. Ruiz, S. R. de Gauna, T. Eftestøl, and J. Kramer-Johansen, "A least mean-square filter for the estimation of the Cardiopulmonary resuscitation artifact based on the frequency of the compressions," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1052–1062, Apr. 2009.

[24] S. Rubertsson *et al.*, "Mechanical chest compressions and simultaneous defibrillation vs. conventional Cardiopulmonary resuscitation in out-of-hospital Cardiac Arrest: The LINC randomized trial," *J. Amer. Med. Assoc.*, vol. 311, pp. 53–61, Jan. 2014.

[25] L. Wik *et al.*, "Manual vs. integrated automatic load-distributing band CPR with equal survival after out of hospital Cardiac Arrest. the randomized CIRC trial," *Resuscitation*, vol. 85, pp. 741–748, Jun. 2014.

[26] G. D. Perkins *et al.*, and P. trial collaborators, "Mechanical versus manual chest compression for out-of-hospital Cardiac Arrest (PARAMEDIC): A pragmatic, cluster randomised controlled trial," *Lancet*, vol. 385, pp. 947–955, Mar. 2015.

[27] I. Isasi *et al.*, "A multistage algorithm for ECG rhythm analysis during piston-driven mechanical chest compressions," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 263–272, Jan. 2019.

[28] I. Isasi *et al.*, "A machine learning shock decision algorithm for use during piston-driven chest compressions," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1752–1760, Jun. 2019.

[29] E. Aramendi, U. Irusta, U. Ayala, H. Naas, J. Kramer-Johansen, and T. Eftestøl, "Filtering mechanical chest compression artefacts from out-of-hospital cardiac arrest data," *Resuscitation*, vol. 98, pp. 41–47, Jan. 2016.

[30] X. Jaureguibeitia *et al.*, "Impedance based automatic detection of ventilations during mechanical Cardiopulmonary resuscitation," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc.*, Jul. 2019, pp. 19–23.

[31] G. Nichol *et al.*, "Regional variation in out-of-hospital Cardiac Arrest incidence and outcome," *J. Amer. Med. Assoc.*, vol. 300, pp. 1423–1431, Sep. 2008.

[32] U. Ayala *et al.*, "Automatic detection of chest compressions for the assessment of CPR-quality parameters," *Resuscitation*, vol. 85, pp. 957–963, Jul. 2014.

[33] Y. Xiao and Y. Tadokoro, "LMS-based notch filter for the estimation of Sinusoidal signals in noise," *Signal Process.*, vol. 46, no. 2, pp. 223–231, 1995.

[34] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[35] D. Krstajic, L. J. Buturovic, D. E. Leahy, and S. Thomas, "Cross-validation pitfalls when selecting and assessing regression and classification models," *J. Cheminformatics*, vol. 6, no. 1, pp. 1–15, 2014.

[36] B. Gregorutti, B. Michel, and P. Saint-Pierre, "Correlation and variable importance in random forests," *Statist. and Comput.*, vol. 27, no. 3, pp. 659–678, 2017.

[37] I. Isasi *et al.*, "Automatic Cardiac rhythm classification with concurrent manual chest compressions," *IEEE Access*, vol. 7, pp. 115 147–115 159, 2019.

[38] G. Sanson *et al.*, "Impact of 'synchronous' and 'asynchronous' cpr modality on quality bundles and outcome in out-of-hospital Cardiac Arrest patients," *Internal Emergency Medicine*, vol. 14, pp. 1129–1137, Oct. 2019.

[39] L. Wik *et al.*, "Quality of Cardiopulmonary resuscitation during out-of-hospital cardiac arrest," *J. Amer. Med. Assoc.*, vol. 293, pp. 299–304, Jan. 2005.

A.1.2 First conference paper

	Publication in international conference
Reference	Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Erik Alonso, Pamela Owens, Henry Wang, Ahamed Idris, "Impedance based automatic detection of ventilations during mechanical cardiopulmonary resuscitation", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019*, pp. 19-23.
Quality indices	• **Type of publication**: International conference in SJR • **Impact factor**: 0.309

Impedance Based Automatic Detection of Ventilations During Mechanical Cardiopulmonary Resuscitation

Xabier Jaureguibeitia[1], Unai Irusta[1,*], Elisabete Aramendi[1], Erik Alonso[2], Pamela Owens[3], Henry Wang[4], Ahamed Idris[3]

***Abstract*— Monitoring ventilation rate is key to improve the quality of cardiopulmonary resuscitation (CPR) and increase the probability of survival in the event of an out-of-hospital cardiac arrest (OHCA). Ventilations produce discernible fluctuations in the thoracic impedance signal recorded by defibrillators. Impedance-based detection of ventilations during CPR is challenging due to chest compression artifacts. This study presents a method for an accurate detection of ventilations when chest compressions are delivered using a piston-driven mechanical device. Data from 223 OHCA patients were analyzed and 399 analysis segments totaling 3101 minutes of mechanical CPR were extracted. A total of 18327 ventilations were annotated using concurrent capnogram recordings. An adaptive least mean squares filter was used to remove compression artifacts. Potential ventilations were detected using a greedy peak detector, and the ventilation waveform was characterized using 8 waveform features. These features were used in a logistic regression classifier to discriminate true ventilations from false positives produced by the greedy peak detector. The classifier was trained and tested using patient wise 10-fold cross validation (CV), and 100 random CV partitions were created to statistically characterize the performance metrics. The peak detector presented a sensitivity (Se) of 99.30%, but a positive predictive value (PPV) of 54.43%. The best classifier configuration used 6 features and improved the mean (sd) Se and PPV of the detector to 93.20% (0.06) and 94.43% (0.04), respectively. When used to measure per minute ventilation rates for feedback to the rescuer, the mean (sd) absolute error in ventilation rate was 0.61 (1.64) $\mathbf{min^{-1}}$. The first impedance-based method to accurately detect ventilations and give feedback on ventilation rate during mechanical CPR has been demonstrated.**

I. INTRODUCTION

Out-of-hospital cardiac arrest (OHCA) is a major public health problem with a yearly incidence of 29-40 cases per 100 000 persons in Europe, and very low survival rates of around 10% [1]. Early defibrillation and cardiopulmonary resuscitation (CPR) are key therapies in the treatment of OHCA patients. The main purpose of CPR is to artificially maintain a minimum/sufficient flow of oxygenated blood to the heart and brain. This is accomplished by delivering chest compressions and ventilations.

Quality of CPR is crucial for survival, and should comply to the recommendations of the resuscitation guidelines [2]. Chest compressions should be delivered at rates between 100 - 120 $\mathrm{min^{-1}}$, and depth between 5 - 6 cm allowing full chest recoil between compressions [2]. The ventilation rate should be of approximately 10 $\mathrm{min^{-1}}$ and hyperventilation should be avoided [3]. Hyperventilation increases intrathoracic pressure and contributes to hemodynamic deterioration, decreasing the probability of survival [4]. However, hyperventilation is common during resuscitation with ventilation rates ranging from moderate ($\sim 14\,\mathrm{min^{-1}}$) to severe ($> 20\,\mathrm{min^{-1}}$) [5].

Great efforts have been made to improve chest compression quality, either using feedback devices [6] and training in manual CPR or through the use of mechanical compression devices [7]. Although the benefits of mechanical CPR for survival are unclear [7], its use is becoming popular, specially in scenarios like transport or invasive procedures. However, there is a need for accurate methods to measure the presence and characteristics of ventilations during CPR. The capnogram, which measures the partial pressure of the expired CO_2, can be used to identify ventilations [8]. However, capnography is available only after an advanced airway is placed, and cannot be used to estimate tidal volumes. The thoracic impedance is available in most equipment early during resuscitation, and air blown into the lungs produces a characteristic ventilation waveform [9].

Chest compressions produce artifacts in the impedance that hinder the detection of ventilations [10]. Ventilation detection is possible either during pauses in compressions [11], or after the removal of the compression artifacts [9], [12]. Suppression of compression artifacts from the impedance is based on adaptive filters conceived to remove CPR artifacts from the ECG [13], [14]. These methods use accelerometer data to track the variable chest compression frequency of manual CPR [9], [12]. Recently, adaptive filters have been tailored to remove piston-driven mechanical compression artifacts from the ECG [15], [16]. In mechanical CPR the frequency of the compressions is fixed by the device, so accelerometer data is not needed for artifact removal [15].

This study introduces the first method for ventilation detection during mechanical CPR using only impedance data. First, chest compression artifacts are removed using an adaptive filter, and then potential ventilation fluctuations

This work was supported by: The Spanish Ministerio de Economía y Competitividad, TEC2015-64678-R, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), UPV/EHU via GIU17/031.

[1]X. Jaureguibeitia, U. Irusta, E. Aramendi and E Alonso are with the Department of Communications Engineering, University of the Basque Country UPV/EHU, Ingeniero Torres Quevedo Plaza, 1, 48013, Bilbao, Spain (e-mail: unai.irusta@ehu.eus).

[2]E Alonso is with the Department of Applied Mathematics, University of the Basque Country UPV/EHU, Ingeniero Torres Quevedo Plaza, 1, 48013, Bilbao, Spain.

[3]A Idris and P Owens are with the University of Texas Southwestern Medical Center, Dallas, Texas, USA.

[4]H Wang is with the University of Texas Health Science Center, Houston, Texas, USA.

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

are detected and their waveform is characterized using a set of features. Those waveform features are fed to a machine learning algorithm to discriminate ventilation waveforms from other fluctuations in impedance. The paper is organized as follows. The study dataset and the data annotation process are described in section II. The impedance filtering scheme and the ventilation detector are presented in section III. Finally, the results and their importance is contextualized in section IV, followed by the main conclusions of the study.

II. MATERIALS

The study dataset originated from the DFW Center of Resuscitation Research (UTSW, Dallas) that maintains a large repository of OHCA electronic files. A total of 296 episodes with the following characteristics were screened: concurrent recordings of impedance and capnography and confirmed usage of a LUCAS-2 compression system (Physio-Control Inc/Jolife AB, Lund, Sweden) in the clinical records. The impedance was needed to develop the automatic ventilation detection method, and the capnogram to annotate the ground truth ventilation events. The electronic files came from the MRx monitor-defibrillator (Phillips Medical Systems, Andover, MA, USA). Thoracic impedance was recorded with a 200 Hz sampling rate and a resolution of 0.74 mΩ. The capnogram was obtained using Microstream technology (sidestream acquisition) and recorded with a 40 Hz sampling rate and a resolution of 0.004 mmHg.

MRx data were converted to an open file format and reviewed using custom Matlab (MathWorks Inc., Natick, MA) tools. First, intervals in which the LUCAS-2 device was used were traced by identifying a steady chest compression rate of $100\,\mathrm{min}^{-1}$. The compression rate was determined using an impedance based automatic compression detection algorithm [17]. Then analysis intervals during mechanical CPR were selected with the following inclusion criteria: minimum duration of 100 s with compression pauses no longer than 20 s, and concurrent and interpretable impedance and capnography signals. The final dataset consisted of 399 segments from 223 patients. There were a total of 3 101 min of signal recordings for analysis. The median (IQR) analysis time per patient was 13 min (9-19).

Ground truth ventilations were semi-automatically annotated using the capnogram. Inspiration downstroke and expiration upstroke instants were automatically detected using a recently proposed algorithm [8]. Then, all automatic annotations were visually reviewed by an experienced biomedical engineer. The capnogram signal was time advanced (2 - 4 s) to compensate for the gas transport delay of the sidestream technology, aligning the inspiratory downstroke (drop in CO_2) to the onset of impedance fluctuations (start of lung air filling). Figure 1 shows an example, in which the inspiration phase annotated in the capnogram is shaded. A total of 18 327 ground truth ventilations were annotated, and the median (IQR) number of ventilations per patient was 72 (45-112).

III. METHODS

The ventilation detection method is composed of three stages. First, the impedance is filtered to obtain the ventilatory waveform. Then, a peak detection algorithm is used to detect potential ventilation events, which are characterized using eight impedance waveform features. This stage was conceived to maximize the detection of potential ventilation events, at the risk of producing many false positives. Finally, the classification stage discriminates these false positives from actual ventilation events using the impedance waveform features.

A. Ventilation Enhancement

The raw impedance signal, $s_{\mathrm{imp}}(n)$, was filtered to obtain the ventilation component, $s_{\mathrm{V}}(n)$. The signal was first downsampled ($f_s = 50$ Hz) to ease the design of the filters. All linear filters were applied in a forward-backward configuration to avoid phase distortion and time delays. Zero

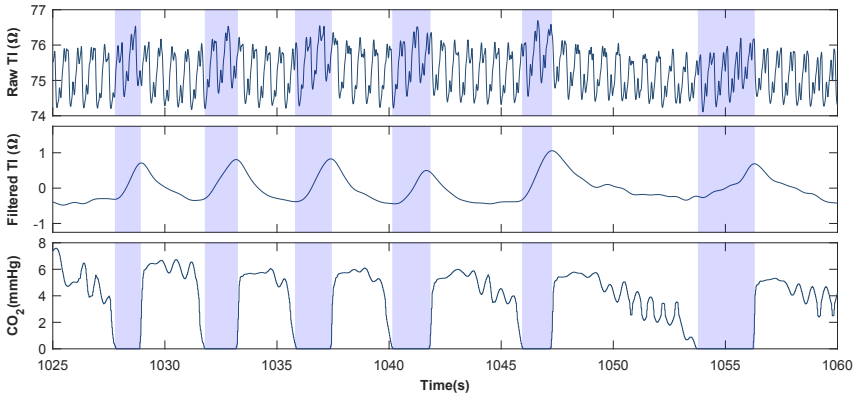


Fig. 1. Example of the signals used in the study dataset. The raw impedance with mechanical compression artifacts is shown on top. The middle trace shows the ventilatory impedance component used to detect ventilations. The bottom trace shows the capnogram for ground truth ventilation annotations. The inspiration phase between the inspiratory downstroke and the expiratory upstroke are highlighted in the capnogram.

phase filtering is possible as 1-min signal windows are normally analyzed for ventilation rate feedback.

Three filters were used, two linear 4th-order Chebyshev filters with 1 dB passband ripple, and a least mean squares (LMS) adaptive filter to remove the chest compression artifacts. First, a high-pass filter with a cutoff frequency of 0.05 Hz was used to remove the DC component and the slow baseline drifts. Then the LMS filter was applied, followed by a low-pass filter with a 1.5 Hz cutoff to remove high frequency residuals from the LMS filter.

The design of the LMS filter is based on a method introduced to remove manual chest compression artifacts from the ECG [14], and has been recently adapted to remove mechanical chest compression artifacts [16]. The chest compression artifact, $s_{\text{cc}}(n)$, is modeled as a N harmonic quasi periodic interference of slowly time-varying amplitudes, $a_k(n)$ and $b_k(n)$:

$$\hat{s}_{\text{cc}}(n) = \sum_{k=1}^{N} a_k(n)\cos(k\omega_{\text{c}}n) + b_k(n)\sin(k\omega_{\text{c}}n) \quad (1)$$

where $\omega_{\text{c}} = 0.209$ is the discrete angular frequency of the LUCAS-2 100 min^{-1} compression rate for a 50Hz sampling frequency. This equation can be expressed compactly as $\hat{s}_{\text{cc}}(n) = \mathbf{w}^{\text{T}}(n)\mathbf{x}(n)$ arranging the in-phase and quadrature harmonic components as a known input vector $\mathbf{x}(n)$, and the time-varying amplitudes as filter coefficients $\mathbf{w}(n)$:

$$\mathbf{x}(n) = [\cos(\omega_{\text{c}}n), \sin(\omega_{\text{c}}n) \ldots \cos(N\omega_{\text{c}}n), \sin(N\omega_{\text{c}}n)]^{\text{T}} \quad (2)$$

$$\mathbf{w}(n) = [a_1(n), b_1(n), a_2(n) \ldots a_N(n), b_N(n)]^{\text{T}} \quad (3)$$

In the LMS model, a gradient descent is used to minimize the mean-square error using instantaneous estimates for the input signal correlation matrix and the cross-correlation vector between input and desired signals. Under these assumptions, the problem is equivalent to minimizing the instantaneous squared error and filter weights are updated as:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\nabla_{\mathbf{w}}e^2(n) \quad (4)$$

$$e(n) = d(n) - \hat{s}_{\text{cc}}(n) = d(n) - \mathbf{w}^{\text{T}}(n)\mathbf{x}(n) \quad (5)$$

$$\nabla_{\mathbf{w}}e^2(n) = -2e(n)\mathbf{x}(n) \quad (6)$$

where $d(n) = s_{\text{imp}}(n)$ and $e(n) = \hat{s}_{\text{V}}(n)$ is an estimate of the ventilation component. We chose $N = 6$ and $\mu = 0.22$ after some preliminary tests, since they produced a good visual removal of the artifact while ensuring the convergence of the LMS filter.

B. Detection of potential ventilations

A greedy peak detector was used to identify potential ventilations in $s_{\text{V}}(n)$. Initially, all peaks in $s_{\text{V}}(n)$ were detected with the only restriction that peaks should be separated by at least 1.4 s (i.e. maximum ventilation rate of 40 min^{-1}). Let $t_{p,i}$ denote the time instant at which peak i occurs. The detector searched for start time $t_{s,i}$ for ventilation $t_{p,i}$ in the interval $t_{si} \in (t_{p,i-1}\,,\, t_{p,i} - 0.55)$. Start time was defined as a timepoint of a local minima, or of an steep change in positive slope. A series of slope projection tests were done to ensure that $t_{s,i}$ did not fall in the ascending phase of the impedance. If a start time could not be determined for $t_{p,i}$ in the prescribed interval, or if $t_{s,i}$ was separated by more than 6 s from $t_{p,i}$ the peak was discarded. Once all peak and start times were determined, ventilation end times were detected, $t_{e,i}$ using an analogous procedure. The definitions of the search intervals for $t_{s,i}$ and $t_{e,i}$ is not critical and was heuristically determined using a few samples from the dataset.

C. Feature Extraction

The waveform of $s_{\text{V}}(n)$ in the interval $t_{s,i} \leqslant t \leqslant t_{e,i}$ was used to compute the 8 waveform features shown in Fig. 2. These features are the amplitude, duration, area and curve-length of $s_{\text{V}}(n)$ for the inspiration (upstroke) and expiration (downstroke) phases. Let us denote by $n_{s,i}$, $n_{p,i}$, $n_{e,i}$ the sample indexes corresponding to $t_{s,i}$, $t_{p,i}$, $t_{e,i}$, respectively. Then inspiration starts at $n_1 = n_{s,i}$ and ends at $n_2 = n_{p,i}$, and expiration starts at $n_1 = n_{p,i}$ and ends at $n_2 = n_{e,i}$. The features are simply calculated as:

$$Z_i = s_{\text{V}}(n_2) - s_{\text{V}}(n_1), \qquad T_i = T_s \cdot (n_2 - n_1) \quad (7)$$

$$A_i = T_s\sum_{n=n_1}^{n_2} s_{\text{V}}(n) - s_{\text{V}}(n_1), \qquad cl_i = \sum_{n=n1}^{n_2} T_s^2 + \dot{s}_{\text{V}}^2(n) \quad (8)$$

where $\dot{s}_{\text{V}}(n)$ in the curve length calculation is the first difference of $s_{\text{V}}(n)$. So, for a peak detected at $t_{p,i}$ the following feature vector was obtained:

$$\boldsymbol{x}_i = [T_{ui},\, Z_{ui},\, A_{ui},\, cl_{ui},\, T_{di},\, Z_{di},\, A_{di},\, cl_{di}] \quad (9)$$

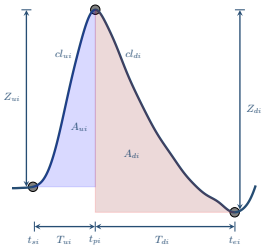


Fig. 2. Visual representation of the ventilation waveform features

D. Classification

The peak detector was designed to identify all possible ventilations, so it produces many false positive detections. A classifier was designed to discriminate the false positives from the actual ventilations. All detections from the peak detector were labeled as either $y_i = 1$, ventilation (true positive), or $y_i = 0$, no ventilation (false positive). The peak at $t_{p,i}$ was considered a ventilation if it fell within the inspiration phase (extended by 1-s) of the capnogram. When more than one peak was detected in that region, the peak closest to inspiration onset was labeled as true positive, and the rest as false positives.

The peaks output from the detector formed a dataset of instance-labels $\{(\boldsymbol{x}_1, y_1), ..., (\boldsymbol{x}_n, y_n)\} \in \mathbb{R}^{K\times(0,1)}$, where $K = 8$ is the number of features. A simple logistic regression classifier was adjusted to obtain the probability that peak i corresponds to a ventilation, $p(\boldsymbol{x}_i)$, which is:

$$p(\boldsymbol{x}_i) = \frac{1}{1 + e^{-\boldsymbol{\beta}^T \boldsymbol{x}_i}} \tag{10}$$

With coefficient and feature vectors $\boldsymbol{\beta} = [\beta_0, \ldots \beta_K]$ and $\boldsymbol{x}_i = [1, x_{i1}, \ldots, x_{iK}]$, respectively. A peak was considered a ventilation if $p(\boldsymbol{x}_i) \geqslant 0.5$.

The dataset was partitioned patientwise into 10-fold cross validation (CV) partitions to estimate the accuracy of the classifier. For each fold the remaining folds were used to train the classifier, which was then tested on the selected fold. The process was repeated until all instances of the dataset were classified. A total of 100 CV partitions were randomly obtained to statistically characterize the accuracy of the classifier. In addition, only quasi-stratified partitions were allowed by enforcing that each fold should deviate by less than 15% from the proportion of the classes found in the whole dataset. During training the instances were weighted so that all patients contributed equally to the model.

The performance of the classifier was evaluated in terms of sensitivity (Se), positive predictive value (PPV) and F_1-score (F_1). Correctly identified ventilations were true positives (TP), missed ventilations were false negatives (FN) and peak detections that did not correspond to actual ventilations false positives (FP). These metrics were computed for the last stage (classification), but also for the complete solution since the peak detection stage missed some true ventilations (FN for the complete solution). Patients were also equally weighted when computing the performance metrics.

IV. RESULTS

There were 33 432 peak detections output from the peak detector, 15 234 false positives and 18 198 actual ventilations. The peak detector missed 129/18 327 ventilations, it had a high Se of 99.30% but a low PPV of 54.43%. The classification stage removed most of the false positive detections. The mean (sd) performance metrics for the 100 CV partitions for a classifier with 8 classification features were, 93.58% (0.07), 94.33% (0.05) and 93.95% (0.05) for the Se, PPV and F_1. The metrics for the complete solution (peak detection + classifier) were 92.92% (0.07), 94.33% (0.05) and 93.62% (0.05). Fig 3 shows some typical cases of false detections and missed detections during peak detection (cases b/c) and caused by the classifier (case d).

All possible combinations of features were tested for classifiers with $K = 1, 2, \ldots 8$ features. The performance of the best classifier for each K is shown in Fig. 4. The best results were obtained for a classifier with the following features: $(T_u, Z_u, A_u, cl_u, T_d, A_d)$. This resulted in Se, PPV and F_1 scores of 93.86% (0.06), 94.43% (0.04) and 94.14% (0.04) for the classifier and of 93.20% (0.06), 94.43% (0.04) and 93.81% (0.04) for the complete system. The classifier worked best with upstroke features (inspiration phase), so Table I shows the mean performance metrics when only the upstroke or downstroke features were used. A simpler peak detector and classifier could be built only using the inspiration phase of the ventilation waveform.

TABLE I

BEST CLASSIFIERS USING UPSTROKE / DOWNSTROKE FEATURES.

	Classifier			Complete system		
Model	Se	PPV	F_1	Se	PPV	F_1
Upstroke						
Z_u	87.8	94.0	90.8	87.2	94.0	90.5
Z_u, A_u	87.8	94.3	90.9	87.1	94.3	90.6
T_u, Z_u, cl_u	92.8	93.5	93.2	92.1	93.5	92.8
T_u, Z_u, A_u, cl_u	93.1	94.0	93.6	92.6	94.0	93.3
Downstroke						
Z_d	81.7	91.3	86.2	81.2	91.3	86.0
Z_d, cl_d	82.8	90.5	86.5	82.3	90.5	86.2
T_d, Z_d, cl_d	86.7	90.6	88.6	86.1	90.6	88.3
T_u, Z_u, A_u, cl_u	88.4	91.4	89.9	87.9	91.4	89.6

Then, the detector's capacity to time-locate the start of a ventilation was determined for the best solution. The mean

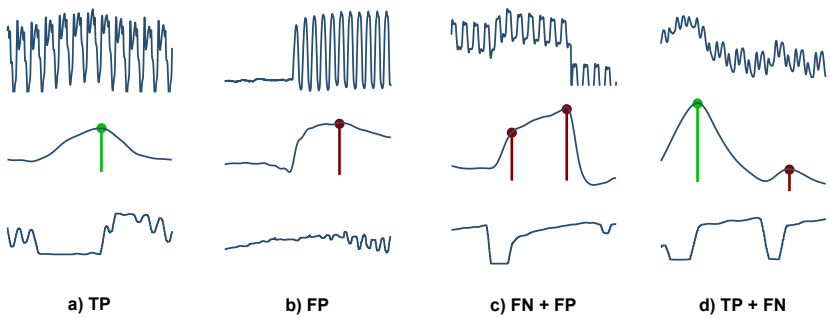


Fig. 3. Examples of correct (green) an incorrect (red) detections. From top to bottom the raw impedance, ventilation component and capnogram are depicted. The first case shows a typical ventilation (TP). The second case corresponds to a FP caused by the transition from compression pauses to compressions with edge transient filtering effects. In the third case the actual ventilation had no downstroke and was missed (FN), and a peak was labeled as ventilation later but corresponded to an artifact in the impedance (FP). Finally, the last example shows two ventilations the first is correctly classified (TP), the second peak was detected but later removed by the classifier (FN) because of its low amplitude.

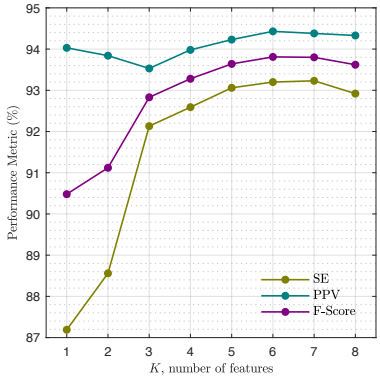


Fig. 4. Performance metrics for the best classifier with K features. All possible combinations of K features were tested.

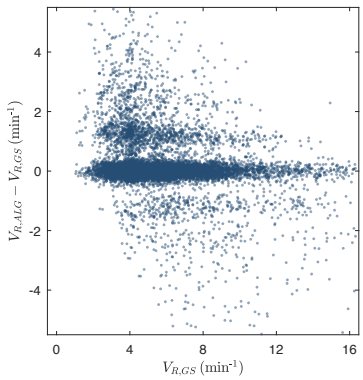


Fig. 5. Bland-Altman plot of the error in the per minute ventilation rate estimation using the automatically detected ventilations.

(sd) absolute error of these time differences was 0.20 s (0.08) when averaged per patient, or 0.20 s (0.22) for all ventilations. This shows that our solution locates ventilations very accurately in time.

Finally, a typical application scenario was evaluated in which the detector was used to give feedback on the per minute ventilation rate (V_R), estimated as:

$$V_R = \frac{60}{\overline{\Delta T_v}} \, (\text{min}^{-1}) \tag{11}$$

where $\overline{\Delta T_v}$ is the mean interval between consecutive ventilation onsets in the 1-min analysis window. A sliding window of 15 s was used for feedback. The procedure was repeated for the ventilations detected the most accurate system, and the ventilation rates from the ground truth ($V_{R,GS}$) and algorithm ($V_{R,ALG}$) were compared. The results are shown as a Bland-Altman plot in Fig 5. The mean (sd) absolute error in ventilation rate was 0.61 (1.64) min^{-1}, and the 90% level of agreement ranges were -1.83 – 1.74 min^{-1}. In 85.9% of feedbacks the absolute error was under 1 min^{-1}, and in 95.5% under 2 min^{-1}.

V. CONCLUSIONS

A system to detect ventilations during mechanical chest compressions was demonstrated, and its value to give an accurate feedback on ventilation rate was shown. The system is based solely on the impedance and could therefore be used in any monitor-defibrillator during resuscitation. This is, to the best of our knowledge, the first accurate solution for automatic ventilation rate feedback during mechanical compressions.

REFERENCES

[1] C. Atwood, M. S. Eisenberg, J. Herlitz, and T. D. Rea, "Incidence of EMS-treated out-of-hospital cardiac arrest in Europe," *Resuscitation* vol. 67, no. 1, pp. 75–80, 2005.

[2] G. D. Perkins, A. J. Handley, R. W. Koster, et al, "European resuscitation council guidelines for resuscitation 2015: Section 2. adult basic life support and automated external defibrillation," *Resuscitation* vol. 95, pp. 81–99, 2015.

[3] J. Soar, J. P. Nolan, B. W. Böttiger, et al, "European resuscitation council guidelines for resuscitation 2015," *Resuscitation* vol. 95, pp. 100–147, 2015.

[4] T. Henlin, P. Michalek, T. Tyll, J. D. Hinds, and M. Dobias, "Oxygenation, ventilation, and airway management in out-of-hospital cardiac arrest: a review." *BioMed research international*, vol. 2014, p. 376871, 2014.

[5] T. P. Aufderheide and K. G. Lurie, "Death by hyperventilation: a common and life-threatening problem during cardiopulmonary resuscitation," *Crit. Care. Med.*, vol. 32, no. 9, pp. S345–351, 2004.

[6] S. O. Aase and H. Myklebust, "Compression depth estimation for CPR quality assessment using DSP on accelerometer signals." *IEEE Trans. Biomed. Eng.*, vol. 49, pp. 263–268, 2002.

[7] G. D. Perkins, R. Lall, T. Quinn, et al, "Mechanical versus manual chest compression for out-of-hospital cardiac arrest (paramedic): a pragmatic, cluster randomised controlled trial," *The Lancet*, vol. 385, no. 9972, pp. 947–955, 2015.

[8] E. Aramendi, A. Elola, E. Alonso, et al, "Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation." *Resuscitation*, vol. 110, pp. 162–168, 2017.

[9] M. Risdal, S. O. Aase, M. Stavland, and T. Eftestol, "Impedance-based ventilation detection during cardiopulmonary resuscitation," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2237–2245, 2007.

[10] E. Alonso, J. Ruiz, E. Aramendi, et al, "Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics," *Resuscitation* vol. 88, pp. 28–34, 2015.

[11] E. Aramendi, Y. Lu, M. P. Chang, et al, "A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation," *Resuscitation*, vol. 132, pp. 41–46, 2018.

[12] D. P. Edelson, J. Eilevstjønn, E. K. Weidman, et al, "Capnography and chest-wall impedance algorithms for ventilation detection during cardiopulmonary resuscitation," *Resuscitation*, vol. 81, no. 3, pp. 317–322, 2010.

[13] J. H. Husoy, J. Eilevstjonn, T. Eftestol, S. O. Aase, H. Myklebust, and P. A. Steen, "Removal of cardiopulmonary resuscitation artifacts from human ECG using an efficient matching pursuit-like algorithm," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1287–1298, 2002.

[14] U. Irusta, J. Ruiz, S. R. de Gauna, T. Eftestøl, and J. Kramer-Johansen, "A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions." *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 1052–1062, 2009.

[15] I. Isasi, U. Irusta, E. Aramendi, et al, "A multistage algorithm for ECG rhythm analysis during piston-driven mechanical chest compressions," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 263–272, 2019.

[16] I. Isasi, U. Irusta, A. Elola, et al, "A machine learning shock decision algorithm for use during piston-driven chest compressions." *IEEE Trans. Biomed. Eng.*, 2019 (In press).

[17] U. Ayala, T. Eftestøl, E. Alonso, et al, "Automatic detection of chest compressions for the assessment of CPR-quality parameters." *Resuscitation*, vol. 85, pp. 957–963, 2014.

A.1.3 Second journal paper

	Publication in international journal
Reference	Xabier Jaureguibeitia, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris, "Impedance-based ventilation detection and signal quality control during out-of-hospital cardiopulmonary resuscitation", *IEEE Journal of Biomedical and Health Informatics 2023*, vol. 27, no. 6, pp. 3026-3036.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (3/31) based on Web of Science Rank 2022 • **Impact factor**: 7.700

Impedance-Based Ventilation Detection and Signal Quality Control During Out-of-Hospital Cardiopulmonary Resuscitation

Xabier Jaureguibeitia, Elisabete Aramendi, *Member, IEEE*, Henry E. Wang, and Ahamed H. Idris

***Abstract*—Feedback on ventilation could help improve cardiopulmonary resuscitation quality and survival from out-of-hospital cardiac arrest (OHCA). However, current technology that monitors ventilation during OHCA is very limited. Thoracic impedance (TI) is sensitive to air volume changes in the lungs, allowing ventilations to be identified, but is affected by artifacts due to chest compressions and electrode motion. This study introduces a novel algorithm to identify ventilations in TI during continuous chest compressions in OHCA. Data from 367 OHCA patients were included, and 2551 one-minute TI segments were extracted. Concurrent capnography data were used to annotate 20724 ground truth ventilations for training and evaluation. A three-step procedure was applied to each TI segment: First, bidirectional static and adaptive filters were applied to remove compression artifacts. Then, fluctuations potentially due to ventilations were located and characterized. Finally, a recurrent neural network was used to discriminate ventilations from other spurious fluctuations. A quality control stage was also developed to anticipate segments where ventilation detection could be compromised. The algorithm was trained and tested using 5-fold cross-validation, and outperformed previous solutions in the literature on the study dataset. The median (interquartile range, IQR) per-segment and per-patient F_1-scores were 89.1 (70.8–99.6) and 84.1 (69.0–93.9), respectively. The quality control stage identified most low performance segments. For the 50% of segments with highest quality scores, the median per-segment and per-patient F_1-scores were 100.0 (90.9–100.0) and 94.3 (86.5–97.8). The proposed algorithm could allow reliable, quality-conditioned feedback on ventilation in the challenging scenario of continuous manual CPR in OHCA.**

***Index Terms*—Adaptive filter, cardiac arrest, cardiopulmonary resuscitation (CPR), quality control, recurrent neural network (RNN), thoracic impedance, ventilation.**

I. Introduction

OUT-OF-HOSPITAL cardiac arrest (OHCA) is a major cause of death in industrialized countries. Emergency medical services (EMS) assess about 350,000 cases each year in the US alone, and survival rates to resuscitation efforts are mostly low, around or below 10% [1]. A patient in cardiac arrest loses spontaneous circulation and breathing, leading to death within minutes if not treated. High-quality cardiopulmonary resuscitation (CPR), consisting of chest compressions and ventilations, maintains a minimum flow of blood and oxygen, and is critical to improve survival from OHCA [2], [3]. Thus, considerable effort has been made to improve the overall quality of CPR. Resuscitation guidelines [4], [5], [6] are periodically updated with the latest evidence-based recommendations for CPR delivery. Supportive technologies have also been developed, such as portable accelerometers to estimate the depth of compressions [7], and many algorithms have been proposed to extract CPR information from different biomedical signals [8], [9], [10]. When integrated into field equipment, these solutions enable real-time feedback to the rescuer, improving adherence to guideline recommendations [11]. When used retrospectively, they facilitate the annotation and analysis of large OHCA registries for either quality programs or research [12]. However, most technical advances in CPR monitoring, analysis and feedback have focused on chest compressions. The importance of measuringventilation during resuscitation is strongly supported by evidence [13], but current technology to monitor ventilation in OHCA is limited, and the optimal ventilation strategy remains unclear [14], [15].

Ventilation in OHCA is typically assessed using end-tidal capnography, which monitors the partial pressure of CO_2 in exhaled gases [16], [17]. However, capnography is not usually available until late phases of resuscitation, once an advanced airway is placed, nor does it provide information on insufflated air volumes. Thoracic impedance (TI) is sensitive to air volume changes in the lungs, and has been extensively used to monitor respiratory events [18], [19]. Most basic defibrillators acquire TI along with the electrocardiogram through the defibrillation

Manuscript received 15 September 2022; revised 4 January 2023; accepted 26 February 2023. Date of publication 7 March 2023; date of current version 6 June 2023. This work was supported in part by the "ERDF A way of making Europe" and MCIN/AEI /10.13039/501100011033 under Grant PID2021-122727OB-I00, in part by Basque Government under Grants PRE-2021-2-0126 and IT-1717-22, in part by the University of the Basque Country (UPV/EHU) under Grant COLAB20/01, and in part by the National Heart Lung and Blood Institute (the parent PART trial) under Grant UH2/UH3-HL125163. *(Corresponding author: Elisabete Aramendi.)*

Xabier Jaureguibeitia and Elisabete Aramendi are with the Department of Communications Engineering, University of the Basque Country UPV/EHU, 48013 Bilbao, Spain (e-mail: xabier.jaureguibeitia@ehu.eus; elisabete.aramendi@ehu.eus).

Henry E. Wang is with the Ohio State University, Columbus, OH 43210 USA (e-mail: Henry.Wang@osumc.edu).

Ahamed H. Idris is with the University of Texas Southwestern Medical Center, Dallas, TX 75390-7208 USA (e-mail: Ahamed.Idris@UTSouthwestern.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/JBHI.2023.3253780, provided by the authors.

Digital Object Identifier 10.1109/JBHI.2023.3253780

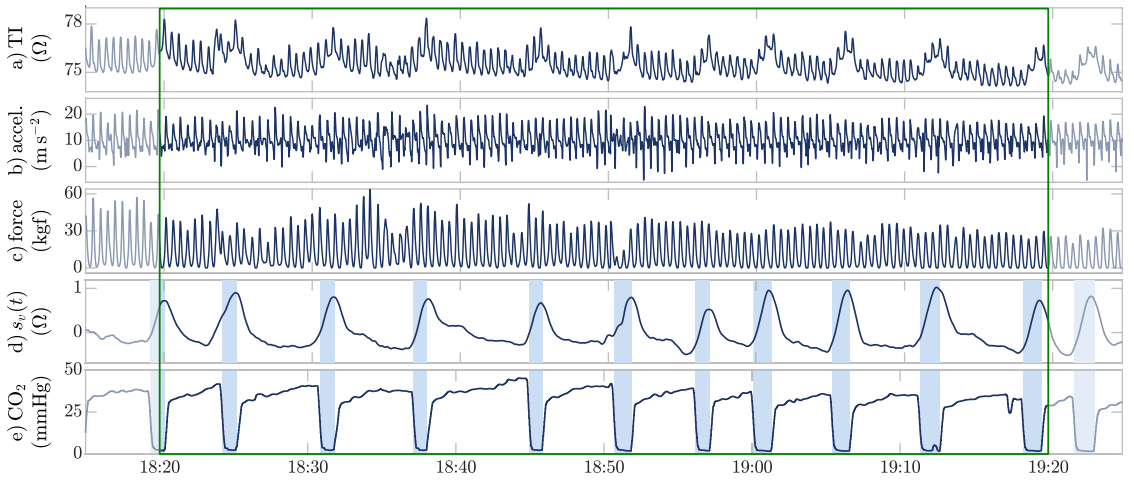


Fig. 1. Example of a one-minute segment, with all the signals used in the algorithm functioning and development: (a) the raw thoracic impedance (TI); (b) and (c), chest force and acceleration, used as references to remove compression artifacts from the TI; (d) $s_v(t)$, the ventilation component of the TI, obtained after the preprocessing stage of the algorithm; and (e) the time-aligned capnogram, used as ground truth for training and evaluation. The ventilations annotated in the capnogram are shaded in blue, and closely match the inflation of fluctuations in $s_v(t)$. Additional 5 s of padding (blurred, outside the green box) were included at both ends to allow the full characterization of fluctuations taking place near the edges.

pads, making it one of the earliest signals available in OHCA. Thus, it could be used in a range of scenarios, from monitoring ventilations early during compression pauses [20], to fine-tuning capnogram readings after patient intubation [16]. Moreover, the amplitude of the TI fluctuations due to ventilations correlates with tidal volume [21] and, while patient-dependent, could offer insights on the effectiveness of ventilation [22].

Impedance-based ventilation detection can be challenging, though. Ventilatory waves may adopt a wide range of amplitudes and durations [23]. The signal itself is very sensitive to electrode motion [24], [25], frequent in ambulatory scenarios such as OHCA. Moreover, during late phases of resuscitation, ventilations are delivered continuously, concurrently with chest compressions; these produce a large artifact which has to be removed for a reliable ventilation detection. Current solutions include harsh static filtering [26], and adaptive filtering based on different compression reference signals, such as those from accelerometers [16], [23].

This study introduces a novel solution for impedance-based ventilation detection during continuous chest compressions in OHCA. Inspired by a previous work on mechanical CPR [27], this study proposes an algorithm for the more general case of rescuer-delivered CPR, where compression artifacts are far more irregular and motion noise levels larger than in mechanical. A preliminary version of this study has been reported [28]. The present work comprises more than twice as many OHCA cases, and improves on its performance by introducing bidirectional adaptive filtering and time-series classification. A signal quality control stage is also presented, which anticipates segments where ventilation detection could be compromised.

II. Data Sources and Preparation

Study data included the de-identified files from 367 OHCA patients treated by EMS between March 2016 and November 2017 in the Dallas - Fort Worth area (Texas, US), all enrolled in the Pragmatic Airway Resuscitation Trial (PART, NCT02419573) [29]. Data collection was approved under US federal rules for Exception From Informed Consent for emergency research (21 CFR 50.24). The files were acquired using a HeartStart MRx monitor-defibrillator (Philips Medical Systems, Andover, MA, US), and included TI, capnography, chest force, and chest acceleration recordings. TI was recorded with a sampling rate of 200 Hz and a resolution of 2.5 $\mathrm{m}\Omega$. The capnogram was acquired using Microstream (sidestream) technology, and recorded with a sampling rate of 40 Hz and a resolution of 0.004 mmHg. Force and acceleration were acquired using a Q-CPR assist pad, and recorded with a sampling rate of 100 Hz and resolutions of 0.01 kgf and 0.01 $\mathrm{m/s^2}$, respectively. Fig. 1 shows an example of the signals involved in the algorithm development. All files were converted and processed using Matlab (MathWorks Inc., Natick, MA, US).

Ventilations were identified in the capnogram and used as ground truth to develop the impedance-based detection algorithm. Ventilations were first automatically annotated using a state-of-the-art solution [17], and then manually reviewed. Capnogram intervals that could not be reliably reviewed were deemed uninterpretable and excluded from the study. On the finally included intervals, the automatic pre-annotation showed a sensitivity of 91.1% and a positive predictive value of 95.6%. The time-delay of the capnogram was also manually corrected by aligning expiration upstrokes with TI fluctuations during

compression-free intervals. A default time-delay of 3.5 s was considered when no clean fluctuations could be identified [16]. The observed mean (standard deviation, SD) time-delay was of 3.5 (0.3) s.

Impedance intervals suitable for the study design were then selected, which included concurrent recordings of acceleration, force, and interpretable capnogram as per the manual review. Chest compression pauses longer than 20 s were excluded, in order to consider mostly CPR artifacted TI. Abrupt TI excursions and other unusually large artifacts were also discarded. Given that ventilation rates are typically measured over one-minute periods [30], the intervals were sub-divided into non-overlapping 60 s segments, with additional 5 s of starting and ending signal padding, as shown in Fig. 1. The final dataset comprised 2551 one-minute segments and 20724 ventilations, median (interquartile range, IQR) of 6 (3–10) minutes and 45 (23–78) ventilations per patient. The 97.1% of ventilations were concurrent with chest compressions, thus reproducing the CPR conditions.

III. Methods

A ventilation detection algorithm was developed using the three-block architecture introduced in previous works [27], [28]. Fig. 2 shows a block-diagram of this layout: First, the TI signal was preprocessed to remove chest compression artifacts and enhance the ventilation waveform. Then, impedance fluctuations potentially due to ventilations were identified at instants t_p and characterized by a set of waveform features $\boldsymbol{x}$. This stage was designed to over-estimate the number of ventilations, resulting in many false positives. Finally, fluctuation data were modeled as a time series, and fed to a recurrent neural network (RNN) to discriminate the actual ventilations. The algorithm was trained and tested using non-overlapping one-minute TI segments. An additional quality control stage was also considered to prevent erroneous feedback under low signal quality or heavy noise conditions.

A. Signal Preprocessing

First, the TI, force and acceleration signals were resampled to a common frequency of $f_s = 50$ Hz. Slow baseline drifts and high frequency components were removed from all three signals using a 0.06 Hz–5 Hz band-pass filter (Butterworth, 4th order) [23]. Then, an adaptive Kalman filter and smoother, with force and acceleration as reference signals, was applied to remove compression artifacts from the TI. Finally, the TI was smoothed using an order 100 finite impulse response low-pass filter; a cut-off frequency of 1 Hz was selected, which is frequently used for impedance smoothing in respiration-related studies [19], [25]. Both static filters were applied in a forward-backward configuration to avoid phase distortion and delay.

In order to set up the adaptive filter, the high-pass filtered TI signal $s(t)$ was assumed to be the sum of a component due to chest compressions $s_{cc}(t)$ and a component due to ventilations $s_v(t)$, such that $s(t) \approx s_{cc}(t) + s_v(t)$. In addition, $s_{cc}(t)$ was modeled as a linear combination of the neighboring force and acceleration samples, so for a time instant t_j:

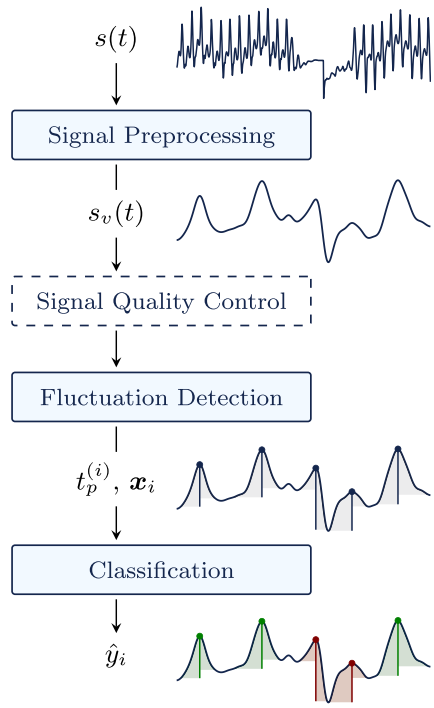


Fig. 2. Block diagram of the ventilation detection algorithm. The raw impedance signal $s(t)$ is first filtered to obtain the ventilation component $s_v(t)$. Then, fluctuations potentially due to ventilations are located at instants t_p^i, and characterized by a waveform feature vector $\boldsymbol{x}_i$. Finally, a recurrent neural network jointly classifies fluctuations as ventilations ($\hat{y}_i = 1$, shaded in green) or false positives ($\hat{y}_i = 0$, shaded in red). A discretionary signal quality control block allows to identify impedance segments where ventilation detection could be compromised.

$$s_{cc}(t_j) = \sum_{k=-M}^{M} a_k(t_j) s_a(t_j + kT_s) + b_k(t_j) s_f(t_j + kT_s) \tag{1}$$

where $T_s = 1/f_s$ is the sampling period, and $s_a(t)$ and $s_f(t)$ the downsampled and high-pass filtered force and acceleration signals. The slowly time-varying coefficients $a_k(t)$ and $b_k(t)$ were also assumed to follow gaussian processes with Ornstein-Uhlenbeck type covariances of length-scale λ^{-1} [31], [32]. The process equation for the Kalman recursion follows:

$$\mathbf{x}_{j+1} = \mathbf{F}_{j+1,j}\mathbf{x}_j + \mathbf{w}_j, \tag{2}$$

where the state vector $\mathbf{x}_j$ is given by

$$\mathbf{x}_j = [a_{-M}(t_j), \ldots, a_M(t_j), b_{-M}(t_j), \ldots, b_M(t_j)], \tag{3}$$

the transition matrix by $\mathbf{F}_{j+1,j} = \exp(-\lambda T_s) \cdot \mathbf{I}_{2M+1}$, and $\mathbf{w}_j$ follows a gaussian process with zero mean and covariance $\mathbf{Q}_j = q\,(1 - \exp(-2\lambda T_s)) \cdot \mathbf{I}_{2M+1}$. Similarly, the measurement equation follows:

$$\mathbf{y}_j = \mathbf{H}_j\mathbf{x}_j + v_j, \tag{4}$$

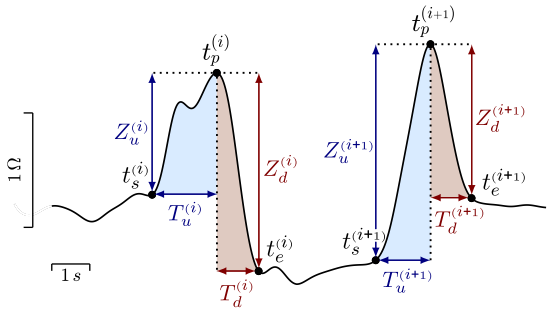


Fig. 3. Example of two $s_v(t)$ fluctuations, as identified by the fluctuation detection stage. The amplitude (Z_u) and duration (T_u) of the upwards or inspiration phase (from t_s to t_p, in blue), and the amplitude (Z_d) and duration (T_d) of the downwards or expiration phase (from t_p to t_e, in red) were computed to characterize each fluctuation.

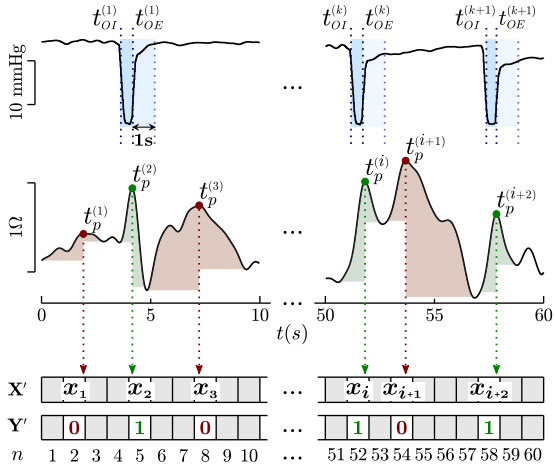


Fig. 4. Example of the fluctuation labeling and time-series composition procedures. Fluctuations were labeled as ventilations ($y_i = 1$, shaded in green) if their peak position t_p^i fell within the bounds (extended by up to 1 s) of a capnogram ground truth ventilation k, and as $y_i = 0$ (in red) otherwise. Fluctuation data were then used to compose the 60-step feature ($\mathbf{X}'$) and label ($\mathbf{Y}'$) time series used in classification. Each time-step in the series represented a one-second interval; fluctuations were mapped to time-steps according to their peak position t_p^i.

with observation vector $\mathbf{H}_j$ given by

$$\mathbf{H}_j = [s_a(t_j - MT_s), \ldots, s_a(t_j + MT_s), \\ s_f(t_j - MT_s), \ldots, s_f(t_j + MT_s)]^T, \quad (5)$$

and $v_j \sim \mathcal{N}(0, R)$. Given these equations, the coefficients $a_k(t_j)$ and $b_k(t_j)$ were obtained using a Rauch-Tung-Striebel smoother, as described in [33]. The ventilation component of interest was finally estimated as $\hat{s}_v(t_j) = s(t_j) - \hat{s}_{cc}(t_j)$, with $\hat{s}_{cc}(t_j) = \mathbf{H}_j \mathbf{x}_j$. Values of $M = 10$, $\lambda = 0.1$, $q = 0.0015$ and $R = 1$ were chosen after some initial experiments. In the following, the term $s_v(t)$ is used to represent the output of the entire preprocessing stage.

B. Fluctuation Detection

Impedance fluctuations potentially due to ventilations were identified in $s_v(t)$. First, the largest local maxima $t_p^{(i)}$ with a minimum separation of 1.5 s were detected. Then, the start of the inflation phase, $t_s^{(i)}$, and the end of the deflation phase, $t_e^{(i)}$, were determined for each fluctuation. Inflation and deflation durations between 0.45 s and 5.5 s were considered [27]. The procedure to determine $t_e^{(i)}$ and $t_s^{(i)}$ was not critical for the overall performance of the algorithm. A detailed description of the heuristic used in this study is included as supplementary material.

C. Feature Extraction

Each fluctuation was characterized in terms of a vector $\boldsymbol{x}$ of 14 waveform features. As shown in Fig. 3, the first four features comprised the amplitudes (Z_u, Z_d) and durations (T_u, T_d) of the inspiration (or upwards, from t_s to t_p) and expiration (or downwards, from t_p to t_e) phases of the fluctuation, given by:

$$Z_u = s_v(t_p) - s_v(t_s), \qquad T_u = t_p - t_s \\ Z_d = s_v(t_p) - s_v(t_e), \qquad T_d = t_e - t_p \quad (6)$$

The remaining 10 features consisted on the curve fit coefficients of each phase in terms of order $m = 0, \ldots, 4$ Legendre polynomials $P_m(z)$. These polynomials form an orthogonal system in the $z \in [-1, 1]$ real domain, and can be recursively obtained through

$$P_{m+1}(z) = (2m - 1)zP_m(z) - mP_{m-1}(z), \quad (7)$$

with $P_0(z) = 1$ and $P_1(z) = z$. Let $\mathbf{s}_u$ be a column vector containing the samples of $s_v(t)$ within the inspiration phase $[t_s, t_p]$. Let also $\mathbf{p}_m$ be another column vector, obtained from evaluating (7) over a finite set of points $\mathbf{z}$, with as many points as samples in $\mathbf{s}_u$, equispaced between -1 and 1. The curve fit coefficients of the inspiration phase $c_{m,u}$ were then computed using:

$$c_{m,u} = \frac{\mathbf{s}_u^{\mathrm{T}} \mathbf{p}_m}{||\mathbf{p}_m||^2} \quad \forall m \in 0, \ldots, 4 \quad (8)$$

An analogous procedure was followed to obtain the expiration phase coefficients $c_{m,d}$.

D. Classification

For classification and evaluation purposes, each final fluctuation output by the previous stage was labeled as either actual ventilation ($y_i = 1$) or false detection ($y_i = 0$), based on the ground truth annotations in the capnogram. As shown in Fig. 4, ventilations in the capnogram were annotated covering the full inspiration cycle, from inspiration onset or downstroke, at time t_{OI}, to expiration onset or upstroke, at time t_{OE}. A fluctuation i was labeled as $y_i = 1$ when its peak position $t_p^{(i)}$ fell within one of these inspiration cycles, with an extra tolerance margin of 1 s. When several fluctuations met this criteria for the same

ground truth ventilation k, only the one with peak $t_p^{(i)}$ closest to $t_{\mathrm{OE}}^{(k)}$ was labeled as $y_i = 1$, and the rest as $y_i = 0$.

Classification was performed over full one-minute segments, which could provide the classifier with comparative context on fluctuation shapes and relative positioning. Each segment was modeled as a time series, where each time-step $n \in \{1..60\}$ represented a one-second interval. As shown in Fig. 4, fluctuations i were mapped to time-steps n according to their peak position $t_p^{(i)}$, such that $n - 1 \leq t_p^{(i)} < n$. Then, a feature time series $\mathbf{X}' = \{\boldsymbol{x'}_n\}$ was constructed, which contained the waveform features $\boldsymbol{x}_i$ of the individual fluctuations at mapped time-steps, and an all-zero feature vector at unmapped ones. A ground truth label series $\mathbf{Y}' = \{y'_n\}$ was constructed using the same procedure to train and evaluate the classifier. In this case, unmapped steps were assigned to the negative class ($y'_n = 0$).

The classification task was performed using an RNN with a two-layer architecture: A bidirectional recurrent layer, comprising 20 gated recurrent units (GRU) [34], to which the time series were fed, and a single neuron output layer, activated by a sigmoid function to produce an output p between 0 and 1. The network worked on a sequence-to-sequence configuration, so a different output p'_n was obtained for each time-step. The outputs p_i associated to fluctuations were recovered from the previously mapped time-steps, and the fluctuations classified as true ventilations ($\hat{y}_i = 1$) if $p_i \geq 0.5$, and as false detections ($\hat{y}_i = 0$) otherwise.

Given the primary interest on the positive class ($y'_n = 1$, the actual ventilations), and the large number of negative class instances (mostly from unmapped time-steps), the models were optimized using the Dice coefficient loss [35], given by:

$$DL\left(y'_n, p'_n\right) = 1 - \frac{2y'_n p'_n + 1}{y'_n + p'_n + 1}. \tag{9}$$

The training of the models was performed over 25 epochs, with a batch size of 32 segments, and using an Adam optimizer with initial learning rate of 0.005. The entire classification stage was implemented in Tensorflow 2.0 [36].

E. Evaluation

Segments were partitioned patient-wise into training and test sets using a 5-fold cross-validation (CV) strategy. Assignment was conducted in a balanced manner, such that approximately one-fifth (maximum deviation of 5%) of the patients, segments and ground truth ventilations were assigned to each fold. A total of 20 random CV partitions were generated to minimize any bias due to data partitioning.

For each one-minute segment, performance was assessed in terms of sensitivity (Se), positive predictive value (PPV) and $\mathrm{F_1}$-score, given by:

$$\mathrm{Se} = \frac{\mathrm{TP}}{N_{\mathrm{GT}}},\ \mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}},\ \mathrm{F_1} = 2\frac{\mathrm{Se} \cdot \mathrm{PPV}}{\mathrm{Se} + \mathrm{PPV}} \tag{10}$$

where TP and FP are the number of true positives ($y_i = 1$, $\hat{y}_i = 1$) and false positives ($y_i = 0, \hat{y}_i = 1$), respectively. Se was computed against the number of ground truth ventilations (N_{GT}), in order to account also for missed fluctuations ($\nexists y_i$). Given the differences in the number of segments per patient, performance metrics were also computed patient-wise, using aggregated data from all segments of each patient. Peformance scores from different partitions were mean-aggregated into a single value for each segment/patient. The distributions of segments and patients for different performance score bands were analyzed. Overall performances by segment/patient were assessed in terms of median (IQR).

1) Comparison With Literature Solutions: The performance of the proposed algorithm was compared to that of various similar solutions in the literature. Risdal et al. proposed an adaptive filter to remove compression artifacts, followed by a machine learning framework for ventilation segmentation [23]. Edelson et al. replaced the machine learning framework with a real-time fluctuation detector and a fixed rule-based discrimination of ventilations [16]. Finally, Alonso et al. proposed the use of static linear filters, and an adaptive rule-based discrimination, using dynamic thresholds based on previous detections [26]. The three methods were implemented in Matlab and trained and tested in the study dataset. Implementation details can be found in Appendix I.

2) Feature Importance: Permutation importance and recursive feature elimination were used to assess the contribution of the different waveform features to classification. First, the RNN models were train-tested on clean data and a mean $\mathrm{F_1}$ of reference (more robust to small differences than the median) was computed. Then, for each feature considered, its values along fluctuations were randomly shuffled and new predictions were obtained. The feature whose permutation caused the smallest decrease in mean $\mathrm{F_1}$ was deemed the least important and was removed from the model. This process was repeated until a single feature remained.

In order to prevent data leakage, the importance analysis was conducted at each regular CV step, using an inner 4-fold CV loop over training data. Different feature rankings were thus obtained for each fold and partition.

F. Signal Quality Control

Impedance is sensitive to many noise sources, such as chest compressions and electrode motion, which may jeopardize the detection of ventilations. A signal quality control solution was designed, which could be integrated in the ventilation detection algorithm to anticipate the reliability of the detection.

The solution used the filtered TI signal $s_v(t)$, downsampled to 5 Hz, to compute the following waveform features:

- The skewness of the sample distribution, which should be high for narrow, positive fluctuations, such as in Fig. 1.
- The first peak amplitude (FPA) of the normalized signal autocorrelation [37], which should be larger for regular ventilation rates and fluctuation shapes.
- SD12, a relational measure between short- and long-term variabilities [38], computed from the Poincaré plot of the first differences of the signal.

These features were used to fit a linear regression model where the $\mathrm{F_1}$ scores (scaled to unit range) of the ventilation detection algorithm worked as target variable. A logit link function was

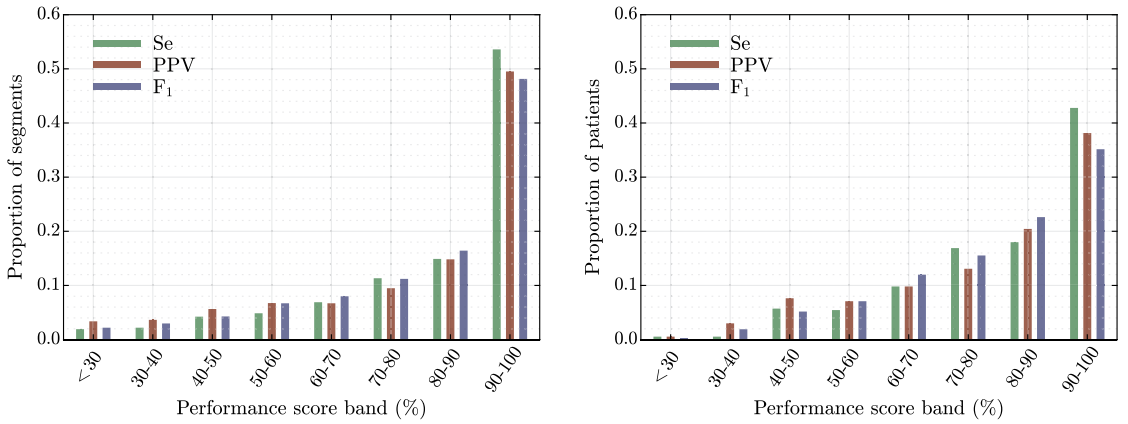


Fig. 5. Distributions of both segments (left) and patients (right) for different sensitivity (Se), positive predictive value (PPV), and F_1 score bands.

applied to produce outputs between 0 and 1. As in the feature importance analysis, an inner 4-fold CV loop (per regular CV step) was conducted on training data to obtain an F_1 score for each training segment.

The regression models were applied to test data to obtain a quality score (QS), a rough estimate of F_1, for each segment. Spearman's correlation was used to measure the ability of the QS to rank segments by F_1. The performance of the ventilation detection algorithm was then reassessed for different segment inclusion rates; that is, when only the target fraction (inclusion rate) of segments with highest QS were evaluated.

Finally, the reliability of the ventilation detection algorithm (combined with the quality control) in providing feedback on ventilation rate (VR) was analyzed. For each segment, the VR was computed as

$$\mathrm{VR} = 60 \;/\; \overline{\Delta t_v^{(i)}} \quad \left(\mathrm{min}^{-1}\right), \tag{11}$$

where ventilation instants $t_v^{(i)}$ were taken at ventilation peaks $t_p^{(i)}$ for estimated VR, or at capnogram $t_{\mathrm{OE}}^{(i)}$ for ground truth VR. Errors and confidence intervals were assessed through a Bland-Altman plot.

IV. Results

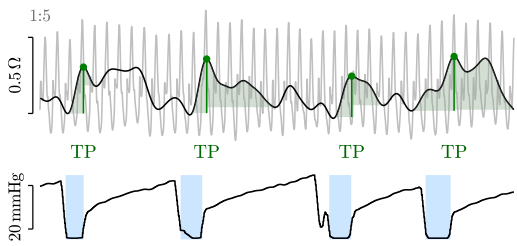
The ventilation detection algorithm, evaluated per segment, achieved performance scores of 93.3 (75.0–100.0) % for Se, 90.0 (68.5–100.0) % for PPV, and 89.1 (70.8–99.6) % for F_1. When computed per patient, scores were 86.5 (71.6–95.1) % for Se, 85.4 (68.3–94.7) % for PPV, and 84.1 (69.0–93.9) % for F_1. Fig. 5 shows the distributions of both segments and patients for different performance decile bands. In both cases, the distributions presented negative skew, with most samples corresponding to the highest performance band. There were many segments, though, for which performance was probably too low for practical use, mostly related to high noise levels and/or low-amplitude ventilations (see Fig. 6). As shown in Fig. 6(b), pauses in chest compressions were also identified as a potential error source. Per-patient metrics, which aggregated information from several segments, were less likely to present extreme performance values.

Lower per-patient median performances were in part due to a higher number of segments for the patients with better overall performances (the 50% and 25% of patients with highest F_1s comprised, respectively, the 56.6% and 26.6% of segments), but also due to within-patient variability: the 63.5% and 36% of patients included, respectively, one or more segments with F_1 scores 10 and 20 points below the patient's aggregate. Note that OHCA treatment may undergo important changes over the course of an episode, such as patient transportation or the placement of an advanced airway. Moreover, rescuers usually take turns delivering CPR, which may result in very different compression artifacts and motion noise levels.

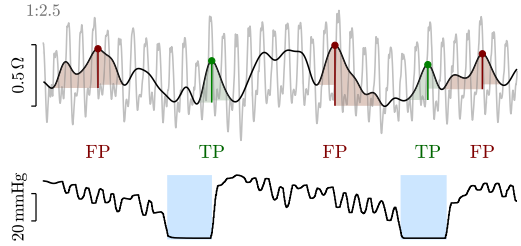
Performance variation between partitions was minimal, with median per-segment and per-patient F_1 SDs of 0.1% and 0.2%, respectively. The variation among the 100 individual test folds was much larger, though, with median per-segment and per-patient F_1s ranging from 82.4% to 91.1%, and from 74.1% to 89.5%, respectively. This was most likely due to the random assignment of patients with very different TI signal qualities.

A. Comparison With Literature Solutions

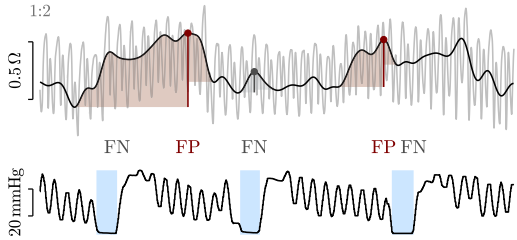
A previous version of this study [28] used a harmonic chest compression model [39], [40] and a least-mean-squares (LMS) filter to remove compression artifacts, and a Random Forest (RF) classifier [41] to individually discriminate fluctuations. As shown in Table I, the preprocessing and classification stages introduced in the present work outperformed that preliminary design. Replacing the Kalman smoother with the LMS filter from the preliminary study resulted in median per-segment and per-patient F_1s of 86.5% and 81.5%, respectively. Similarly, replacing the RNN with an RF classifier resulted in median per-segment and per-patient F_1s of 85.8% and 81.0%, showing the benefits of a broader context on TI fluctuations.



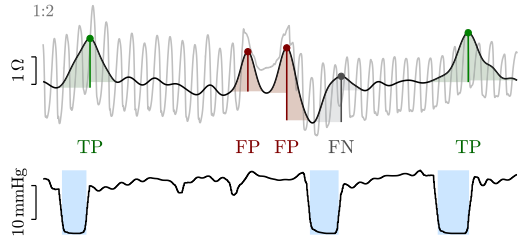
(a) Part of a segment with good performance (F_1 of 95%), despite low amplitude and distorted ventilation waveforms.



(b) Low ventilation amplitudes, with ventilation-level noise producing many false positives.



(c) Very low amplitude ventilations, obscured by stronger noise. Fluctuations could not be located for two out of three ventilations.



(d) False positives due to baseline changes during a compression pause. An actual ventilation was also distorted and missed.

Fig. 6. Examples of error-prone ventilation impedance waveforms. Intervals of 20 s are presented, along with their corresponding ground truth capnogram waveform. The unfiltered impedance signal, depicted in the background in gray, is scaled (scale shown top-left of each figure) so fluctuations are highlighted. True positives (TP), false positives (FP) and false negatives (FN) are indicated.

TABLE I
MEDIAN (IQR) PERFORMANCE METRICS OF PREVIOUS SOLUTIONS IN THE LITERATURE, EVALUATED ON THE STUDY DATASET

	Performance by segment			Performance by patient		
Algorithm	F_1 (%)	Se (%)	PPV (%)	F_1 (%)	Se (%)	PPV (%)
Proposed algorithm	89.1 (70.8 – 99.6)	93.3 (75.0 – 100.0)	90.0 (68.5 – 100.0)	84.1 (69.0 – 93.9)	86.5 (71.6 – 95.1)	85.4 (68.3 – 94.7)
Preliminary design [28]	85.2 (66.4 – 96.4)	87.5 (64.0 – 100.0)	87.9 (68.8 – 100.0)	80.3 (65.2 – 90.7)	80.0 (59.6 – 91.5)	83.9 (69.0 – 93.4)
Risdal et al. [23]	75.0 (53.3 – 90.0)	94.1 (75.0 – 100.0)	71.4 (46.7 – 91.0)	70.0 (56.3 – 82.8)	87.3 (71.7 – 95.8)	65.2 (48.1 – 82.1)
Edelson et al. [16]	66.7 (40.0 – 85.7)	83.5 (50.0 – 100.0)	71.4 (43.3 – 100.0)	62.1 (46.5 – 70.5)	77.3 (52.2 – 91.2)	64.3 (45.8 – 83.6)
Alonso et al. [26]	75.0 (55.2 – 92.3)	100.0 (85.7 – 100.0)	66.7 (42.9 – 90.0)	68.6 (56.1 – 85.5)	92.5 (84.6 – 97.1)	60.2 (43.4 – 79.5)

The proposed algorithm also outperformed similar solutions in the literature, as implemented and tested in the study dataset (see Table I). The algorithm by Alonso et al. [26] achieved the highest Se, but produced many false positives (low PPV). The dynamic amplitude threshold was able to reject relatively small artifacts; however, its starting value was low (>0.1 Ω), rejecting only the smallest fluctuations, and adaptation failed or was too slow under sustained noise. In the opposite end, the solution by Edelson et al. [16] imposed much more severe amplitude (>0.25 Ω) and, most importantly, slope (>0.2 Ω/s) constraints. This resulted in a lower Se, as many of the low-amplitude ventilations in the dataset were missed (a Se of zero was obtained for 9.1% of the segments). The PPV did not improve much, as false positives were still frequent when noise levels exceeded the thresholds. The most elaborate solution, by Risdal et al. [23], did not substantially improve results either. The segmentation framework used a short window of 1.4 s to identify the inspiration (OI) and expiration (OE) onsets, which may had provided too little context to discriminate ventilations from other artifacts. The final decision ruleset was also rather simple, reliant on a correct segmentation, and did not include information of the expiration phase. When optimized for the study dataset, the algorithm was found to require a significant amplitude constraint (>0.3 Ω). Finally, the solution was also penalized by a 2.3 Hz smoothing filter, which could have let too much noise and compression residuals through. Results were better with a more conservative 1 Hz cut-off (median F_1 of 76.9% and 71.8% for segments and patients, respectively), but still far behind those from the algorithm in this study.

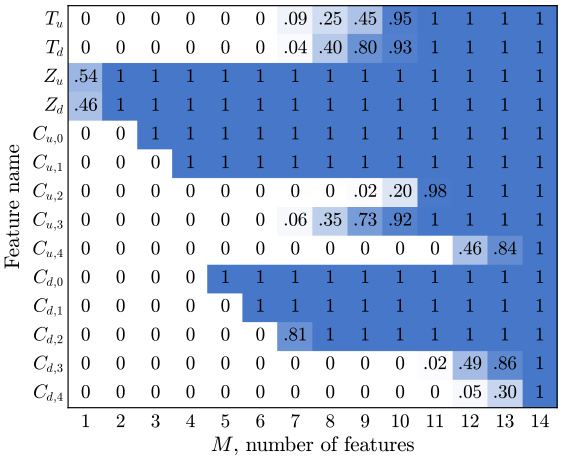


Fig. 7. Probability of selecting a given feature in a simplified M-feature model, computed as the proportion of training folds and partitions for which the feature was selected through recursive feature elimination.

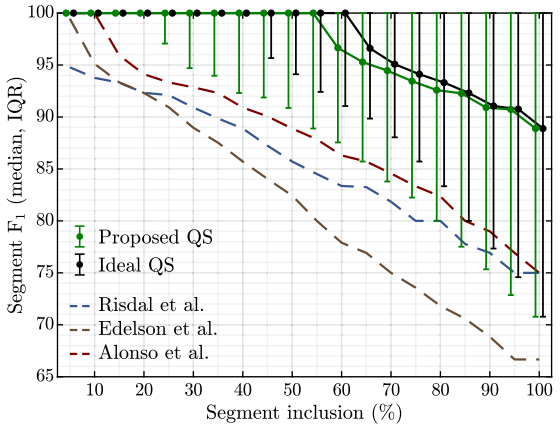


Fig. 8. Median (IQR) per-segment F_1 scores for increasing segment inclusion rates. Segments considered for evaluation were selected according to the proposed quality score (QS), and performance compared to an ideal QS (the F_1 itself). Median and IQR values were averaged between partitions. Median F_1 scores are also shown for the different solutions in the literature, with quality control optimized for each case.

B. Feature Importance

As shown in Fig. 7, feature importance was very consistent across folds and partitions, with the exception of the two most important features, Z_u and Z_d, which were found on par. High order coefficients ($C_{u,4}, C_{d,3}, C_{d,4}$) could likely be removed from the model without affecting performance. All remaining features contributed to performance, although the RNN proved robust to less detailed information. A simplified model using the four most important features ($Z_u, Z_d, C_{u,0}, C_{u,1}$) scored a median per-segment F_1 of 88.6%. Similarly, a model including only the inflation amplitude Z_u scored a median F_1 of 87.1%. Note that this was possible due to the larger context managed by the RNN classifier. An optimal universal amplitude threshold (found at 0.25 Ω), used independently in all fluctuations, scored a much lower F_1 of 78.3%, more in line with previous solutions in the literature.

C. Signal Quality Control

The quality control stage identified segments where ventilation detection was defective. The solution assigned a QS, an estimate of the F_1, to each segment. A Spearman's correlation of $\rho = 0.7$ was measured between QS and F_1 values, proving that the solution could reliably sort segments by performance. Ad-hoc QS thresholds could be defined, each with a different segment inclusion rate and expected performance range.

Fig. 8 shows the median (IQR) per-segment F_1 for different segment inclusion rates considered for evaluation. Median and quartile scores showed monotonic growth as inclusion became more restrictive and less segments were considered. Moreover, the median F_1 scores were close to those of an ideal QS (the F_1 itself), reaching 100% for segment inclusions up to 55%. First quartile values, however, were lower, indicating that some low F_1 segments eluded the control. For an inclusion rate of 50%, with overall F_1 of 100.0 (90.9–100.0) %, 9.3% and 5.1% of the selected segments presented, respectively, F_1s below 80% and 70%. Causes were varied, but often involved a dominant segment section (due to artifacts or larger ventilations) which governed the extraction of quality features. Similarly, 7.8% of all study segments with F_1 of 100% were incorrectly left out, mostly showing distorted fluctuations like in Fig. 6(a).

Per-patient performances were notably lower, as many patients marginally contributed with relatively low F_1 segments. A segment inclusion of 50% comprised data from 73.2% of all patients, for a per-patient F_1 of 94.3 (86.5–97.8) %. However, 21.7% of these patients were very partially included (less than 25% of each patient's available data) and represented only the 6.4% of the segments considered. These showed, in general, lower performances, with F_1 of 85.7 (61.7–93.4) %.

All the impedance features considered in the quality control model were statistically significant, ($p < 0.01$). The individual Spearman's correlations were 0.40 for FPA, 0.47 for Skewness, and −0.64 for SD12. The regression coefficients (using normalized features) were 0.28, 0.27 and −0.59, respectively, denoting SD12 as the most relevant feature. The quality control model was also implemented for the solutions in the literature. In all three cases, the model was effective at sorting segments by F_1, but the relative importance of the features (as given by regression coefficients) varied. Note that quality features were computed on the filtered TI, which was different in most cases. Moreover, some solutions included harsh amplitude thresholds, whereas features were intentionally amplitude independent.

D. Feedback on Ventilation Rates

Finally, the reliability of the ventilation detection algorithm to provide VR feedback was assessed. Fig. 9 shows a Bland-Altman plot on VR, in which 90% levels of agreement (LoA)

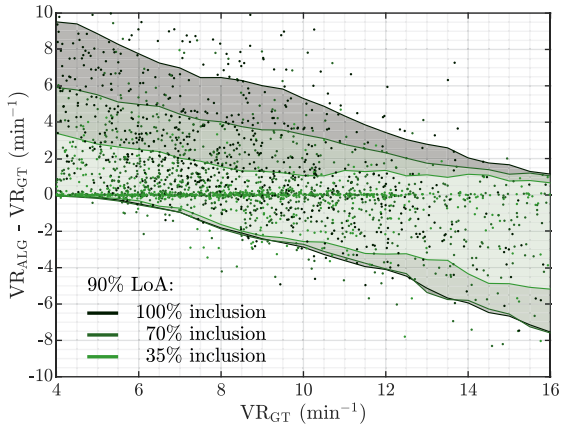


Fig. 9. Bland-Altman plot, comparing estimated (VR_{ALG}) and ground truth (VR_{GT}) ventilation rates. The 90% levels of agreement (LoA) were computed for different VR_{GT} intervals (interval width of 3 /min and step of 0.5 /min) and segment inclusion rates (according to quality control results). Individual VR_{ALG} and LoA values were averaged partition-wise.

are depicted for the whole dataset, as well as 70% and 35% data inclusion in combination with the quality control. Errors were larger without quality control, with a partition-averaged global LoA of (−3.8, 6.8) /min. Overestimation was frequent for low to moderate rates, the median (IQR) ground truth VR being 8.4 (6.1–11) /min, whereas underestimation was prevalent for high rates (>12 /min). Errors were larger for extreme rates, probably affected by data imbalance; 90% of the segments were in the (4.1–16.2) /min range.

Quality control helped prevent many of the errors. Global LoA were (−3.4, 3.9) /min and (−2.0, 2.0) /min for 70% and 35% data inclusion, respectively. However, the improvement was highly biased, avoiding mostly overestimation errors in low to moderate VR segments.

V. Discussion

This study presents a novel algorithm for impedance-based ventilation detection during continuous chest compressions in OHCA. Evaluated on ground truth annotations taken entirely in the capnogram, the algorithm achieved median F_1 scores of 89.1% and 84.1% for one-minute segments and full patients, respectively, which represents about 4 points of improvement over its preliminary version [28]. Two main differences may explain this improvement: First, the Kalman smoother used in this study, with force and acceleration as reference signals, proved superior at removing chest compression artifacts compared to the previous LMS approach and other unidirectional adaptive filters. And second, the RNN classifier was fed with combined information from all fluctuations within the one-minute segment, which provided added context to discriminate ventilations.

The proposed method also outperformed previous solutions in the literature [16], [23], [26] by more than 10 points of F_1. For all of them, performance was lower in our study dataset than in the original study. In general, these solutions treated each potential ventilation independently, and relied ultimately in amplitude thresholding; this would require ventilation levels to be overall larger than noise ones, which may not be the case in our data or in a real setting. Only the solution by Alonso et al. [26] included some case-specific context, but adapted too slowly and produced many false positives in noisy scenarios. Differences in data sources could explain these performance disparities. All the data in Edelson et al. [16] as well as part of the data in Risdal et al. [23] came from an in-hospital setting, in which noise levels could have been lower. In contrast, our dataset was collected entirely out-of-hospital, and 21.6% of the included segments were acquired during transport, which further hindered the detection of ventilations (median F_1 of 84.0%, versus 90.3% for non-transport segments). Similarly, endotracheal intubation may have been the norm in hospital data [42], whereas 13.1% of our segments had no advanced airway in place, and 56.5% corresponded to laryngeal tubes, which have been shown to produce lower TI amplitudes [43]. Risdal et al. reported a median (IQR) patient average inflation amplitude of 1.0 (0.8–1.3) Ω, while our dataset, considering only flawlessly classified segments, showed a much lower 0.4 (0.3–0.7) Ω. Our results show that, in these circumstances, a universal amplitude threshold alone is not feasible; a broader context on the surrounding fluctuations and/or a more detailed characterization is needed to reliably identify ventilations.

Despite the improved performance, we still identified many TI segments where ventilation detection was unreliable. This is in contrast with our previous study on mechanical CPR [27], in which noise levels were found minimal and capnography-grade performance was obtained. Thanks to the quality control we introduced, we could anticipate low-performance segments and prevent erroneous feedback. While different approaches such as template matching [25] or convolutional networks [37] could be explored in the future, our linear regression solution proved robust enough while being highly interpretable, computationally cheap, and easy to integrate in the detection algorithm. No fixed binary quality labels, but the continuous F_1 scores of the detection algorithm were used to fit the model, so ad-hoc quality threshold could easily be defined according to performance requirements. All the quality features were also amplitude-independent. Given the high correlation between F_1 and quality scores, our detection algorithm should be able to identify low-amplitude ventilations as long as not significantly distorted.

The impact of ventilation therapy in the outcome of cardiac arrest patients has been documented, but little is known about the details of ventilation during OHCA because it has not been adequately measured either in research studies or in practice [13]. Our ventilation detection algorithm could be used in two main scenarios: a) to provide real-time feedback on ventilation during OHCA treatment, and b) for the retrospective analysis of ventilation effectiveness in large datasets. Notice that real-time solutions could analyze overlapping segments to increase the frequency of readings. Other modifications could also be considered, such as performing filtering and fluctuation detection in shorter sub-segments to reduce memory requirements, or

optimizing the signal padding to minimize delay. Whatever the application mode, a complete solution should also include logic to identify isolated large artifacts, which were manually excluded in this study. As a concrete application, the algorithm could be used to detect specific ventilation patterns, such as potentially harmful hyper-ventilation events [44]. Notice that the RNN in this study was trained using all data available, with no further considerations. This may have caused performances to be lower for minority ventilation patterns, and could be potentially alleviated using oversampling or data augmentation techniques. In the context of ongoing studies on ventilation effectiveness [22], the algorithm could also be used to provide TI amplitude measures as air volume surrogates. Given its ability to detect low amplitude ventilations, diverse enough study data could be selected and reliably analyzed with minimal human effort.

A. Limitations

This study has some limitations. First, all data came from a single device model, the Philips HeartStart MRx. Moreover, the algorithm relied on force and acceleration data, acquired with an external Q-CPR pad that may often not be available and may not be compatible with other manufacturers' devices. Different filters should be explored to broaden the applicability of the solution. Second, the evaluation was carried out with a cross-validation strategy, with all data coming from a single resuscitation site. Thus, the algorithm could be overfitted to specific ventilation patterns, which could be different for other EMS agencies. Further validation should be conducted on an independent dataset, preferably from different agencies and, if possible, different devices. Third, the evaluation was also limited to episode intervals with a capnogram readable enough to annotate the ground truth ventilations. Although we observed no correlation between end-tidal CO_2 levels in the capnogram and ventilation amplitudes in the TI (Pearson's R of -0.11), other relationships may have existed between the two signals that resulted in a selection bias. And fourth, the algorithm was designed for continuous chest compression CPR. In early resuscitation stages, however, 30:2 CPR is usually practiced, where ventilations are delivered during pauses. Other solutions could be more appropriate for this scenario [20].

VI. Conclusion

This study introduces a novel algorithm for the detection of ventilations in TI during continuous chest compression CPR in OHCA. The algorithm improves on previous solutions through the use of an enhanced filtering of compression artifacts and a recurrent neural network to leverage on signal context. The study also introduces a quality control solution to anticipate TI segments where ventilation detection could be defective. The algorithm could be used to facilitate research on ventilation effectiveness, and potentially integrated in resuscitation equipment to provide real-time feedback on ventilation.

Appendix I
Implementation of Literature Solutions

In this section we highlight the details of our implementation of the algorithms proposed by others. Missing information might have conditioned our implementation and thus resulted in suboptimal performance metrics.

The original MC-RAMP filter by Risdal et al. [23] used, on top of force and acceleration, the ECG common-mode channel, which was not available in the study dataset and was therefore ignored. Moreover, the training of the algorithm used OI and OE annotations taken directly on the filtered TI, which had to be derived from the capnogram: First, the largest TI local maxima, with minimum separation of 1.5 s, were located, and annotated as OE when they fell within 0.5 s of a capnogram t_{OE}. Then, for each OE, an OI was searched within 1 s of the corresponding t_{OI}. For each possible OI point, straight lines were fitted to the 0.5 s TI sections before and after the point; the OI was finally selected as the point with largest positive angle change between both lines. A window of 1.4 s and five hidden nodes were used, as found optimal in the original study. Given possible inaccuracies in OI and OE points, a duration of 0.2–5 s was considered; this was specified in the original work for a rule-based solution of reference. The decision thresholds on minimum amplitude and segmentation outputs ($g_{\mathrm{OI}}, g_{\mathrm{OE}}$) were optimized through a grid search to maximize the median per-segment F_1.

The study by Edelson et al. [16] missed key information on both static and adaptive filters, so the same TI preprocessing as in Risdal et al. was applied. No procedural definition was given for the potential start/end times of a ventilation, so they were defaulted to local minima. No validity ranges were given either for the expiration duration and the expiration/inflation ratio, so no related checks were performed.

Finally, Alonso et al. [26] introduced a dynamic amplitude threshold, based on information from previous detections. Amplitude measures were saved and applied between consecutive segments from the same patient, but the average case duration was about five times shorter than in the original study. This may have prevented a successful long run adaptation.

Acknowledgment

We express our deepest gratitude to late Prof. Unai Irusta for his contribution in the conceptualization and early development of this study. We also thank Andoni Elola for his advice on signal filtering and time-series classification.

References

[1] C. W. Tsao et al., "Heart disease and stroke statistics-2022 update: A report from the american heart association," *Circulation*, vol. 145, pp. e153–e639, Feb. 2022.
[2] G. Ritter et al., "The effect of bystander CPR on survival of out-of-hospital cardiac arrest victims," *Amer. heart J.*, vol. 110, pp. 932–937, Nov. 1985.
[3] J. Christenson et al., "Chest compression fraction determines survival in patients with out-of-hospital ventricular fibrillation," *Circulation*, vol. 120, pp. 1241–1247, Sep. 2009.
[4] A. R. Panchal et al., "Part 3: Adult basic and advanced life support: 2020 american heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care," *Circulation*, vol. 142, pp. S366–S468, Oct. 2020.

[5] T. M. Olasveengen et al., "European resuscitation council guidelines 2021: Basic life support," *Resuscitation*, vol. 161, pp. 98–114, Apr. 2021.

[6] J. Soar et al., "European resuscitation council guidelines 2021: Adult advanced life support," *Resuscitation*, vol. 161, pp. 115–151, Apr. 2021.

[7] S. O. Aase and H. Myklebust, "Compression depth estimation for CPR quality assessment using DSP on accelerometer signals," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 3, pp. 263–268, Mar. 2002.

[8] U. Ayala et al., "Automatic detection of chest compressions for the assessment of CPR-quality parameters," *Resuscitation*, vol. 85, pp. 957–963, Jul. 2014.

[9] A. Elola et al., "Finger photoplethysmography to monitor chest compression rate during out-of-hospital cardiac arrest," in *Proc. Comput. Cardiol.*, 2018, pp. 1–4.

[10] J. Coult, J. Blackwood, T. D. Rea, P. J. Kudenchuk, and H. Kwok, "A method to detect presence of chest compressions during resuscitation using transthoracic impedance," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 3, pp. 768–774, Mar. 2020.

[11] D. Hostler et al., "Effect of real-time feedback during cardiopulmonary resuscitation outside hospital: Prospective, cluster-randomised trial," *Brit. Med. J. (Clin. Res. ed.)*, vol. 342, Feb. 2011, Art. no. d512.

[12] X. Jaureguibeitia et al., "Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets," *Resuscitation*, vol. 168, pp. 44–51, Nov. 2021.

[13] A. H. Idris, "The importance of measuring ventilation during resuscitation," *Resuscitation*, vol. 177, pp. 41–42, Jul. 2022.

[14] V. N. Mosesso, "Ventilation during cardiopulmonary resuscitation-only mostly dead!," *Resuscitation*, vol. 141, pp. 200–201, Aug. 2019.

[15] M. R. Neth, A. Idris, J. McMullan, J. L. Benoit, and M. R. Daya, "A review of ventilation in adult out-of-hospital cardiac arrest," *J. Amer. College Emerg. Physicians open*, vol. 1, pp. 190–201, Jun. 2020.

[16] D. P. Edelson et al., "Capnography and chest-wall impedance algorithms for ventilation detection during cardiopulmonary resuscitation," *Resuscitation*, vol. 81, pp. 317–322, Mar. 2010.

[17] E. Aramendi et al., "Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation," *Resuscitation*, vol. 110, pp. 162–168, Jan. 2017.

[18] J. J. Freundlich and J. C. Erickson, "Electrical impedance pneumography for simple nonrestrictive continuous monitoring of respiratory rate, rhythm and tidal volume for surgical patients," *Chest*, vol. 65, pp. 181–184, Feb. 1974.

[19] D. Blanco-Almazán, W. Groenendaal, F. Catthoor, and R. Jané, "Chest movement and respiratory volume both contribute to thoracic bioimpedance during loaded breathing," *Sci. Rep.*, vol. 9, Dec. 2019, Art. no. 20232.

[20] E. Aramendi et al., "A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation," *Resuscitation*, vol. 132, pp. 41–46, Nov. 2018.

[21] H. Losert et al., "Thoracic impedance changes measured via defibrillator pads can monitor ventilation in critically ill patients and during cardiopulmonary resuscitation," *Crit. Care Med.*, vol. 34, pp. 2399–2405, Sep. 2006.

[22] M. P. Chang et al., "Association of ventilation with outcomes from out-of-hospital cardiac arrest," *Resuscitation*, vol. 141, pp. 174–181, Aug. 2019.

[23] M. Risdal, S. O. Aase, M. Stavland, and T. Eftestol, "Impedance-based ventilation detection during cardiopulmonary resuscitation," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2237–2245, Dec. 2007.

[24] S. Ansari, K. R. Ward, and K. Najarian, "Motion artifact suppression in impedance pneumography signal for portable monitoring of respiration: An adaptive approach," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 2, pp. 387–398, Mar. 2017.

[25] P. H. Charlton et al., "An impedance pneumography signal quality index: Design, assessment and application to respiratory rate monitoring," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102339.

[26] E. Alonso et al., "Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics," *Resuscitation*, vol. 88, pp. 28–34, Mar. 2015.

[27] X. Jaureguibeitia et al., "Automatic detection of ventilations during mechanical cardiopulmonary resuscitation," *IEEE J. Biomed. Health Inform.*, vol. 24, pp. 2580–2588, Sep. 2020.

[28] X. Jaureguibeitia, U. Irusta, E. Aramendi, H. Wang, and A. Idris, "An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation," in *Proc. Comput. Cardiol.*, 2020, pp. 1–4.

[29] H. E. Wang et al., "Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: A randomized clinical trial," *JAMA*, vol. 320, pp. 769–778, Aug. 2018.

[30] J. Kramer-Johansen, D. P. Edelson, H. Losert, K. Köhler, and B. S. Abella, "Uniform reporting of measured quality of cardiopulmonary resuscitation (CPR)," *Resuscitation*, vol. 74, pp. 406–417, Sep. 2007.

[31] Z. Zhao, S. Särkkä, and A. B. Rad, "Spectro-temporal ECG analysis for atrial fibrillation detection," in *Proc. IEEE 28Th Int. Workshop Mach. Learn. Signal Process.*, 2018, pp. 1–6.

[32] A. Elola, E. Aramendi, U. Irusta, P. O. Berve, and L. Wik, "Multimodal algorithms for the classification of circulation states during out-of-hospital cardiac arrest," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 6, pp. 1913–1922, Jun. 2021.

[33] S. Haykin, *Kalman Filtering and Neural Networks*. Hoboken, NJ, USA: Wiley, 2004, vol. 47.

[34] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," in *Proc. 8th Workshop Syntax, Semantics Struct. Stat. Transl.*, 2014, pp. 103–111.

[35] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. Fourth Int. Conf. 3D Vis.*, 2016, pp. 565–571.

[36] M. Abadi et al., "Tensorflow: A system for large-scale machine learning," in *Proc 12th USENIX Conf. Oper. Syst. Des. Implement*, 2016, pp. 272–283.

[37] A. Rozo et al., "Data augmentation and transfer learning for data quality assessment in respiratory monitoring," *Front. Bioeng. Biotechnol.*, vol. 10, 2022, Art. no. 806761.

[38] M. Nardelli et al., "Reliability of lagged poincaré plot parameters in ultrashort heart rate variability series: Application on affective sounds," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 3, pp. 741–749, May 2018.

[39] U. Irusta, J. Ruiz, S. R. de Gauna, T. EftestØl, and J. Kramer-Johansen, "A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1052–1062, Apr. 2009.

[40] I. Isasi et al., "Automatic cardiac rhythm classification with concurrent manual chest compressions," *IEEE Access*, vol. 7, pp. 115147–115159, 2019.

[41] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[42] J. A. Penketh et al., "Airway management during in-hospital cardiac arrest: An international, multicentre, retrospective, observational cohort study," *Resuscitation*, vol. 153, pp. 143–148, Aug. 2020.

[43] M. M. J. Nassal et al., "Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the pragmatic airway resuscitation trial," *Resuscitation*, vol. 168, pp. 58–64, Nov. 2021.

[44] T. P. Aufderheide and K. G. Lurie, "Death by hyperventilation: A common and life-threatening problem during cardiopulmonary resuscitation," *Crit. Care Med.*, vol. 32, pp. S345–S351, Sep. 2004.

Supplementary Materials
Heuristic Ruleset for Fluctuation Detection

Xabier Jaureguibeitia, Elisabete Aramendi, *Member, IEEE*, Henry E. Wang, and Ahamed H. Idris

This section describes the procedure to identify fluctuations in the filtered TI signal $s_v(t)$. First, the largest local maxima $t_p^{(i)}$, with minimum separation of 1.5 s, were located. Then, starting from the largest maxima and in decreasing order, the starting and ending points, $t_s^{(i)}$ and $t_e^{(i)}$, of the corresponding fluctuation were identified. In the following, the procedure to identify $t_s^{(i)}$ is detailed, being analogous for $t_e^{(i)}$.

- An inflation duration between 0.45 s and 5.5 s was considered, setting initially the start of the fluctuation $t_s^{(i)}$ at $t_p^{(i)} - 5.5\,\text{s}$. From there, it was advanced to the last time-point before $t_p^{(i)}$, if any, for which $s_v(t) \geq s_v(t_p^{(i)})$, and then advanced again, this time to the global minimum of $s_v(t)$ up to $t_p^{(i)}$.
- Let $r(t)$ be a line segment connecting $s_v(t)$ between $t_s^{(i)}$ and $t_p^{(i)}$. The start of the fluctuation was then updated to:

$$t_s^{(i)} \leftarrow \arg\max \{r(t) - \alpha s_v(t)\} \quad t_s^{(i)} \leq t \leq t_p^{(i)} \tag{1}$$

  As shown in Fig. 1, a value of $\alpha = 2$ was able to skip flat signal plateaus and naturally split different fluctuations. The procedure was also fairly insensitive to minor local minima, which may act as confounders for other alternative approaches. However, as shown in Fig. 1c, this may also result in failing to split ventilations, specially if close to each other and a baseline drift exists.
- In case other fluctuation peaks $t_p^{(k)}$ were contained within the interval $[t_s^{(i)}, t_p^{(i)}]$, a final rule was applied to prevent merging ventilations. Starting from the furthest secondary peak $t_p^{(k)}$, the global minimum of $s_v(t)$ between $t_p^{(k)}$ and $t_p^{(i)}$ was first identified at instant $t_{min}^{(k)}$. Then, the start of the main fluctuation $t_s^{(i)}$ was updated to $t_{min}^{(k)}$ if both the following were met:

$$Z_{down}^{(k)} \geq C_1\, Z_{up}^{(k)} \tag{2}$$

$$Z_{down}^{(k)} \geq C_2\, Z_{up}^{(i)} \tag{3}$$

  where, as shown in Fig1c,

$$Z_{down}^{(k)} = s_v(t_p^{(k)}) - s_v(t_{min}^{(k)})$$
$$Z_{up}^{(k)} = s_v(t_p^{(k)}) - s_v(t_s^{(i)})$$
$$Z_{up}^{(i)} = s_v(t_p^{(i)}) - s_v(t_{min}^{(k)})$$

  Values of $C_1 = C_2 = 0.35$ were considered after some preliminary experiments. This process was repeated for all peaks $t_p^{(k)}$ within $[t_s^{(i)}, t_p^{(i)}]$, from furthest from $t_p^{(i)}$ to closest, using each time the updated value of $t_s^{(i)}$.

If, at any point, $t_s^{(i)}$ was updated beyond the minimum inflation duration, $t_p^{(i)} - 0.45\,\text{s}$, the process was aborted and $t_p^{(i)}$ was discarded as a possible ventilation peak. Otherwise, an analogous procedure was followed to locate the end of the fluctuation $t_e^{(i)}$. If both $t_s^{(i)}$ and $t_e^{(i)}$ were located, the fluctuation was confirmed as a potential ventilation, characterized, and forwarded to the classification stage. Any secondary peak $t_p^{(k)}$ within $[t_s^{(i)}, t_e^{(i)}]$ was automatically discarded.

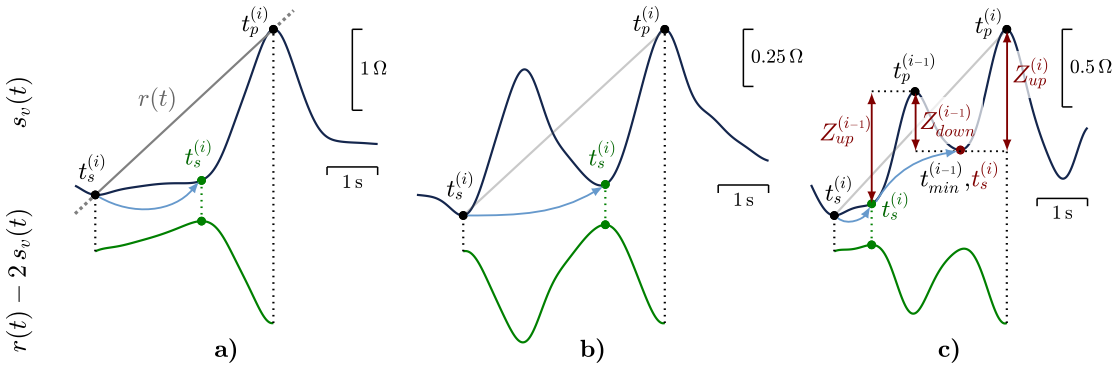


Fig. 1. Examples on determining the start of a fluctuation t_s. The point is first updated (t_s updates in light blue) from the global minimum of $s_v(t)$ preceding t_p to the maximum of $r(t) - 2\,s_v(t)$ (green), where $r(t)$ is a straight line segment connecting the initial t_s and t_p (gray). This procedure was able to: a) skip flat signal plateaus, and b) naturally separate fluctuations. As shown in panel c, a last rule based on relative inflation and deflation amplitudes (red) was included for those cases in which the split of ventilations was not achieved.

A.1.4 Second conference paper

	Publication in international conference
Reference	Xabier Jaureguibeitia, Unai Irusta, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris, "An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation", *Computing in Cardiology, 2020*, pp. 1-4.
Quality indices	• **Type of publication**: International conference in SJR • **Impact factor**: 0.257

An Impedance-based Algorithm to Detect Ventilations During Cardiopulmonary Resuscitation

X Jaureguibeitia[1], U Irusta[1], E Aramendi[1], HE Wang[2], AH Idris[3]

[1] University of the Basque Country (UPV/EHU), Bilbao, Spain
[2] University of Texas Health Science Center, Houston (Texas), USA
[3] University of Texas Southwestern Medical Center, Dallas (Texas), USA

Abstract

Cardiopulmonary resuscitation (CPR) is a core therapy to treat out-of-hospital cardiac arrest (OHCA). Thoracic impedance (TI) can be used to assess ventilations during CPR, but the signal is also affected by chest compression (CC) artifacts. This study presents a method for TI-based ventilation detection during concurrent manual CCs.

Data from 152 OHCA patients were analyzed. A total of 423 TI segments of at least 60 s during ongoing CCs were extracted. True ventilations were annotated using the capnogram. The final dataset comprised 1210 min of TI recordings and 9665 ground truth ventilations.

A three-stage detection algorithm was developed. First, the TI signal was filtered to obtain ventilation waveforms, including a least mean squares filter to remove artifacts due to CCs. Potential ventilations were then identified with a heuristic detector and characterized by a set of 16 features. These were finally fed to a random forest classifier to discriminate between true ventilations and false positives.

Patients were split into 100 distinct training (70%) and test (30%) partitions. The median (interquartile range) sensitivity, PPV and F-score were 83.9 (70.2-91.2) %, 86.1 (75.0-93.3) % and 84.3 (72.1-91.4) %. Our method would allow feedback on ventilation rates as well as surrogate measures of insufflated air volume during CPR.

1. Introduction

High-quality cardiopulmonary resuscitation (CPR) is a key element in the treatment of out-of-hospital cardiac arrest (OHCA). Chest compressions (CC) and ventilations induce an artificial flow of oxygenated blood in the patient which delays the progress of ischemia. This buys time for other therapies to be applied and greatly improves the odds of survival [1]. Resuscitation guidelines recommend CCs to be delivered at rates of 100 - 120 min^{-1} [1]. When CCs are performed without interruption, ventilations should be delivered at a rate of about 10 min^{-1}, avoiding hyperventilation [2]. Adherence to these recommendations improves when rescuers are provided with real-time feedback on CPR therapy [3]. Many solutions have been implemented to enhance CC quality, from integrated accelerometers to monitor rates and depths, to automated devices that compress the chest mechanically. In contrast, few solutions exist to provide feedback on ventilation quality.

Capnography, which monitors the concentration of CO_2 in expired gases, is the recommended method to assess ventilations, but requires advanced airway management and is therefore rarely available early during resuscitation. Thoracic impedance (TI), a measure of body resistance to current flow, fluctuates with air volume changes in the lungs and can also be used to monitor ventilations. TI is typically available much earlier, as it is recorded by most defibrillation equipment concurrently with the ECG. Moreover, the amplitudes of the ventilation waveforms in the TI can be taken as surrogate measures of insufflated air volumes and can be associated with survival outcomes [4].

Detecting ventilations in the TI is challenging even in intervals free of CCs [5]. TI fluctuations vary significantly between cases, and the signal itself is sensitive to noise sources like patient or cable movements. Moreover, CCs produce large artifacts that have to be removed to detect ventilations. State-of-the-art solutions for this scenario make use of adaptive filters to suppress CC artifacts [6]. The present work was built on a previous machine learning solution for ventilation detection during mechanical CCs [7]. The challenge was in adapting those methods to manual CPR, in which CC patterns are much more variable due to rescuer and patient diversity.

2. Data materials

The study dataset included data from 152 OHCA cases treated by the emergency medical services of the Dallas - Fort Worth area and enrolled in the Pragmatic Airway Resuscitation Trial [8]. All cases were treated and recorded with a Philips MRx monitor-defibrillator (Philips Medical Systems, Andover, MA), which acquires the TI signal with

a 200 Hz sampling frequency and a resolution of 0.74 mΩ. The electronic recordings were converted to a MATLAB (MathWorks Inc., Natick, MA) format, and explored to locate TI segments suitable for the study design. Segments of at least 60 s with ongoing CCs were considered. Intervals presenting abrupt TI excursions or CC pauses longer than 20 s were discarded. Only segments with synchronous recordings of compression depth (CD) and capnography were included. CC instants were annotated in the CD and used as reference to filter out the CC artifact. Ventilations were annotated in the capnogram and set as ground truth to develop the detection algorithm. Time delays between the capnogram and the TI were corrected. Figure 1 shows an example of the signals involved in the annotation process. A total of 423 TI segments were extracted, comprising 1210 min of signal recordings and 9665 ground truth ventilations (97.7% during CCs).

3. Methods

A three-stage method was designed to detect ventilations. First, the TI was preprocessed to filter out CC artifacts. A peak detection procedure was then applied to identify potential ventilations in the resulting $s_v(n)$ signal (see Fig 1c). The waveform fluctuation around each peak was delimited and characterized by a set of 16 features. Finally, a random forest classifier was used to discriminate between true ventilations and the many false positives output by the previous stage.

3.1. Signal preprocessing

The impedance signal was first downsampled to 50 Hz. A 0.06 Hz cutoff high-pass IIR filter was then applied to remove the DC offset, followed by a moving average filter ($L = 50$) to remove high-frequency components, including most of the CC artifact. The resulting signal $s(n)$ is assumed to be given by $s(n) = s_v(n) + s_{cc}(n)$, where $s_v(n)$ contains the desired ventilation information and $s_{cc}(n)$ represents the residual of the CC artifact. A Least Mean Squares (LMS) adaptive filter, following the model proposed by Irusta et al [9], was applied to remove CC artifacts. The CC component was approximated by a quasi-periodic signal of the form:

$$\hat{s}_{cc}(n) = A(n) \sum_{k=1}^{N} [a_k(n)\cos(k\phi(n)) + b_k(n)\sin(k\phi(n))]$$

where N is the number of harmonics. The phase term $\phi(n)$ is built by producing a 2π phase shift between each pair of consecutive compressions. $A(n)$ differentiates intervals with ($A = 1$) and without ($A = 0$) CCs, with a pause in compressions being defined as an interval longer than 1.5 s in absence of CCs. The time varying in-phase and quadrature amplitude coefficients, a_k and b_k, were then estimated using the LMS algorithm for harmonic interference suppression [10]. In this setting, the error signal $e(n) = s(n) - \hat{s}_{cc}(n)$ is an estimate of the desired $s_v(n)$ component. Values of $N = 3$ and step-size $\mu = 0.05$ were selected according to some initial experiments.

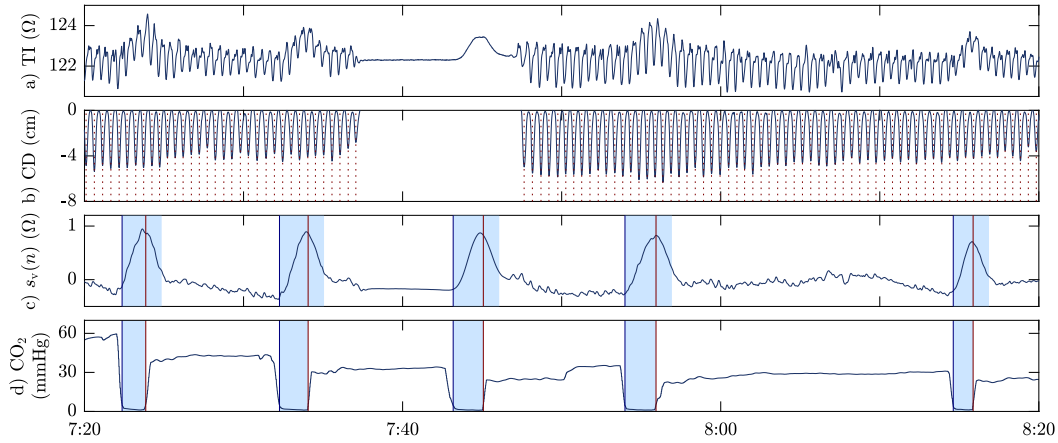


Figure 1: An example of the signals involved in the detection process. (a) The raw TI signal, as stored in defibrillator data. (b) The CD signal, including the annotated CC instants as dotted vertical lines. (c) The ventilation component of the TI signal, obtained through adaptive filtering using the CC instants shown in (b). (d) The time-aligned capnogram signal, including true ventilation annotations as shaded intervals. These intervals, associated to air insufflation phases, are reproduced in (c) and extended by up to 1 s to determine which peaks correspond to true ventilations.

3.2. Ventilation waveform processing

Potential ventilations were detected as the highest local maxima of $s_{\mathrm{v}}(n)$ with a refractory period of 1.5 s (ventilation rate of 40 min^{-1}). For each detected peak position t_{p_i}, the start and end points of the ventilation waveform, t_{s_i} and t_{e_i}, were then estimated. The location of t_{s_i} was first set to the minimum $s_{\mathrm{v}}(n)$ point in the interval $(t_{p_i} - 5.5\,, t_{p_i} - 0.45)$, and then adjusted following a series of heuristic rules. Amplitude relations between local maxima and minima within $(t_{s_i}\,, t_{p_i})$ were evaluated to avoid merging several waveforms into a single detection. An initial slope condition was also applied to prevent t_{s_i} from being set on a near flat signal region, far away from the actual waveform. An analogous procedure was followed to locate t_{e_i} within the interval $(t_{p_i} + 0.45\,, t_{p_i} + 5.5)$. Any peak t_{p_i} for which t_{s_i} or t_{e_i} could not be determined was discarded as potential ventilation.

3.3. Feature extraction

A total of 16 features were used to characterize each potential ventilation waveform. The first four corresponded to the amplitude and duration of the insufflation and exhalation phases, that is the intervals $(t_{s_i}\,, t_{p_i})$ and $(t_{p_i}\,, t_{e_i})$, respectively [7]. The next 10 features, 5 for each phase, were given by the coefficients of a least-squares approximation of the phase waveform in terms of up to fourth order Legendre polynomials. Curve fitting was approached assuming the samples of each phase to be equispaced over a [-1,1] transformed time domain. Finally, the signal excursions $\Delta_{90} = P_{95}(s_{\mathrm{v}}) - P_{5}(s_{\mathrm{v}})$ and $\Delta_{95} = P_{97.5}(s_{\mathrm{v}}) - P_{2.5}(s_{\mathrm{v}})$, which were computed in an up to 1 min interval around the peak position. P_{x} represents the x-th amplitude percentile. This was deemed possible as windows of about 1 min are typically analyzed to report ventilation rates, and allowed the classifier to adapt to the large ventilation amplitude differences between cases.

3.4. Classification

The waveform detection stage was conceived to miss few true ventilations, so it produced a large number of false positives. A random forest classifier [11] was employed to discriminate between both cases. In order to train the classifier and test the algorithm, each potential ventilation i reported by the detection stage was labeled as either true ventilation ($y_i = 1$) or false positive ($y_i = 0$) according to the ground truth annotations in the capnogram. Those waveforms with peak position t_{p_i} within a ground truth interval (shaded regions in Figure 1c) were labeled as $y_i = 1$. When several detections were reported within the same interval, only the one closest to the expiration onset (red lines in Figure 1c) was labeled as true ventilation. The outputs of the detection stage formed a set of instance-label pairs $\{\boldsymbol{x_i}, y_i\} \in \mathbb{R}^{K\times(0,1)}$, where $K = 16$ is the number of features considered. A 100 tree ensemble size and a minimum leaf-size of 5 were used as classifier hyperparameters. Data was patient-wise weighted to avoid biasing results towards longer cases with a larger amount of ventilations.

3.5. Model evaluation

Data were randomly partitioned patient-wise into train (70%) and test (30%) sets. Due to large differences in the number of ventilations, the assignment was carried out in a quasi-stratified way, forcing the train set to contain between 65% and 75% of the ground truth ventilations. Performance was assessed in terms of sensitivity (Se, proportion of correctly identified true ventilations), positive predictive value (PPV, proportion of true ventilations among reported detections) and F-score (F_1, harmonic mean of Se and PPV). To avoid data partition bias in the results, 100 different train/test splits were generated and performance metrics were averaged patient-wise.

4. Results

The waveform detector overestimated ventilations with a median (interquartile range, IQR) patient-wise Se of 94.1 (89.0-98.4) % and a PPV of 56.0 (44.9-66.7) %. The classification stage corrected most of the false positives, resulting in Se, PPV and F_1 scores of 83.9 (70.2-91.2) %, 86.1 (75.0-93.3) % and 84.3 (72.1-91.4) % respectively. The patient-wise distribution of performance metrics is shown in Figure 2.

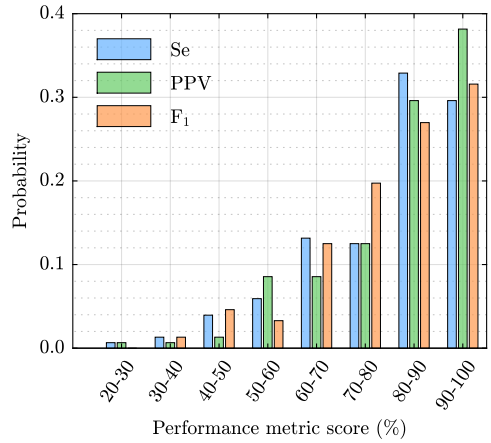


Figure 2: Patient-wise performance metric scores for the full algorithm, as probability distribution.

Motion artifacts caused by rescuers during treatment were identified as the major error source. As shown in Figure 3, these may produce many false positives, and distort or even completely mask the actual ventilations.

5. Conclusions

An algorithm for impedance-based ventilation detection during concurrent manual CCs was demonstrated. With median patient-wise performance metric scores of about 85%, the method could allow accurate feedback on ventilation rates and possibly on insufflated air volumes in the majority of cases. Performances could be improved in the future by employing more complex filters and feature selection techniques. For some cases, however, the algorithm presented a severely degraded performance due to motion artifacts, a phenomena not present during mechanical CPR [7], presumably due to the compression device being tightly attached to the patient. Devising a signal quality index to identify these situations may improve performance in the future. This would allow to anticipate problematic cases, avoiding inaccurate feedback, but more importantly, preventing the classifier from being fed with corrupted training examples which may severely hinder the overall performance.

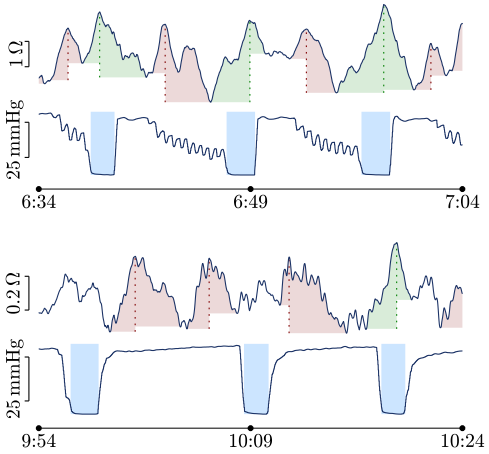


Figure 3: Examples of TI segments corrupted by motion artifacts. The detections output by the algorithm are displayed in in green for true ventilations and in red for false positives. For large ventilations (top), artifacts may result in additional detections and distorted waveforms. During shallow ventilations (bottom), artifacts might be dominant and result in extremely degraded performances.

Acknowledgments

This work was supported by the Spanish Ministerio de Ciencia, Innovacion y Universidades through grant RTI2018-101475-BI00, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), and by the Basque Government through grants IT1229-19 and PRE-2019-1-0209.

References

[1] G. D. Perkins, A. J. Handley, R. W. Koster *et al.*, "European resuscitation council guidelines for resuscitation 2015: Section 2. adult basic life support and automated external defibrillation." *Resuscitation*, vol. 95, pp. 81–99, Oct. 2015.

[2] J. Soar, J. P. Nolan, B. W. Böttiger *et al.*, "European resuscitation council guidelines for resuscitation 2015: Section 3. adult advanced life support." *Resuscitation*, vol. 95, pp. 100–147, Oct. 2015.

[3] D. Hostler, S. Everson-Stewart, T. D. Rea *et al.*, "Effect of real-time feedback during cardiopulmonary resuscitation outside hospital: prospective, cluster-randomised trial." *BMJ (Clinical research ed.)*, vol. 342, p. d512, Feb. 2011.

[4] M. P. Chang, Y. Lu, B. Leroux *et al.*, "Association of ventilation with outcomes from out-of-hospital cardiac arrest." *Resuscitation*, vol. 141, pp. 174–181, Aug. 2019.

[5] E. Aramendi, Y. Lu, M. P. Chang *et al.*, "A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation." *Resuscitation*, vol. 132, pp. 41–46, Nov. 2018.

[6] M. Risdal, S. O. Aase, M. Stavland *et al.*, "Impedance-based ventilation detection during cardiopulmonary resuscitation." *IEEE Trans Biomed Eng*, vol. 54, pp. 2237–2245, Dec. 2007.

[7] X. Jaureguibeitia, U. Irusta, E. Aramendi *et al.*, "Automatic detection of ventilations during mechanical cardiopulmonary resuscitation," *IEEE JBHI*, 2020.

[8] H. E. Wang, R. H. Schmicker, M. R. Daya *et al.*, "Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial," *JAMA*, vol. 320, no. 8, pp. 769–778, 2018.

[9] U. Irusta, J. Ruiz, S. R. de Gauna *et al.*, "A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions." *IEEE Trans Biomed Eng*, vol. 56, pp. 1052–1062, Apr. 2009.

[10] Y. Xiao and Y. Tadokoro, "Lms-based notch filter for the estimation of sinusoidal signals in noise," *Signal Processing*, vol. 46, no. 2, pp. 223–231, 1995.

[11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

Address for correspondence:

Xabier Jaureguibeitia
Plaza Ingeniero Torres Quevedo 1. 48013 - Bilbao, Spain
xabier.jaureguibeitia@ehu.eus

A.1.5 Third conference paper

	Publication in national conference
Reference	Xabier Jaureguibeitia, Elisabete Aramendi, Henry E. Wang, Ahamed H. Idris, "Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar", *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica, 2022*, pp. 175-178.
Quality indices	• **Type of publication**: National conference (not indexed)

Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar

X. Jaureguibeitia Lara[1], E. Aramendi Ecenarro[1], H.E. Wang[2], A.H. Idris[3]

[1] BioRes, Dpto. de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {xabier.jaureguibeitia, elisabete.aramendi}@ehu.eus

[2] Wexner Medical Center, Ohio State University, Columbus, Ohio, EEUU, Henry.Wang@osumc.edu

[3] University of Texas Southwestern Medical Center, Dallas, Texas, EEUU, ahamed.idris@utsouthwestern.edu

Resumen

Una resucitación cardiopulmonar (RCP) de calidad es esencial para mejorar la supervivencia a la parada cardiorrespiratoria extra-hospitalaria (PCREH), pero poco se sabe sobre la terapia óptima de ventilación. La impedancia torácica (IT), medida a través de los parches de desfibrilación, permite monitorizar las ventilaciones de cara al análisis retrospectivo de grandes registros de PCREH, pero presenta artefactos debidos a las compresiones torácicas y al movimiento de los electrodos. El objetivo de este trabajo fue evaluar si las arquitecturas de aprendizaje profundo podrían ofrecer ventajas respecto de los métodos disponibles para la detección de ventilaciones en IT con compresiones concurrentes. Se analizó un total de 367 episodios de PCREH, y se anotaron 20724 ventilaciones de referencia en capnografía. Se implementó una arquitectura de segmentación U-Net a la que se alimentó con segmentos no filtrados de IT de un minuto de duración. El desempeño de la red se evaluó usando una estrategia de validación cruzada. Las métricas de desempeño globales fueron de un 79,8% de sensibilidad, un 83,4% de valor predictivo positivo, y un 81,6% de valor F_1, superando a soluciones anteriores propuestas en la literatura, aunque con una variabilidad importante entre casos/pacientes. Para las ventilaciones correctamente detectadas, los instantes de inspiración y expiración anotados en el capnograma se reprodujeron con un error medio (desviación típica) de 0,02s (0,51s) y –0,002s (0,43s), respectivamente.

1. Motivación

La parada cardiorrespiratoria extra-hospitalaria (PCREH) constituye un problema de salud pública de primer orden, con en torno a 56 casos atendidos por 100.000 habitantes y año, y una tasa de supervivencia inferior al 10% [1]. Un paciente en PCREH pierde, generalmente de forma súbita, la circulación y respiración espontáneas, lo que desemboca en la muerte en ausencia de tratamiento. La resucitación cardiopulmonar (RCP), consistente en compresiones torácicas y ventilaciones, suministra un flujo mínimo de sangre y oxígeno al paciente, y es clave para mejorar las opciones de supervivencia a la PCREH [2]. En ese sentido, se han destinado importantes esfuerzos a mejorar la calidad de la RCP: los parámetros óptimos de compresión han sido estudiados extensamente [3], y se han desarrollado también nuevas tecnologías tales como dispositivos autónomos de compresión mecánica. La mayoría de avances han estado, sin embargo, centrados en las compresiones torácicas; la tecnología actual para monitorizar la ventilación durante la PCREH es limitada, y poco se sabe aún sobre su aplicación óptima [4]. Son necesarias nuevas técnicas y tecnologías que permitan un mejor análisis y una mayor comprensión del efecto de la ventilación durante la PCREH.

La capnografía, que monitoriza la concentración de CO_2 en los gases expirados, es la técnica habitual para evaluar la ventilación en la PCREH [5]. Sin embargo, ésta no suele estar disponible hasta fases avanzadas de la resucitación, tras la inserción de una vía aérea. La impedancia torácica (IT), que mide la resistencia del cuerpo al paso de corriente, es sensible a cambios de volumen de aire en los pulmones, y supone una alternativa para la detección de ventilaciones. La IT se adquiere junto al electrocardiograma a través de los parches de desfibrilación, siendo así una de las primeras señales en estar disponible durante la PCREH. Además, la amplitud de las formas de ondas de ventilación se relaciona con el volumen de aire insuflado [6], una información no disponible mediante las técnicas habituales de capnografía extra-hospitalaria. La detección de ventilaciones en IT no está, sin embargo, exenta de problemáticas: las formas de onda de ventilación son heterogéneas [7], y la propia señal es susceptible de ser contaminada por diversas fuentes de ruido, como el movimiento de electrodos [8]. Más aún, la IT es sensible a las compresiones administradas como parte de la RCP, lo que resulta en un artefacto que enmascara las ventilaciones. La supresión de este artefacto requiere el uso de filtros muy exigentes [9], con la consiguiente pérdida de información de ventilación, o de filtros adaptativos en base a señales adicionales [7,10], no siempre disponibles.

En los últimos años, las técnicas de aprendizaje profundo han obtenido resultados prometedores en diversos campos de la medicina, incluyendo la PCREH [11]. En esta línea, el objetivo de este trabajo fue evaluar si dichas técnicas podrían superar a los métodos actuales en la identificación de ventilaciones en IT durante compresiones concurrentes.

2. Materiales

Se analizaron los registros electrónicos de 367 pacientes de PCREH, tratados por los servicios de emergencia de Dallas - Fort Worth (Texas, EEUU) entre 2016 y 2017, y adscritos al ensayo clínico PART (Pragmatic Airway Resuscitation Trial) [12]. Todos los registros fueron adquiridos usando un monitor-desfibrilador HeartStart MRx (Philips Medical Systems, Andover, MA, EUU), e incluían grabaciones de IT (200Hz, resolución de 2,5mΩ) y capnografía, así como de fuerza, aceleración y profundidad de compresión.

ISBN: 978-84-09-45972-8, pags. 175- 178

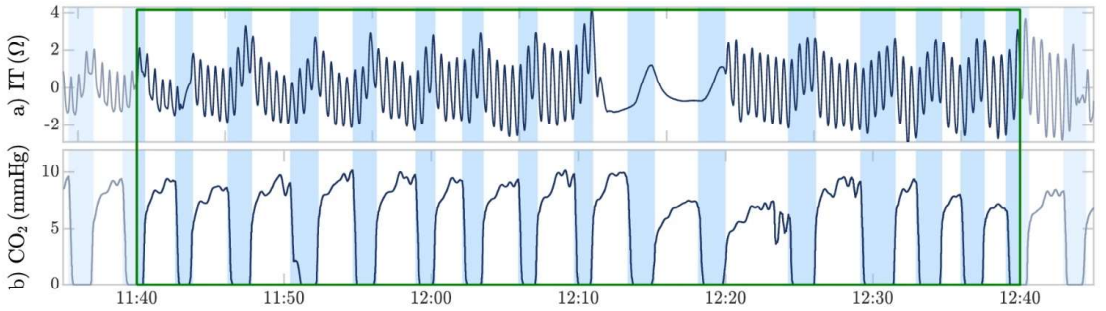


Figura 1. *Segmento de un minuto de impedancia torácica (IT) y su capnograma asociado. Las ventilaciones de referencia, anotadas en la capnografía y correspondientes a la fase de inspiración, se muestran en azul. Ambas señales se alinearon temporalmente haciendo coincidir los flancos de subida del capnograma con los picos de las fluctuaciones de IT.*

Los registros se procesaron y analizaron empleando Matlab (MathWorks Inc., Natick, MA, EEUU), y se seleccionaron intervalos de señal apropiados para el estudio. Se extrajeron intervalos durante compresiones torácicas manuales, y con concurrencia de todas las señales previamente citadas. Se descartaron pausas de compresiones superiores a 20s, así como excursiones abruptas en la IT y otros artefactos dominantes aislados. La capnografía se utilizó para anotar las ventilaciones de referencia para el entrenamiento y la evaluación de la red. Se corrigió también de forma manual el retardo temporal de la capnografía respecto de la IT (ver Figura 1). Las señales de fuerza, aceleración y profundidad de compresión se incluyeron exclusivamente para la implementación de soluciones similares en la literatura [7, 9, 10].

Los intervalos de señal seleccionados se subdividieron en segmentos no superpuestos de un minuto, sobre los que se ejecutaron los diferentes algoritmos. A fin de facilitar la detección de ventilaciones parcialmente incluidas en los segmentos, se incluyó también un relleno no evaluable de aproximadamente 4,5s en cada extremo de los mismos. La base de datos final incluyó un total de 2551 segmentos de un minuto y 20724 ventilaciones de referencia, el 97.1% de las mismas durante compresiones torácicas.

3. Métodos

3.1. Preprocesado de señal

Se apostó por un preprocesado mínimo, evitando el uso de señales adicionales y delegando sobre la red de aprendizaje

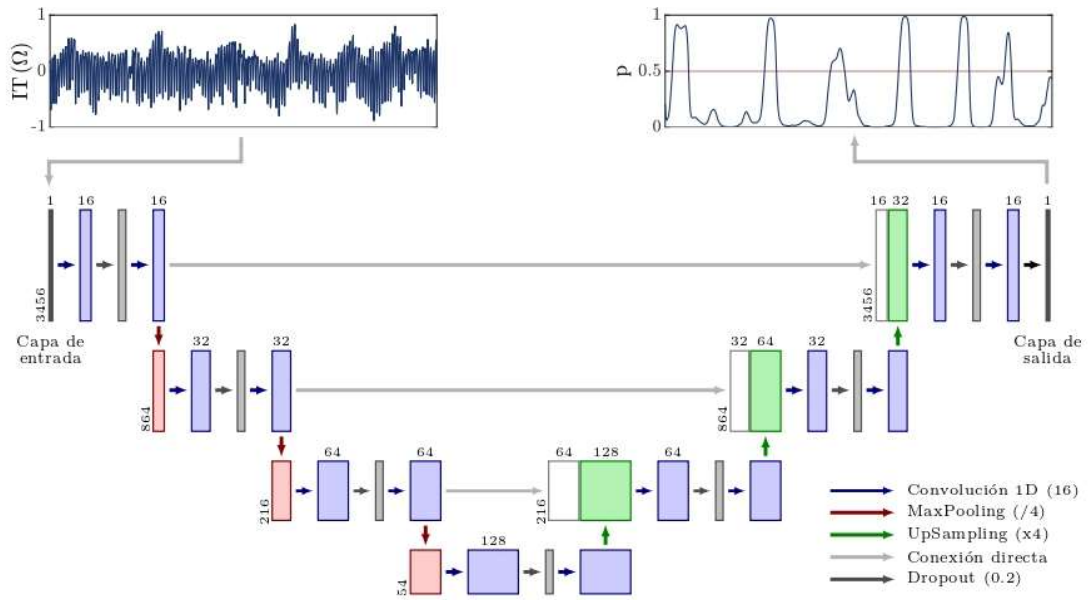


Figura 2. *Arquitectura de la red U-Net implementada. Los valores numéricos sobre cada capa indican el número de filtros o canales a la salida de la misma. Los valores en la parte inferior derecha de cada capa/nivel indican el número de muestras por canal a la salida de la misma. La red se alimenta con segmentos de un minuto de impedancia (IT), y devuelve por cada muestra de señal la probabilidad p de que esta corresponda a una ventilación.*

profundo el tratamiento del artefacto de compresiones. Los segmentos de IT se remuestrearon a 50Hz, y se eliminó su componente continua mediante un filtrado digital paso alto de fase de cero (Butterworth, orden 4, frecuencia de corte de 0,06Hz).

3.2. Arquitectura de aprendizaje profundo

Se escogió una arquitectura de segmentación U-Net [13], adaptada para operar sobre señales unidimensionales. Una red U-Net sigue un esquema codificador-decodificador, en el que se incluyen conexiones directas entre niveles pares de codificación y decodificación. Esto permite reconstruir el nivel de detalle perdido durante la etapa de contracción o codificación, al tiempo que se mantiene su contribución en términos de contexto. La red implementada (ver Figura 2) constó de cuatro niveles, con dos capas de convolución por nivel. A fin de reducir el número de pesos entrenables, se emplearon convoluciones separables, todas con filtros de 16 muestras y función de activación rectificadora lineal (ReLU). Las capas de deconvolución del diseño original se sustituyeron también por capas de sobremuestreo o *upsampling*, sin parámetros entrenables. Dada la limitación de los datos y el considerable tamaño de la red, se incluyó en cada nivel una capa de olvido o *dropout* (20%) para combatir el sobreajuste. La red se alimentó con los segmentos de un minuto de IT, incluyendo 228 muestras (~ 4,5s) de relleno de señal, no evaluables, en cada extremo. La capa de salida consistió en una convolución puntual con activación sigmoide, produciendo una secuencia de valores de probabilidad p con tantas muestras como la entrada.

La optimización de la red se llevó a cabo en 50 épocas, con tandas de 64 segmentos de IT. Se utilizó como función de coste la entropía cruzada binaria, y un optimizador Adam con tasa de aprendizaje de 0,001. La implementación y la optimización de la red se realizaron en Tensorflow 2.8.

3.3. Evaluación como detector

El desempeño de la red de segmentación se evaluó usando una estrategia de validación cruzada de cinco grupos. Los segmentos de IT se asignaron a los distintos grupos en base a paciente, a fin de mantener separados segmentos con similitudes morfológicas. La asignación se realizó también de forma cuasi-estratificada, de modo que cada grupo recogiese un quinto (tolerancia máxima del 5%) de pacientes, segmentos y ventilaciones de referencia.

Dada la salida de la red, se consideró como una ventilación potencial cualquier intervalo continuo con valores $p \geq 0{,}5$. Estas detecciones se evaluaron como verdaderos positivos (VP) o falsos positivos (FP) en base a las anotaciones de referencia realizadas en el capnograma. Se tomó como instante de detección el instante con mayor valor de p dentro de cada intervalo, y se consideró VP si quedaba comprendido dentro de una ventilación de referencia, con un margen de toleracia de 0,5s. Este margen se redujo, de ser necesario, para evitar el solape de las ventilaciones de referencia. En caso de que varias detecciones calificasen para la misma ventilación de referencia, aquella más cercana a su punto medio se consideró VP, y el resto FP.

El desempeñó se evaluó en términos de sensibilidad (Se), valor predictivo positivo (VPP), y valor F_1, dados por:

$$\mathrm{Se} = \frac{\mathrm{VP}}{\mathrm{N_{ref}}} \qquad \mathrm{VPP} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FP}} \qquad \mathrm{F_1} = 2\,\frac{\mathrm{Se}\cdot\mathrm{VPP}}{\mathrm{Se} + \mathrm{VPP}}$$

siendo $\mathrm{N_{ref}}$ el número de ventilaciones de referencia. Las métricas se calcularon para toda la base de datos, así como para cada paciente individual. Las métricas por paciente se caracterizaron en términos de mediana (rango intercuartil). El desempeñó se comparó con el de varias soluciones de la literatura [7, 9, 10], las cuales fueron implementadas en Matlab y evaluadas contra los datos de estudio.

3.4. Evaluación como segmentador

Para cada detección evaluada como correcta, se estableció el comienzo de la inspiración t_{insp} como el instante inmediatamente anterior con salida de la red $p < p_{insp}$, o como el instante con p mínima desde la ventilación precedente, lo que resultase más cercano. El valor p_{insp} se obtuvo usando como referencia las predicciones de la red sobre los datos de entrenamiento. Se escogió p_{insp} tal que se minimizara el error cuadrático medio entre los instantes t_{insp} y el inicio de la referencia en capnografía asociada (flanco de bajada). Se siguió un procedimiento análogo para determinar el valor frontera p_{exp} y establecer los inicios de expiración t_{exp}.

Se evaluaron, sobre las detecciones correctas en los datos de testeo, el error medio (desviación típica) de los instantes t_{insp} y t_{exp}, así como de la duración total de la inspiración, respecto de las referencias en capnografía asociadas.

4. Resultados

Tal como muestra la Tabla 1, la red implementada superó, en desempeño como detector, a las soluciones previas en la literatura, incluidas aquellas que requerían de señales adicionales en la etapa de filtrado [7, 10]. Las soluciones por Risdal et al. y Alonso et al. obtuvieron una mejor Se, pero a costa de reportar un elevado número de FP (VPP bajo), resultando en un valor F_1 muy inferior. Esto se atribuyó principalmente a una falta de contexto para discriminar pequeñas fluctuaciones debidas al ruido de ventilaciones más

Algoritmo	Datos completos			Por paciente		
	Se (%)	VPP (%)	F_1 (%)	Se (%)	VPP (%)	$F_{1\,(\%)}$
U-Net	79.8	83.4	81.6	83.1 (65.3-94.2)	84.9 (71.4-95.0)	82.2 (66.7-93.3)
Risdal et al. [7]	80.8	62.0	70.2	87.3 (71.7-95.8)	65.2 (48.1-82.1)	70.0 (56.3-82.8)
Alonso et al. [9]	89.0	59.1	71.1	92.5 (84.6-97.1)	60.2 (43.4-79.5)	68.6 (56.1-85.5)
Jaureguibeitia et al. [10]	76.3	82.4	79.2	79.7 (58.6-91.3)	83.3 (69.1-93.8)	80.5 (63.9-91.0)

Tabla 1. *Métricas de desempeño de la arquitectura U-Net y de soluciones previas en la literatura.*

prominentes en el mismo segmento. La solución previa por Jaureguibeitia et al., que incluía un algoritmo de aprendizaje máquina y varias características de contexto, ofreció mejores resultados, pero inferiores a la U-Net propuesta pese a requerir de detecciones precisas de los instantes de compresión medidas en la profundidad de compresión.

Resulta importante observar que, con independencia del algoritmo empleado, existió entre pacientes una variabilidad considerable de las métricas desempeño. Algunos segmentos de IT, e incluso registros completos, podrían no ser aptos para la identificación de ventilaciones. En este sentido, el enfoque alternativo ofrecido por el aprendizaje profundo podría servir para analizar relaciones de concordancia con los algoritmos clásicos y desarrollar índices de calidad de señal para la IT [8].

En lo referente a la segmentación, el procedimiento empleado obtuvo distribuciones de error muy centradas, con errores medios marginales, pero desviaciones importantes. Los errores medios para t_{insp} y t_{exp} fueron de de –0,002s (0,43s) y 0,02s (0,51s), respectivamente. El error medio en la duración de la inspiración fue de –0,02s (0,60s).

5. Conclusiones

La red U-Net propuesta mostró resultados prometedores de cara al uso de aprendizaje profundo para la segmentación de ventilaciones en la señal de IT. Se mejoró el desempeño de soluciones previas en la literatura, sin utilizar, además, señales adicionales para la cancelación del artefacto de compresiones. Esto representa una mejora significativa en cuanto a su aplicabilidad a un mayor rango de escenarios. Alternativamente, la red podría también alimentarse con dichas señales para una mejora potencial de los resultados. La red es también susceptible a otras mejoras. Por ejemplo, la frecuencia de muestro de la IT (50Hz) podría reducirse, reduciendo el uso de memoria y permitiendo acortar los filtros para una menor carga computacional. Si bien el uso de la red actual parece limitado al análisis retrospectivo de registros de PCREH, una red suficientemente ligera podría integrarse en el equipamiento de campo para proporcionar realimentación en tiempo real. Una red más ligera sería también menos susceptible a un potencial sobreajuste, algo que podría lograrse también mediante técnicas de aumento de datos, no empleadas en este estudio. Finalmente, cabría explorar diferentes arquitecturas, incluyendo algunas de las múltiples variantes de U-Net, pero también otros enfoques radicalmente diferentes que podrían adaptarse mejor a la segmentación de ventilaciones.

Agradecimientos

Este trabajo ha recibido ayuda financiera del Ministerio de Ciencia, Innovación y Universidades, proyecto RTI2018-101475-BI00, junto con el Fondo Europeo de Desarrollo Regional (FEDER), así como del Gobierno Vasco a través de la subvención a grupos de investigación IT-1229-19 y la beca predoctoral PRE-2021-2-0126, y de la Universidad del País Vasco a través del proyecto COLAB20/01.

Referencias

[1] Gräsner JT, Wnent J, Herlitz J, et al. Survival after out-of-hospital cardiac arrest in Europe – Results of the EuReCa TWO study. *Resuscitation*, vol 148, 2020, pp 218-226.

[2] Cummins RO, Eisenberg MS, Hallstrom AP, et al. Survival of out-of-hospital cardiac arrest with early initiation of cardiopulmonary resuscitation. *The American Journal of Emergency Medicine*, vol 3, 1985, pp 114-119.

[3] Panchal AR, Bartos JA, Cabañas JG, et al. Part 3: Adult Basic and Advanced Life Support: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation*, vol 142, 2020, pp S366-S468.

[4] Neth MR, Idris A, McMullan J, et al. A review of ventilation in adult out-of-hospital cardiac arrest. *Journal of the American College of Emergency Physicians open*, vol 1, 2020, pp 190-201

[5] Aramendi E, Elola A, Alonso E, et al. Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation. *Resuscitation*, vol 110, 2017, pp 162-168.

[6] Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. *Resuscitation*, vol 141, 2019, 174-181.

[7] Risdal M, Aase SO, Stavland M, et al. Impedance-based ventilation detection during cardiopulmonary resuscitation. *IEEE Transactions on Biomedical Engineering*, vol 54, 2007, pp 2237-2245.

[8] Charlton PH, Bonnici T, Tarassenko L, et al. An impedance pneumography signal quality index: Design, assessment and application to respiratory rate monitoring. *Biomedical Signal Processing and Control*, vol 65, 2021.

[9] Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance for measuring cardiopulmonary resuscitation quality metrics. *Resuscitation*, vol 88, 2015, pp 28-34.

[10] Jaureguibeitia X, Irusta U, Aramendi E, et al. An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation. *Proc. Computing in Cardiology,* 2020, pp 1-4.

[11] Jaureguibeitia X, Zubia G, Irusta U, et al. Shock decision algorithms for automated external defibrillators based on convolutional networks. *IEEE Access*, vol 8, 2020.

[12] Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. *JAMA*, vol 320, 2018, pp 769-778.

[13] Ronneberger O, Fischer P, y Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv*, 2015.

A.2 Publications associated to objective 2

A.2.1 First journal paper

	Publication in international journal
Reference	Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Erik Alonso, Tom P. Aufderheide, Robert H. Schmicker, Matthew Hansen, Robert Suchting, Jestin N. Carlson, Ahamed H. Idris, Henry E. Wang, "Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets", *Resuscitation 2021*, vol. 168, pp. 44-51.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (3/32) based on Web of Science Rank 2021 • **Impact factor**: 6.251

RESUSCITATION 168 (2021) 44–51

Available online at ScienceDirect

Resuscitation

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

Clinical paper

Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets

Xabier Jaureguibeitia[a], *Elisabete Aramendi*[a,b,*], *Unai Irusta*[a,b], *Erik Alonso*[c], *Tom P. Aufderheide*[d], *Robert H. Schmicker*[e], *Matthew Hansen*[f], *Robert Suchting*[g], *Jestin N. Carlson*[h,i], *Ahamed H. Idris*[j], *Henry E. Wang*[k]

[a] *Communications Engineering Department, University of the Basque Country UPV/EHU, Bilbao, Spain*
[b] *Biocruces Bizkaia Health Research Institute, Barakaldo, Spain*
[c] *Department of Applied Mathematics, University of the Basque Country UPV/EHU, Bilbao, Spain*
[d] *Department of Emergency Medicine, Medical College of Wisconsin, Milwaukee, WI, United States*
[e] *Clinical Trial Center, Department of Biostatistics, University of Washington, Seattle, WA, United States*
[f] *Department of Emergency Medicine, Oregon Health and Science University, Portland, OR, United States*
[g] *Department of Psychiatry and Behavioral, Sciences University of Texas Health Science Center at Houston, Houston, TX, United States*
[h] *Department of Emergency Medicine, Saint Vincent Hospital, Allegheny Health Network, Erie, PA, United States*
[i] *Department of Emergency Medicine, University of Pittsburgh, Pittsburgh, PA, United States*
[j] *Department of Emergency Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States*
[k] *Department of Emergency Medicine, Ohio State University, Columbus, OH, United States*

Abstract

Background: Out-of-hospital cardiac arrest (OHCA) data debriefing and clinical research often require the retrospective analysis of large datasets containing defibrillator files from dierent vendors and clinical annotations by the emergency medical services.

Aim: To introduce and evaluate a methodology to automatically extract cardiopulmonary resuscitation (CPR) quality data in a uniform and systematic way from OHCA datasets from multiple heterogeneous sources.

Methods: A dataset of 2236 OHCA cases from multiple defibrillator models and manufacturers was analyzed. Chest compressions were automatically identified using the thoracic impedance and compression depth signals. Device event time-stamps and clinical annotations were used to set the start and end of the analysis interval, and to identify periods with spontaneous circulation. A manual audit of the automatic annotations was conducted and used as gold standard. Chest compression fraction (CCF), rate (CCR) and interruption ratio were computed as CPR quality variables. The unsigned error between the automated procedure and the gold standard was calculated.

Results: Full-episode median errors below 2% in CCF, 1 min^{-1} in CCR, and 1.5% in interruption ratio, were measured for all signals and devices. The proportion of cases with large errors (>10% in CCF and interruption ratio, and >10 min^{-1} in CCR) was below 10%. Errors were lower for shorter sub-intervals of interest, like the airway insertion interval.

Conclusions: An automated methodology was validated to accurately compute CPR metrics in large and heterogeneous OHCA datasets. Automated processing of defibrillator files and the associated clinical annotations enables the aggregation and analysis of CPR data from multiple sources.

Keywords: Automated methods, Chest compressions, Debriefing, Quality metrics, Validation

* Corresponding author at: Communications Engineering Department, University of the Basque Country UPV/EHU, Escuela de Ingeniería de Billbao, Plaza Ingeniero Torres Quevedo 1, Bilbao, Spain.

E-mail address: elisabete.aramendi@ehu.eus (E. Aramendi).

https://doi.org/10.1016/j.resuscitation.2021.09.005

Received 7 February 2021; Received in Revised form 1 September 2021; Accepted 3 September 2021

Introduction

Quality cardiopulmonary resuscitation (CPR) is essential for the survival of patients in out-of-hospital cardiac arrest (OHCA). Current and previous resuscitation guidelines have emphasized the importance of good chest compressions (CCs) provided continuously, or with minimal pauses, ensuring adequate rate and depth, and avoiding hyperventilation.[1,2] Many large multicentre research projects and clinical trials focus on CPR analysis of large numbers of OHCA cases. These datasets contain two sources of data. First, clinical annotations, including Utstein template data and ad-hoc annotations by emergency medical services (EMS) providers, typically collected in an external spreadsheet file. And second, defibrillator files with detailed information on prehospital treatment and patient response.[3–9]

Current technology enables the detailed analysis and the systematic information retrieval from defibrillator files, including the ECG and the thoracic impedance (TI) signal, the CPR-pad signals such as acceleration or compression depth (CD), the capnogram, or photoplethysmography and oxygen saturation. Many algorithms based on signal processing and artificial intelligence have been described to measure CPR quality, for compressions[10–14] and ventilations,[15,12,16–18] to determine perfusion states or return of spontaneous circulation (ROSC),[19–24] to predict shock success,[25–27] or to determine the patient's heart rhythm.[28,29] These algorithms use a single biomedical signal or a combination of them to automatically obtain CPR quality variables and patient response data. The integration of these advances in large scale OHCA research projects is still a challenging objective.

Multicentre studies that collect data from multiple EMS agencies may use defibrillators from different vendors, with differences in the format of the information, the software for the analysis and visualization of the cases, or the data export tools. All this software is proprietary and vendor-specific. So they differ on the use of signals (CD or TI) to derive CPR events like CCs, on how CPR quality variables are computed, on the definition of start/end of the analysis intervals, and they do not exclude ROSC periods and uninterpretable or noisy signal periods from the analysis. For these reasons, CPR analysis of large numbers of OHCA cases frequently involves an onerous labor intensive manual annotation and revision of the data, for either complete episodes or intervals of interest. Consequently, there is a need for a uniform methodology and a framework to facilitate the analysis of large, complex and heterogeneous OHCA datasets.

The aim of this study was to derive and validate a methodology to automatically obtain CPR quality data in a uniform and systematic way from large OHCA datasets with clinical annotations and defibrillator files from multiple vendors. Since both the TI and CD are useful to monitor CCs during CPR,[11–13,30] we propose a systematic and uniform way to obtain CPR quality variables from the CD/TI signals stored in the defibrillator electronic files together with clinical annotations of ROSC. We demonstrate this new methodology on a large heterogenous OHCA dataset of over 2,000 cases with files from the three major defibrillator manufacturers.

Materials and methods

Data materials

The study dataset comprised the de-identified electronic defibrillator files and the clinical annotations, in spreadsheet format, from the adult OHCA cases enrolled in the Pragmatic Airway Resuscitation Trial (PART).[8] Data were collected by multiple EMS and fire agencies within the Resuscitation Outcome Consortium (ROC) initiative, and included files from the three major defibrillator vendors: Philips (HeartStart MRx. Koninklijke Philips N.V., Eindhoven, Netherlands), Stryker (Lifepak models 12, 15 and 500. Stryker, Kalamazoo, MI, US) and ZOLL (X and E series, AED Pro and AED plus. ZOLL Medical, Chelmsford, MA, US). As shown in Fig. 1, all files were exported using custom and proprietary software. Then, using an automated batch process, they were converted into a common MATLAB (MathWorks Inc., Natick, MA, US) format to ensure uniform data processing. Defibrillator data included biomedical signal waveforms as well as time-stamps for relevant events like device power-on, shocks, rhythm analyses or signal discontinuity, but also CPR data like the CC time-stamps calculated by the proprietary software of each vendor. The clinical annotations provided extensive additional information including ROSC occurrences and advanced airway insertion attempt and success/abandon times. Only the clinical annotations needed to define the intervals to calculate the CPR quality variables were used in this study, that is, ROSC annotations and times associated to airway insertion.

Only files with signals suitable for CC annotation were included, that is, impedance (Stryker and Philips) or sternal accelerometer based compression depth (Philips and ZOLL). Cases with less than 1-min of usable CC data were discarded. Cases involving several defibrillator files were also discarded or limited to the most representative file when the synchronization was not reliable. Of the initial 2731 defibrillator files, 51 could not be exported to an open format, and 319 contained no usable CC information. The final dataset included data from 2236 OHCA cases (2361 defibrillator files), 390 (390) Stryker, 925 (926) Philips, and 921 (1045) ZOLL cases, respectively.

Methods

A uniform automatic procedure (see Section 'Automatic annotation') was used to obtain CPR quality variables from all files, as shown in Fig. 1. State-of-the-art automatic algorithms were applied to detect CCs in the TI and/or CD signals.[11,12] The boundaries of the analysis interval, the ROSC periods, and the unreliable signal data periods were automatically identified. Then, a manual audit with an ad-hoc tool was conducted to correct CC annotations, ROSC periods, unreliable signal periods, and the boundaries of the analysis interval. These audited data were used as gold standard to compute the CPR quality variables. The agreement in CPR quality variables between the uniform automatic procedure and the gold standard was analyzed to evaluate the reliability of the automated procedure.

Automatic annotation

CCs were detected in the TI signal (Stryker and Philips files) using the algorithm proposed by Alonso et al.[12] A 2.1 Hz cut-off low-pass filter and a 0.1–0.25 Ω minimum amplitude adaptive threshold were selected as configuration parameters. Due to waveform similarities, the same algorithm was also used to detect CCs in ZOLL's CD signal. In this case, the adaptive threshold was set to 1–2 cm and a 1–5 Hz bandpass filter was employed, as the signal was found sensitive to low frequency noise and did not involve as many harmonic components as the TI. Finally, CC detection on the Philips CD signal was carried out using a standard peak detection method, with minimum peak separation and depth of 0.35 s and 1.5 cm respectively.[11] The CD algorithm was used when TI and CD signals were

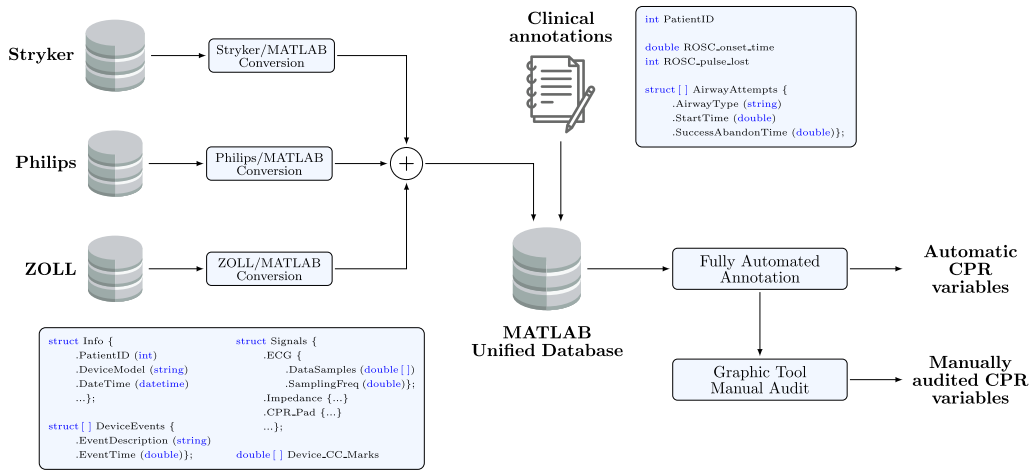


Fig. 1 – General workflow for data conversion and the computation of cardiopulmonary resuscitation (CPR) quality metrics. Each defibrillator file, along with the corresponding clinical annotations, was converted into a common MATLAB format structure. The association of the clinical information with the device file data was performed using the patient ID embedded in the defibrillator record filename. The common format files were subjected to a series of automatic annotation procedures to compute the CPR quality variables. A graphical tool permitted the manual audit/annotation of the cases.

simultaneously available. Compression series were defined as sequences of at least five consecutive and uninterrupted CCs. Interruptions were defined as intervals longer than 3 s without CCs, as customarily done for ROC data analysis.[31]

Intervals of signal unavailability and ROSC were automatically annotated and excluded from CPR quality calculations. Unreliable signal intervals were identified from the device's event time-stamps. Depths resulting from small accelerations (< 5 m s^{-2}) and thoracic impedances outside normal values (30–200 Ω) were also excluded. For cases involving several files, any time lapse between records was also marked as unavailable. When files overlapped in time, the annotations of the latest one prevailed for the overlapping interval. ROSC occurrences were identified using the clinical annotations. The onset of a ROSC event was moved to the start of the longest CC interruption around the annotated ROSC time. When pulse loss was clinically confirmed (transient ROSC), the end of ROSC was set at the resumption of CCs.

Following Kramer-Johansen et al.[32] the start of an episode was initially set at the first therapeutic event, i.e. first CC, shock or rhythm analysis. Then, it was advanced until a signal suitable for CC annotation was available. The end of the episode was set to the last CC or the onset of sustained ROSC. The analysis interval was defined as the time period between the start and end of the episode, but excluding periods with ROSC and periods with unreliable signals.

Manual audit of the episodes

A graphic user interface was developed in MATLAB to obtain the gold standard by manually auditing the annotations from the uniform automatic procedure. The tool allowed the display of all available signals, the modification of CC annotations (both individually and over user selected time intervals), and the adjustment of the boundaries of the analysis interval, including episode start/end times as well as ROSC intervals. Noise regions were also annotated when the CD/TI signals were uninterpretable. Fig. 2 shows a screenshot of the graphic user interface during the review of a Philips case.

CPR quality variables

Chest compression fraction (CCF), rate (CCR) and interruption ratio were computed. The CCR was defined as the frequency of CCs during compression series,[32] and was calculated as:

$$\mathrm{CCR} = \frac{60}{\mathrm{median}\{\Delta t_{cc}\}} \quad (\mathrm{min}^{-1}) \qquad (1)$$

where Δt_{cc} is the time in seconds between two consecutive CCs. The CCF was defined as the proportion of time with ongoing CCs in the analysis interval.[31] The interruption ratio was computed as the time-proportion with interruptions in CCs between the first and last CC of the analysis interval, that is, the complement of the CCF but only from start to end of CCs.

Performance evaluation and comparative analysis

The CPR quality variables were computed using the uniform automatic method and compared to the gold standard. The fully automatic method annotated CCs and the start/end of the analysis interval according to signal and event time-stamps in defibrillator files, and ROSC intervals according to the time-stamps provided by clinicians. The gold standard included reviewer corrections to all the former, plus noise intervals whenever signals were deemed uninterpretable. The agreement in CCF, CCR and interruption ratio was evaluated in terms of the unsigned error, the unsigned difference in the variables between the uniform automatic procedure and the gold standard.

When available, CPR quality variables were also computed using the CC annotations provided by the corresponding defibrillator vendor software. This allowed comparison of the uniform automatic procedure and the vendor-specific procedures.

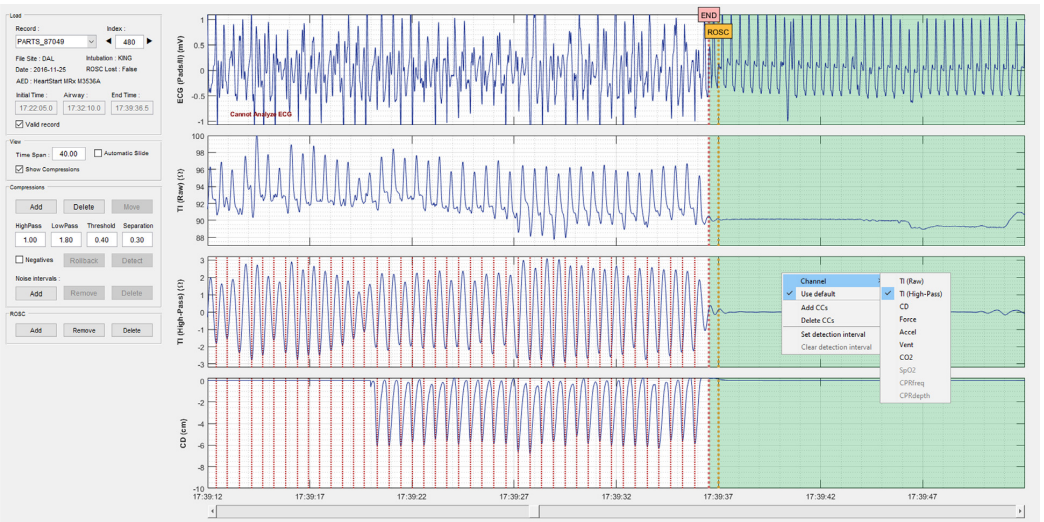


Fig. 2 – Screenshot of the revision tool during the analysis of a Philips case. Traces of the ECG, compression depth (CD) and thoracic impedance (TI), both in raw format and high-pass filtered for chest compression (CC) annotation, are presented. CCs (dotted red vertical lines) were automatically annotated using the CD signal, or the TI when the CD was absent. An interval with return of spontaneous circulation (ROSC) was annotated (green) using the clinical information (orange label). In this case, ROSC was sustained, so the onset of ROSC was the end of the analysis interval. Controls on the left permitted the manual annotation of CCs as well as ROSC and noisy intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The PART study focused on the comparison of different advanced airway strategies, and included clinical annotations on airway management. CPR quality variables were computed over airway insertion intervals, as a case study to show the flexibility of the proposed methodology. The airway insertion interval was defined as the period between the start of the first airway insertion attempt and the success/abandon time of the last one. These airway insertion attempt times were obtained from the clinical annotations of the PART study.

Results

Signal availability

The data converted in the uniform automatic procedure was used to measure, in each device type, the availability of the CD and TI signals for the calculation of the CPR quality variables. Availability was measured from device power-on to the last pads/leads disconnection. A total of 2236 cases were suitable for analysis, 390 from Stryker, 925 from Philips and 921 from ZOLL. For Stryker and Philips cases, median (interquartile range, IQR) TI availabilities were 92.9 (84.8–96.6) % and 93.4 (85.4–96.2) %, respectively. ZOLL's CD availability was 94.9 (88.0–97.1) %. This value is similar to that of the TI because their technology integrates the accelerometer into the defibrillation pads. CD availability for Philips was a lower 89.5 (59.4–96.9) %, because it requires an external CPR assist pad, which may not be present from the beginning of the episode or might have to be disconnected under certain circumstances, like in the use of mechanical compression devices. Overall data availability was largest for Philips records, which often included both TI and CD (from the 925 cases, 900 had TI and 886 CD). The blind interval from device power-on until the first signal source became available was also shorter for Philips, 0.5 (0.2–1.1) minutes, compared to those measured for Stryker and ZOLL, 1.0 (0.5–1.6) and 0.9 (0.5–1.5) minutes, respectively.

Episode-wide CPR quality

Table 1 shows, for each vendor and for the complete dataset, the median (IQR) error in CCF, CCR and interruption ratio between the uniform automatic procedure and the gold standard. Median error values below 2% in CCF, 1 min^{-1} in CCR and 1.5% in interruption ratio were measured for all devices. The proportion of episodes with large errors (>10% for CCF and interruption ratio, and >10 min^{-1} for CCR) were below 10% in all cases.

Large errors in CCF and interruption ratio were mostly caused by an incorrect delineation and identification of ROSC intervals in cases where the clinical annotations were inaccurate or incomplete. Of the 761 episodes with clinical ROSC annotations, ROSC onset times were reported with 1-min precision in 334 (42.9%) cases, and were missing in 10 (1.3%) cases. There were no time-stamps for ROSC termination, so for 329 cases with transient ROSC annotations, the end of ROSC had to be marked at the time CCs were resumed. Moreover, the clinical annotations only contained information about the first occurrence of ROSC, so secondary ROSC events were missed by the uniform automatic procedure. In the manual audit of the data (gold standard), secondary ROSC events were identified in 132 (40.1%) cases with clinically diagnosed pulse loss. These events occurred in long pauses in CCs (>1 min) with an organized heart rhythm and no other plausible explanation. Transient ROSC events were also added to 87 cases with no clinical ROSC

Table 1 – Median (IQR) unsigned error in CPR quality variables for the uniform automatic procedure. Full episodes of the three defibrillator vendors were considered.

		CCF (%)		CCR (min^{-1})		Interruptions (%)	
	N	error	> 10%	error	> 10 min^{-1}	error	> 10%
Stryker	389	1.6 (0.6–3.9)	36 (9.3%)	0.8 (0.3–2.5)	12 (3.1%)	1.4 (0.5–3.8)	36 (9.3%)
Philips	925	0.5 (0.1–1.9)	92 (9.9%)	0.3 (0.1–0.7)	1 (0.1%)	0.4 (0.1–1.4)	71 (7.7%)
ZOLL	918	0.7 (0.3–2.0)	68 (7.4%)	0.2 (0.1–0.6)	6 (0.7%)	0.6 (0.3–1.8)	67 (7.3%)
TOTAL	2232	0.7 (0.2–2.4)	196 (8.8%)	0.3 (0.1–0.9)	19 (0.9%)	0.6 (0.2–2.0)	174 (7.8%)

annotations. The episodes with undeclared ROSC events showed the largest errors in CCF, with a median (IQR) of 6.6 (2.2–17.8) %, and 37.4% of the cases presenting large errors (>10%) in CCF. Errors in CCF were large in 28.4% of cases with a single transient ROSC event, and in 9.9% of cases with sustained ROSC. In contrast, only 2.5% of the episodes without ROSC presented large errors in CCF. A detailed analysis of the error distributions is shown in the Bland-Altman plots of Fig. 3. As shown in the figure, incorrect ROSC annotations resulted in large underestimations of the CCF (overestimations of the interruption ratio). CCR estimations were found much more robust to misannotations of the analysis intervals.

Specific intervals, advanced airway device insertion

Table 2 shows the errors in CPR quality variables for a specific analysis sub-interval, the airway insertion interval in the PART study. The inclusion criteria for this sub-analysis were: 1) clinical time annotations for the start of the first insertion effort and the success/abandon of the last effort, and 2) available signal data, suitable for CC annotation, in over 50% of the airway insertion interval. A total of 1139 cases met the inclusion criteria, with a median (IQR) airway insertion duration of 1.2 (0.8–3.0) min. The estimated errors in CPR quality variables were lower for the airway insertion interval than for the episode analysis interval (Table 2).

Vendor-specific procedure

The CC detection algorithms are proprietary in the vendor-specific software, and their output determine the values of the CPR quality variables. Differences between vendor-specific CC-detectors and the ones used in our uniform automatic procedure were evaluated by comparing the errors in CPR quality variables. To evaluate only the effect of CC detection algorithms the manually audited analysis intervals (manual audit of episode start/end, ROSC and noise periods) were used for all the algorithms. As shown in Table 3, estimated errors were small either way, although the errors in CCF and interruption ratio, and the proportion of cases with large errors, were all significantly smaller for the algorithms used in this study ($p < 0.05$ for the Wilcoxon signed rank test). In addition, errors due to CC detection were larger for TI based algorithms than for those based on CD, as reported in previous studies.[11] Impedance is more sensitive to noise sources like patient or electrode movement, has more waveform variability,[33] and frequently presents low amplitude variations associated with CCs.

Discussion

This paper describes and evaluates a methodology to automatically process and extract information in a uniform and systematic way from

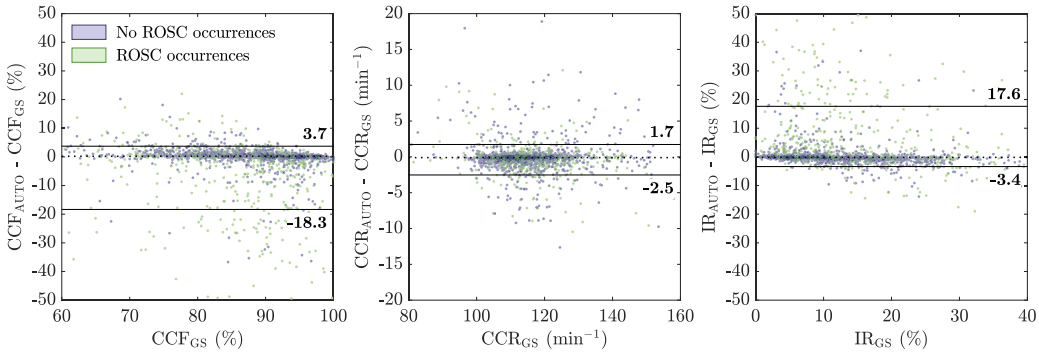


Fig. 3 – Bland-Altman plot for chest compression fraction (CCF), chest compression rate (CCR), and interruption ratio (IR). The values of the variables obtained for the uniform automatic procedure were compared to the gold standard. For each variable the dotted line shows the median error, and the solid lines the 90% levels of agreement (LOA). The median value of the errors is close to zero for all variables. Errors in the automatic annotation of the analysis interval resulted in large negative errors in CCF (positive for IR) and skewed LOAs. These were mainly related to cases with return of spontaneous circulation (ROSC) occurrences (green dots in the plot). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 – Median (IQR) duration of the advanced airway insertion intervals and unsigned error in CPR quality variables for the uniform automatic procedure.

			CCF (%)		CCR (min^{-1})		Interruptions (%)	
	N	Duration (min)	error	> 10%	error	> 10 min^{-1}	error	> 10%
Stryker	114	1.3 (1.0–3.0)	0.3 (0.0–2.5)	8 (7.0%)	0.5 (0.1–1.7)	5 (4.4%)	0.3 (0.0–2.1)	7 (6.1%)
Philips	607	1.3 (0.8–3.5)	0.0 (0.0–0.0)	21 (3.5%)	0.0 (0.0–0.3)	10 (1.6%)	0.0 (0.0–0.0)	21 (3.5%)
ZOLL	418	1.0 (0.7–3.0)	0.2 (0.0–1.0)	11 (2.6%)	0.1 (0.0–0.4)	7 (1.7%)	0.2 (0.0–0.9)	11 (2.6%)
TOTAL	1139	1.2 (0.8–3.0)	0.0 (0.0–0.7)	40 (3.5%)	0.0 (0.0–0.5)	22 (1.9%)	0.0 (0.0–0.6)	39 (3.4%)

Table 3 – Median (IQR) unsigned error in CPR quality variables for the uniform automatic procedure compared to the annotations of the manufacturer's software (vendor-specific procedure). Variables were computed over the reviewed analysis intervals, where the ROSC events and the start/end times were corrected. The manufacturer's software includes CC annotations based on impedance (Stryker) or depth (Philips and ZOLL).

		CCF (%)		CCR (min^{-1})		Interruptions (%)	
	N	error	> 10%	error	> 10 min^{-1}	error	> 10%
Uniform automatic CC annotations							
Stryker (TI)	390	0.9 (0.4–2.1)	5 (1.3%)	0.7 (0.3–2.2)	10 (2.6%)	0.9 (0.3–2.1)	6 (1.5%)
Philips	925	0.2 (0.0–0.6)	10 (1.1%)	0.2 (0.1–0.7)	1 (0.1%)	0.2 (0.0–0.5)	10 (1.1%)
- TI	894	0.8 (0.3–1.9)	33 (3.7%)	0.5 (0.2–1.0)	2 (0.2%)	0.8 (0.3–1.8)	33 (3.7%)
- CD	531	0.1 (0.0–0.3)	2 (0.4%)	0.3 (0.1–0.6)	0 (0.0%)	0.1 (0.0–0.2)	1 (0.2%)
ZOLL (CD)	921	0.4 (0.2–0.7)	4 (0.4%)	0.2 (0.1–0.4)	1 (0.1%)	0.4 (0.2–0.8)	4 (0.4%)
Vendor-specific automatic CC annotations							
Stryker (TI)	390	1.8 (0.7–4.0)	27 (6.9%)	1.4 (0.4–3.4)	24 (6.2%)	1.8 (0.7–3.9)	25 (6.4%)
Philips (CD)	531†	0.4 (0.2–1.0)	3 (0.6%)	0.3 (0.1–0.6)	0 (0.0%)	0.4 (0.1–0.9)	2 (0.4%)
ZOLL (CD)	921	0.4 (0.1–1.1)	20 (2.2%)	0.5 (0.2–1.4)	14 (1.5%)	0.4 (0.1–1.1)	20 (2.2%)

† From the 886 Philips cases with CD only 531 contained vendor-specific CC annotations.

large repositories of OHCA data, which contain clinical annotations and defibrillator files from multiple vendors. Defibrillator files and clinical annotations were obtained from the PART trial, and the analysis was performed for complete episodes and for specific sub-intervals of interest, such as the insertion period of the advanced airway device. The automated method provided accurate CPR quality variables (errors below 10 min^{-1} in CCR, and below 10% in CCF and interruption ratio) in more than 90% of the cases. Given the advent of numerous automatic detection algorithms for OHCA data based on signal processing and artificial intelligence, this methodology could be extended in the future to add information on ventilation rates[12,15–18] and volumes,[34,35] perfusion states,[22,24,36] or heart rhythm transitions,[28,37–39] depending on the available signals in the defibrillator files.

For CPR data processing, a key procedure is an accurate and systematic definition of the analysis interval, including the start/end of the treatment and the detection of ROSC events (non cardiac arrest intervals). The automated procedure used to set the start/end of the analysis interval was based on the event time-stamps of the device and its signal recordings. If start/end times as defined by Kramer-Johansen et al.[32] were not automatically identified, and instead the power on/power off (or last signal recording) time-stamps of the device were used, the error in CCF raised to 23.0 (10.9–42.6) %, a problem that has recently been reported.[40] The missannotation of periods with ROSC accounted for the majority of large errors in the automated procedure. Inaccurate clinical annotations and undocumented ROSC events explain these errors, and were one of the major challenges of the automatic procedure. Several methods have been proposed to automatically identify pulsed rhythms using the ECG alone[21] or in combination with other signals such as the impedance[19,20,23] or the capnogram.[22] The use of these algorithms to automatically detect ROSC intervals could contribute to a more rapid and precise annotation procedure, but those algorithms have not yet been validated for large repositories of OHCA data.

Regarding the automatic annotation of CCs, the detection algorithms used in this study were based on TI and CD, and in both cases resulted in lower errors in CPR quality variables compared to vendor's proprietary algorithms. Estimated errors were smaller for CCR than for CCF, which is strongly associated to survival.[41,42] However, errors in CCF were low (0.7%, see Table 1), and CCF estimates were more accurate using our procedure than the vendor specific software (Table 3). Also, CC depth, which is the stronger correlate of outcome,[43] was not measured in this study because there is no ground truth for depth (CD is a derived measure from accelerometer data), and thus errors could not be evaluated. Furthermore, CD was not available for one of the vendors. However, CC depth derived from the CD signal could be easily calculated in the automated procedure, and used in models to estimate patient outcome.[44]

Accuracy depended on the signals used to estimate the CPR quality metrics. The CD signal is derived from sternal accelerometers, either as part of an external assist pad (Philips) or integrated in the defibrillation pads (ZOLL), and is more stable and less artifact prone than the TI signal. Patient and/or electrode connection movement, short disconnections and low/high frequency noise, all affect the TI waveform and compromise the accuracy of the CC detection algorithms.[12] Still, the accuracy for the TI algorithm was high, in line with previous studies,[10,11] and the algorithm worked well for different devices. It must be stated that 192 episodes, 8.6% of the cases, included intervals with mechanical CCs. In the PART data used in this

study, we found no significant differences ($p > 0.05$, Chi square test) in proportion of cases with large errors for the CPR quality variables between the manual and mechanical cases. There were no differences either in median errors for CCF ($p = 0.07$, Mann Whitney U test), but differences between manual and mechanical CPR were significant for CCR (0.3% vs 0.2%, $p < 0.05$), and interruption ratio (0.6% vs 0.7%, $p < 0.05$). However, the differences in CCR and interruption ratio were of 0.1-percent points, which have no clinical importance. For details of this secondary analysis consult the supplementary data.

Our analysis includes, as a case study, the assessment of CPR during airway insertion, an interval of potential interest given the long interruptions in CPR observed during advanced airway insertion in previous studies.[45] This example shows the flexibility of the proposed methodology to 1) analyze CPR data in non-standard analysis intervals, 2) include device or clinical annotations to define the boundaries of sub-intervals of interest, such as the airway insertion interval, and 3) define ad-hoc CPR variables. This flexibility opens a wide variety of options to personalize the studies, add new covariates to the explanatory models in large clinical trials, and answer specific research questions. How the CPR quality variables in certain time intervals or around certain treatment procedures affect outcome could be analyzed. Many large clinical trials and research studies could benefit from this methodology, which allows the automatic analysis of thousands of OHCA cases from heterogeneous sources. Similar procedures are currently performed manually, which involve a very high cost in time and money. The software tools and analysis procedures of our automatic uniform procedure are available through the corresponding author for clinical OHCA research groups.

Limitations of the study

Although various devices from three vendors were considered, the availability of signals and the accuracy of the algorithms depend on their technology to acquire signals. Moreover, device models were closely linked to emergency agencies, which may operate under different protocols, and thus introduce some bias in the results. We did not include CC depth measures, nor any assessment of ventilation. Clinical annotations of ROSC were inaccurate, and in cases like transient ROSC and secondary ROSC events were frequently missing.

Conclusions

An automatic and uniform (vendor independent) methodology to compute CPR quality variables in large repositories of OHCA defibrillator files from multiple vendors was presented and validated. Median episode-wide errors below 2% in CCF, 1 min^{-1} in CCR, and 1.5% in interruption ratio, were measured for all signals and device models. The proposed automated method permits an accurate analysis of CPR data in large datasets with files from multiple defibrillator vendors, integrating relevant clinical information. This methodology could be extended to answer other research questions, by integrating different automated algorithms and considering ad-hoc intervals of interest for the analysis.

Conflicts of interest

Ahamed H. Idris is member of the Stryker (HeartSine) Belfast Clinical Advisory Board and of the American Heart Association National Emergency Cardiovascular Care Committee. Robert H. Schmicker is funded by the National Heart, Lung, and Blood Institute (NHLBI). Tom P. Aufderheide has an unrestricted research grant from ZOLL Medical, Inc. and was funded by NHLBI as site PI for the PART Trial.

Acknowledgments

This work was supported by the Spanish Ministerio de Ciencia, Innovacion y Universidades through grant RTI2018-101475-BI00, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), by the Basque Government through grants IT1229-19 and PRE_2020_2_0182, and by the University of the Basque Country (UPV/EHU) under grant COLAB20/01.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.resuscitation.2021.09.005.

REFERENCES

1. Perkins GD, Handley AJ, Koster RW, et al. European Resuscitation Council Guidelines for Resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation. Resuscitation 2015;95:81–99.
2. Soar J, Nolan JP, Böttiger BW, et al. European Resuscitation Council Guidelines for Resuscitation 2015: Section 3. Adult advanced life support. Resuscitation 2015;95:100–47.
3. Pellis T, Bisera J, Tang W, Weil MH. Expanding automatic external defibrillators to include automated detection of cardiac, respiratory, and cardiorespiratory arrest. Crit Care Med 2002;30:S176–8.
4. Abella BS, Edelson DP, Kim S, et al. CPR quality improvement during in-hospital cardiac arrest using a real-time audiovisual feedback system. Resuscitation 2007;73:54–61.
5. Wik L, Olsen JA, Persse D, et al. Manual vs. integrated automatic load-distributing band CPR with equal survival after out of hospital cardiac arrest. The randomized CIRC trial. Resuscitation 2014;85:741–8.
6. Nichol G, Leroux B, Wang H, et al. Trial of Continuous or Interrupted Chest Compressions during CPR. New Engl J Med 2015;373:2203–14.
7. Idris AH, Guffey D, Pepe PE, et al. Chest compression rates and survival following out-of-hospital cardiac arrest. Crit Care Med 2015;43:840–8.
8. Wang HE, Schmicker RH, Daya MR, et al. Effect of a Strategy of Initial Laryngeal Tube Insertion vs Endotracheal Intubation on 72-Hour Survival in Adults With Out-of-Hospital Cardiac Arrest: A Randomized Clinical Trial. JAMA 2018;320:769–78.
9. Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. Resuscitation 2019;141:174–81.
10. Zhang H, Yang Z, Huang Z, et al. Transthoracic impedance for the monitoring of quality of manual chest compression during cardiopulmonary resuscitation. Resuscitation 2012;83:1281–6.
11. Ayala U, Eftestøl T, Alonso E, et al. Automatic detection of chest compressions for the assessment of CPR-quality parameters. Resuscitation 2014;85:957–63.
12. Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics. Resuscitation 2015;88:28–34.
13. Kwok H, Coult J, Liu C, et al. An accurate method for real-time chest compression detection from the impedance signal. Resuscitation 2016;105:22–8.

14. Coult J, Blackwood J, Rea TD, Kudenchuk PJ, Kwok H. A Method to Detect Presence of Chest Compressions During Resuscitation Using Transthoracic Impedance. IEEE J Biomed Health Informat 2020;24:768–74.
15. Risdal M, Aase SO, Stavland M, Eftestøl T. Impedance-based ventilation detection during cardiopulmonary resuscitation. IEEE Trans Bio-med Eng 2007;54:2237–45.
16. Aramendi E, Elola A, Alonso E, et al. Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation. Resuscitation 2017;110:162–8.
17. Aramendi E, Lu Y, Chang MP, et al. A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation. Resuscitation 2018;132:41–6.
18. Jaureguibeitia X, Irusta U, Aramendi E, Owens P, Wang H, Idris A. Automatic detection of ventilations during mechanical cardiopulmonary resuscitation. IEEE J Biomed Health Informat 2020.
19. Risdal M, Aase SO, Kramer-Johansen J, Eftesol T. Automatic identification of return of spontaneous circulation during cardiopulmonary resuscitation. IEEE Trans Biomed Eng 2008;55:60–8.
20. Ruiz J, Alonso E, Aramendi E, et al. Reliable extraction of the circulation component in the thoracic impedance measured by defibrillation pads. Resuscitation 2013;84:1345–52.
21. Elola A, Aramendi E, Irusta U, et al. Deep Neural Networks for ECG-Based Pulse Detection during Out-of-Hospital Cardiac Arrest. Entropy 2019;21:305.
22. Elola A, Aramendi E, Irusta U, et al. Capnography: A support tool for the detection of return of spontaneous circulation in out-of-hospital cardiac arrest. Resuscitation 2019;142:153–61.
23. Alonso E, Irusta U, Aramendi E, Daya MR. A Machine Learning Framework for Pulse Detection During Out-of-Hospital Cardiac Arrest. IEEE Access 2020;8:161031–41.
24. Elola A, Aramendi E, Irusta U, Berve PO, Wik L. Multimodal algorithms for the classification of circulation states during out-of-hospital cardiac arrest. IEEE Trans Biomed Eng 2020.
25. Ristagno G, Mauri T, Cesana G, et al. Amplitude spectrum area to guide defibrillation: a validation on 1617 patients with ventricular fibrillation. Circulation 2015;131:478–87.
26. Savastano S, Baldi E, Raimondi M, et al. End-tidal carbon dioxide and defibrillation success in out-of-hospital cardiac arrest. Resuscitation 2017;121:71–5.
27. Chicote B, Aramendi E, Irusta U, Owens P, Daya M, Idris A. Value of capnography to predict defibrillation success in out-of-hospital cardiac arrest. Resuscitation 2019;138:74–81.
28. Rad AB, Eftestøl T, Irusta U, et al. An automatic system for the comprehensive retrospective analysis of cardiac rhythms in resuscitation episodes. Resuscitation 2018;122:6–12.
29. Isasi I, Irusta U, Rad AB, et al. Automatic Cardiac Rhythm Classification With Concurrent Manual Chest Compressions. IEEE Access 2019;7:115147–59.
30. Aramendi E, Ayala U, Irusta U, Alonso E, Eftestøl T, Kramer-Johansen J. Suppression of the cardiopulmonary resuscitation artefacts using the instantaneous chest compression rate extracted from the thoracic impedance. Resuscitation 2012;83:692–8.
31. Iyanaga M, Gray R, Stephens SW, et al. Comparison of methods for the determination of cardiopulmonary resuscitation chest compression fraction. Resuscitation 2012;83:568–71.
32. Kramer-Johansen J, Edelson DP, Losert H, Köhler K, Abella BS. Uniform reporting of measured quality of cardiopulmonary resuscitation (CPR). Resuscitation 2007;74:406–17.
33. Aramendi E, Irusta U, Ayala U, Naas H, Kramer-Johansen J, Eftestøl T. Filtering mechanical chest compression artefacts from out-of-hospital cardiac arrest data. Resuscitation 2016;98:41–7.
34. Terndrup TE, Rhee J. Available ventilation monitoring methods during pre-hospital cardiopulmonary resuscitation. Resuscitation 2006;71:10–8.
35. Berve PO, Irusta U, Kramer-Johansen J, et al. Transthoracic impedance measured with defibrillator pads–New interpretations of signal change induced by ventilations. J Clin Med 2019;8:724.
36. Skogvoll E, Nordseth T, Sutton RM, et al. Factors affecting the course of resuscitation from cardiac arrest with pulseless electrical activity in children and adolescents. Resuscitation 2020;152:116–22.
37. Kvaløy JT, Skogvoll E, Eftestøl T, et al. Which factors influence spontaneous state transitions during resuscitation? Resuscitation 2009;80:863–9.
38. Alonso E, Eftestøl T, Aramendi E, Kramer-Johansen J, Skogvoll E, Nordseth T. Beyond ventricular fibrillation analysis: Comprehensive waveform analysis for all cardiac rhythms occurring during resuscitation. Resuscitation 2014;85:1541–8.
39. Elola A, Aramendi E, Rueda E, Irusta U, Wang H, Idris A. Towards the Prediction of Rearrest during Out-of-Hospital Cardiac Arrest. Entropy 2020;22:758.
40. Gupta V, Schmicker RH, Owens P, Pierce AE, Idris AH. Software annotation of defibrillator files: Ready for prime time? Resuscitation 2021;160:7–13.
41. Christenson J, Andrusiek D, Everson-Stewart S, et al. Chest compression fraction determines survival in patients with out-of-hospital ventricular fibrillation. Circulation 2009;120:1241–7.
42. Vaillancourt C, Petersen A, Meier EN, et al. The impact of increased chest compression fraction on survival for out-of-hospital cardiac arrest patients with a non-shockable initial rhythm. Resuscitation 2020;154:93–100.
43. Vadeboncoeur T, Stolz U, Panchal A, et al. Chest compression depth and survival in out-of-hospital cardiac arrest. Resuscitation 2014;85:182–8.
44. Nichol G, Daya MR, Morrison LJ, et al. Compression depth measured by accelerometer vs. outcome in patients with out-of-hospital cardiac arrest. Resuscitation 2021;167:95–104.
45. Wang HE, Simeone SJ, Weaver MD, Callaway CW. Interruptions in cardiopulmonary resuscitation from paramedic endotracheal intubation. Ann Emergency Med 2009;54:645–652.e1.

A.3 Publications associated to objective 3

A.3.1 First journal paper

	Publication in international journal
Reference	Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Jeffrey L. Jarvis, Jestin N. Carlson, Unai Irusta, Erik Alonso, Tom Aufderheide, Robert H. Schmicker, Matthew L. Hansen, Ryan M. Huebinger, M. Ricardo Colella, Richard Gordon, Robert Suchting, Ahamed H. Idris, "Airway strategy and chest compression quality in the Pragmatic Airway Resuscitation Trial", *Resuscitation 2021*, vol. 162, pp. 93-98.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (3/32) based on Web of Science Rank 2021 • **Impact factor**: 6.251

RESUSCITATION 162 (2021) 93 –98

Available online at www.sciencedirect.com

Resuscitation

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

Clinical paper

Airway strategy and chest compression quality in the Pragmatic Airway Resuscitation Trial

Henry E. Wang[a,*], *Xabier Jaureguibeitia*[b], *Elisabete Aramendi*[b], *Jeffrey L. Jarvis*[c,a], *Jestin N. Carlson*[d], *Unai Irusta*[b], *Erik Alonso*[j], *Tom Aufderheide*[e], *Robert H. Schmicker*[f], *Matthew L. Hansen*[g], *Ryan M. Huebinger*[a], *M. Riccardo Colella*[e], *Richard Gordon*[a], *Robert Suchting*[h], *Ahamed H. Idris*[i]

[a] *Department of Emergency Medicine, The University of Texas Health Science Center at Houston, Houston, TX, United States*
[b] *Department of Communication Engineering, BioRes Group, University of the Basque Country, Bilbao, Spain*
[c] *Williamson County Emergency Medical Services, Georgetown, TX, United States*
[d] *Department of Emergency Medicine, The University of Pittsburgh, Pittsburgh, PA, United States*
[e] *Department of Emergency Medicine, Medical College of Wisconsin, Milwaukee, WI, United States*
[f] *Center for Biomedical Statistics, The University of Washington, Seattle, WA, United States*
[g] *Department of Emergency Medicine, Oregon Health & Science University, Portland, OR, United States*
[h] *Department of Psychiatry and Behavioral Sciences, The University of Texas Health Science Center at Houston, Houston, TX, United States*
[i] *Department of Emergency Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States*
[j] *Department of Applied Mathematics, University of the Basque Country, Bilbao, Spain*

Abstract

Background: Chest compression (CC) quality is associated with improved out-of-hospital cardiopulmonary arrest (OHCA) outcomes. Airway management efforts may adversely influence CC quality. We sought to compare the effects of initial laryngeal tube (LT) and initial endotracheal intubation (ETI) airway management strategies upon chest compression fraction (CCF), rate and interruptions in the Pragmatic Airway Resuscitation Trial (PART).

Methods: We analyzed CPR process files collected from adult OHCA enrolled in PART. We used automated signal processing techniques and a graphical user interface to calculate CC quality measures and defined interruptions as pauses in chest compressions longer than 3 s. We determined CC fraction, rate and interruptions (number and total duration) for the entire resuscitation and compared differences between LT and ETI using t-tests. We repeated the analysis stratified by time before, during and after airway insertion as well as by successive 3-min time segments. We also compared CC quality between single vs. multiple airway insertion attempts, as well as between bag-valve-mask (BVM-only) vs. ETI or LT.

Results: Of 3004 patients enrolled in PART, CPR process data were available for 1996 (1001 LT, 995 ETI). Mean CPR analysis duration were: LT 22.6 ± 10.8 min vs. ETI 25.3 ± 11.3 min ($p < 0.001$). Mean CC fraction (LT 88% vs. ETI 87%, $p = 0.05$) and rate (LT 114 vs. ETI 114 compressions per minute (cpm), $p = 0.59$) were similar between LT and ETI. Median number of CC interruptions were: LT 11 vs. ETI 12 ($p = 0.001$). Total CC interruption duration was lower for LT than ETI (LT 160 vs. ETI 181 s, $p = 0.002$); this difference was larger before airway insertion (LT 56 vs. ETI 78 s, $p < 0.001$). There were no differences in CC quality when stratified by 3-min time epochs.

Conclusion: In the PART trial, compared with ETI, LT was associated with shorter total CC interruption duration but not other CC quality measures. CC quality may be associated with OHCA airway management.

Keywords: Cardiopulmonary arrest, Airway management, Intubation, Emergency medical service

* Corresponding author at: Presented at the National Association of EMS Physicians Annual Meeting, January 2021, Department of Emergency Medicine, The Ohio State University, 750 Prior Hall, 376 W 10th Avenue, Columbus, OH 43210, United States.
E-mail address: henry.wang@osumc.edu (H.E. Wang).

https://doi.org/10.1016/j.resuscitation.2021.01.043
Received 21 November 2020; Received in revised form 15 January 2021; Accepted 28 January 2021

Introduction

Sudden out-of-hospital cardiopulmonary arrest (OHCA) is a major public health problem affecting over 300,000 adults in the United States each year, with only 1 in 10 surviving.[1] Airway management is an important component of resuscitation from OHCA, facilitating delivery of oxygen to lungs for circulation to ischemic organs. Common advanced airway management techniques used by rescuers include endotracheal intubation (ETI–"intubation") and supraglottic airway insertion (SGA) such as the laryngeal tube (LT). In the Pragmatic Airway Resuscitation Trial (PART), LT was associated with better OHCA outcomes than ETI, including higher 72-h survival, hospital survival, and hospital survival with favorable neurologic status.[2] The mechanism for this difference was unknown.

Successful OHCA resuscitation requires high quality chest compressions (CC) to provide effective circulation of oxygen throughout the body. Prior studies highlight the influence of CC rate, fraction and interruptions upon OHCA outcomes.[3–6] For example, OHCA survival is highest when CC fraction (CCF) exceeds 0.60.[3,7,8] A CC rate of 100–120 cpm has also been associated with optimal survival.[9] Because of its complexity, ETI has strong potential to adversely influence CC quality.[10] Prior studies have linked airway management efforts with chest compression quality.[11,12] While SGA insertion entails simpler technique than ETI, only limited data have compared their relative influence upon CC quality.[13]

The objective of this study was to determine the differences in CC quality between airway management strategies in the PART trial.

Methods

Study design

We conducted a *post hoc* analysis of data from the PART trial.[2] Institutional Review Boards of participating institutions approved the parent PART study under federal Exception from Informed Consent rules (21 CFR 50.24).

Setting

The PART trial compared different paramedic airway management strategies in adult OHCA.[2] The 27 emergency medical service (EMS) agencies participating in the trial were associated with the Birmingham (Alabama), Dallas-Fort Worth (Texas), Milwaukee (Wisconsin), Pittsburgh (Pennsylvania) and Portland (Oregon) sites of the Resuscitation Outcomes Consortium. The trial used cluster-randomization with crossover, assigning adult OHCA to strategies of initial-ETI vs. initial-LT. The primary outcome was 72-h hospital survival. Secondary outcomes included return of spontaneous circulation, hospital survival and hospital survival with favorable neurologic function (Modified Rankin Scale $\leq$3).

Selection of participants

The parent trial included all adult OHCA $\geq$18 years (or per local interpretation) requiring advanced airway management or bag-valve-mask (BVM) ventilation. Key exclusion criteria included patients $<$ 18 years, pregnant women, prisoners, traumatic cardiac arrest, and the initial presence of a non-study advanced EMS unit. The parent trial enrolled 3004 patients from December 1, 2015 through November 4, 2017. For this analysis, we included all enrolled patients with available CPR process files.

Methods of measurement – analysis of CPR process files

CPR process data were recorded by standard cardiac monitors used by participating EMS agencies, including monitors manufactured by Physio-Control, Inc. (Life-Pak 15 series, Physio-Control, Redmond, Washington), Zoll, Inc. (X-series, Zoll, Inc., Chelmsford, Massachusetts) and Philips, Inc. (MRX series, Philips Healthcare, Andover, Massachusetts). The monitors detected CC through either accelerometers (Zoll and Philips) or changes in electrical impedance (Physio-Control and Philips), which has been widely used to detect fluctuations due to both CC and ventilations.[14–18] As part of standard protocols, EMS personnel collected CPR process measurements on all adult OHCA. The manufacturer's CPR files were attached electronically to EMS electronic health record systems. A central study team used commercial software (Philips EventReview, PhysioControl CodeStat, and Zoll RescueNet Code Review) to view the CPR process files and to allow cleaning of corrupted or repeated files.

We used previously developed and extensively validated methods for automated importing and analysis of CPR process files.[14–18] All CPR process files were converted to a common MATLAB (MathWorks Inc., Natick, Massachusetts) file format in order to apply a uniform annotation procedure. We first applied waveform processing algorithms to automatically annotate and characterize CCs using both the thoracic impedance and the accelerometer-based compression depth channels.[15,18] We then used a previously developed Graphical User Interface tool to review the CC annotations and to identify intervals or events of interest. Research personnel used the GUI to review each individual file to review or correct key events, including initiation and end of CC, periods with artifact or noise, and ROSC.[14–18] Suitable intervals for CPR quality analysis were identified according to Kramer-Johansen, et al; i.e., between the first therapeutic event (CC, shock or rhythm analysis) and the end of resuscitation (last CC or persistent ROSC).[19]

Primary exposure

The primary exposure of interest was the assigned airway management strategy (initial-LT vs. initial-ETI). We analyzed the data according to intention-to-treat principles.

Outcomes

The primary outcomes of the study were CC fraction, rate and interruptions.[3,8,9,20,21] We defined (1) a CC interruption as a pause in CC longer than 3 s, (2) CC series as a group of at least 5 consecutive CCs without interruption, (3) CCF as the proportion of time (excluding ROSC and signal unavailability intervals) with active CC series in the analysis interval, and (4) CC rate as the number of CCs per minute of resuscitation, computed as the inverse of the median time difference between consecutive CCs within a CC series. We determined the mean and total duration of CC interruptions. Due to variations in the technologies used across participating EMS agencies, CC depth was not consistently measured. Therefore for this analysis we opted not to study CC depth. Because of the absence of validated detection methods, we did not assess ventilations in this analysis.

Analysis

We determined differences between LT and ETI for each CC process measure for a) the entire resuscitation episode, b) the time period

Table 1 – Characteristics of the study population. Includes *n* = 1001 with initial laryngeal tube (LT) and *N* = 995 with endotracheal intubation (ETI).

Characteristics	LT *N* (%)	ETI *N* (%)	*P*-value
Age, Median (IQR)	64 (22)	63 (22)	0.42
Sex			
Male	636 (63.6)	589 (59.2)	0.05
Female	364 (36.3)	406 (40.8)	
Race			
White	540 (54.0)	472 (47.4)	<0.001
Hispanic	53 (5.3)	69 (6.9)	0.23
Black	252 (25.2)	333 (33.5)	<0.001
Asian	22 (2.2)	16 (1.6)	0.35
Pacific Islander	1 (0.1)	1 (0.1)	0.5
Native American	2 (0.2)	2 (0.2)	0.64
Other	6 (0.6)	11 (1.1)	0.37
Witnessed Arrest, n/N (%)	443 (48.8)	450 (49.0)	0.99
EMS witnessed	103 (11.4)	109 (11.9)	0.79
Bystander witnessed	340 (37.5)	341 (37.1)	0.90
Not witnessed	464 (51.2)	469 (51.0)	0.99
Bystander chest compressions, n/N (%)			
Yes	461 (53.9)	447 (54.6)	0.80
No	391 (46.1)	392 (45.4)	
Initial rhythm			
Shockable	175 (16.6)	160 (16.4)	0.40
Non-shockable	803 (83.4)	818 (83.6)	

before airway insertion, and c) the time period after airway insertion. Airway insertion consisted of the point of successful airway placement or abandonment of failed attempts, as reported by EMS personnel. We repeated the analysis stratified by 3-min epochs. We also analyzed the data limited to the time during airway insertion efforts. We defined the "during" airway insertion period as a subset of the "before" period attributed by EMS personnel to airway insertion efforts. EMS personnel reported the beginning and end of airway insertion efforts, including the time of successful airway placement or the abandonment of airway efforts. We did not alter the reported timepoints when conducting the analysis.

We repeated this analysis using imputation in cases where the airway placement time was missing or implausible (for example, zero or negative elapsed time). Where the end of airway insertion efforts was missing, we imputed the value using the median duration of the airway insertion efforts for the corresponding airway type. We similarly imputed the airway insertion time if the reported elapsed time to airway placement was zero or negative. Only efforts with both precise ($\leq$10 s) timestamps were considered to compute the imputable durations. Imputation was necessary for approximately 35% of the cases; 11.7% for missing time points, and 23.8% for implausible values.

Finally, we examined differences in total CC interruptions between cases with single- and multiple airway insertion attempts. We examined CC differences in the subpopulation of patients receiving a single airway attempt. We also examined differences between cases receiving BVM only vs. [LT insertion or ETI].

We used univariate t-tests to facilitate all comparisons between LT and ETI. Because our focus on the association of airway strategy with CC quality (not patient clinical outcomes), we did not adjust the analysis for clinical cofounders such as Utstein variables.[22] We conducted all analyses using MATLAB (MathWorks Inc., Natick, Massachusetts).

Results

Of 3004 patients enrolled in the parent trial, we included 1996 in this analysis, including 1001 LT and 995 ETI. We excluded cases with unavailable or defective monitor files (n=533), no advanced airway insertion attempts (n=191) or unusable impedance signals for over 50% of the CPR process file (n=284). The final CPR process files originated from PhysioControl (n=316), Zoll (n=834) and Philips (n=846) monitors. There were n=191 cases with BVM ventilation only.

LT cases were more likely to involve men and whites. Other baseline characteristics (age, witnessed arrest, bystander chest compressions, and initial cardiac rhythm) were similar between LT and ETI (Table 1). Mean duration of resuscitation for the CC analysis was shorter for LT than ETI (LT 22.6 $\pm$ 10.8 min vs. ETI 25.3 $\pm$ 11.3 min, $p<0.001$).

CCF for the entire resuscitation episode was similar between LT and ETI (LT 87.8% vs. ETI 87.1%, p=0.05) (Table 2, Appendix 1). CCF was similar between LT and ETI for the time before and after airway insertion. There were no differences in CCF between LT and ETI when assessed in 3-min epochs or when limited to the time period during airway insertion efforts (Appendix 2).

CC rate for the entire resuscitation episode was similar between LT and ETI (LT 113.7 vs. ETI 114.0 cpm, p=0.59) (Table 3, Appendix 3). CC rate was similar between LT and ETI for the time before and after airway insertion. There were no differences in CC rate between LT and ETI when assessed in 3-min epochs or when limited to the time period during airway insertion efforts (Appendix 4).

The number of CC interruptions were higher for ETI than LT; LT median 11.0 (IQR 6.0–17.0) vs. ETI median 12.0 (IQR 8.0–19.0), (p=0.001) (Table 4, Appendix 5). While the duration of each CC

Table 2 – Differences in chest compression fraction between airway strategies (LT vs. ETI). Full table in Appendix 1. Airway insertion defined as point of successful airway placement or abandonment of failed attempts. LT = Laryngeal tube. ETI = Endotracheal intubation. CCF = Chest compression fraction.

Time period	LT			ETI			Difference (ETI-LT)	
	N	CCF Mean % (SD)	Analysis time sec (SD)	*N*	CCF Mean % (SD)	Analysis time sec (SD)	ΔCCF % (95% CI)	p-Value
Entire resuscitation episode	1001	87.9 (8.4)	1355.4 (650.3)	995	87.1 (8.7)	1514.3 (680.5)	−0.8 (−1.5 to 0.02)	0.05
Before airway insertion	903	86.9 (12.4)	463.5 (296.2)	886	87.3 (10.6)	659.4 (349.3)	0.4 (−0.7 to 1.4)	0.49
After airway insertion	902	88.8 (9.1)	974.4 (589.6)	866	88.2 (9.8)	953.2 (596.8)	−0.6 (−1.5 to 0.3)	0.17

Table 3 – Differences in chest compression rate between airway strategies (LT vs. ETI). Full table in Appendix 2. Airway insertion defined as point of successful airway placement or abandonment of failed attempts. LT = Laryngeal tube. ETI = Endotracheal intubation. CC = Chest compression. Cpm = compressions per minute.

Time period	LT			ETI			Difference (ETI-LT)	
	N	CC rate Mean cpm (SD)	Analysis time sec (SD)	N	CC rate Mean cpm (SD)	Analysis time sec (SD)	ΔCC rate cpm (95% CI)	p-Value
Entire resuscitation episode	1001	113.7 (9.1)	1355.4 (650.3)	995	114.0 (10.5)	1514.3 (680.5)	0.2 (−0.6 to 1.1)	0.59
Time before airway insertion	903	112.6 (11.2)	463.5 (296.2)	886	113.0 (11.2)	659.4 (349.3)	0.4 (−0.6 to 1.4)	0.45
Time after airway insertion	902	113.6 (11.0)	974.4 (589.6)	866	113.8 (10.9)	953.2 (596.8)	0.2 (−0.8 to 1.2)	0.72

interruption was similar between LT and ETI (LT 12.6 vs. ETI 13.0 s, $p = 0.78$), the total duration of all CC interruptions was longer for ETI than LT (ETI 180.8 vs. LT 160.0 s, $p = 0.002$). The total duration of CC interruptions was higher for ETI than LT prior to (ETI 77.9 vs. LT 56.4 s, $p < 0.001$) but not after airway insertion (LT 109.9 vs. ETI 113.3 s, $p = 0.58$). There were no differences in the duration of CC interruptions when stratified by 3-min epochs (Appendix 6). When limited to the period during airway insertion, ETI was associated with slightly longer total CC interruptions than LT (LT 15.4 vs. ETI 20.9 s, $p < 0.01$; imputed values LT 10.1 vs. ETI 16.5 s, $p < 0.001$). ETI was associated with longer total airway insertion efforts than LT (LT 131.7 vs. ETI 177.2 s, $p < 0.001$; imputed values LT 88.0 vs. ETI 139.0 s, $p < 0.001$).

Total CC interruption duration was shorter for single ($n = 1457$) than multiple ($n = 539$) airway insertion attempts (single 127 vs. multiple 153 s, $p = 0.005$). When limiting the analysis to cases with a single airway insertion effort ($n = 887$ for LT, $n = 570$ for ETI), total CC interruption duration was slightly higher for ETI than LT (ETI 174.6 vs. LT 157.5 s, $p = 0.03$). Total duration of CC interruptions prior to airway insertion was higher for ETI than LT (ETI 68.2 vs. LT 52.2, $p < 0.001$). No differences were observed in CCF and CC rate between airway groups. No significant differences were observed in airway insertion durations and total CC interruptions during airway insertion.

Total CC interruption duration was shorter for BVM-only than LT or ETI ($n = 1996$); BVM $n = 109$, 104 s vs. [LT or ETI] $n = 1996$, 170 s, $p < 0.001$). However, the total resuscitation duration was shorter for BVM than LT or ETI (BVM 698 s vs. [LT or ETI] 1434.6 s, $p < 0.001$). CCF, which is episode duration independent, was higher for LT or ETI than for BVM ([LT or ETI] 87.5% vs. BVM 84.1%, $p < 0.001$). Of note, continuous CCs was the most frequent CPR strategy employed before airway insertion and during BVM ventilation, with 30:2 CPR patterns being observed only on a relatively small portion (approximately 10%) of the cases.

Discussion

The parent PART trial found that an airway management strategy of initial-LT resulted in improved adult OHCA outcomes compared with a strategy of initial-ETI. A postulated reason for the differences in clinical outcomes is the differential effects of airway strategy upon CC quality. In the current analysis we found that compared with ETI, LT was associated with shorter total CC interruption duration but not other CC quality measures. Our analysis is one of the largest to describe linkages between airway management efforts and CC quality. Our results are bolstered by the use of automated signal processing techniques for systematic ascertainment of CC quality measures.

In a series of 100 adult OHCA from Pittsburgh, paramedic ETI efforts were associated with a median of two (2) chest compression interruptions totaling a median of 109.5 s.[11] In an analysis of 2767 adult OHCA resuscitations in the Resuscitation Outcomes Consortium Prehospital Resuscitation using an IMpedance valve and Early versus Delayed (PRIMED) trial, chest compression fraction was higher with supraglottic airway insertion than ETI during the 2 min before and after airway insertion.[12] In an analysis of data from the PARAMEDIC2 trial, Deakin, et al. found no difference in CCF between airway management strategies (SGA vs. ETI vs. both vs. none), but the analysis was limited to only 286 of the total 8000 patients enrolled in the parent trial.[23,24] We note that the CCF observed in our current series (>87%) is very high compared with prior published series, suggesting that EMS personnel were very adept at maintaining CC quality.[3,20] Differences in CC quality may have been more evident with EMS agencies or settings with lower baseline CC quality.

While we did not observe interval differences in CC rate, fraction, or individual interruptions between airway techniques, we did observe that ETI was associated with total CC interruptions almost 20 s longer than with LT. We note that total resuscitation duration was 3 min longer for ETI than LT. The exact reasons for the longer resuscitation duration for ETI is not known. Variations in practice or selection bias may have resulted in longer resuscitation efforts, enabling extended airway insertion effort. However, the ETI-based strategy may have also directly extended the direction of care. The latter is plausible given that ETI technique entails more complex technique than LT insertion. The extended duration of resuscitation would also explain why total CC interruptions were greater for ETI than LT despite similar CCF.

Of note, the ETI success rate observed in PART was approximately 53%; the sensitivity analysis limited to cases with a single airway insertion effort suggests that multiple ETI attempts may have explained some—but not all—of the increased CC interruptions. As emphasized in the parent trial, we believe that it is the strategy of airway management—not the mechanics of the individual devices—that likely influenced OHCA and outcomes. When we limited the analysis to ETI and LT cases with a single insertion attempt, total CC interruption was still longer for ETI than LT, suggesting that there could be other aspects of ETI technique influencing CC continuity. While our study suggests shorter CC interruptions with BVM than [LT or ETI], we believe that this is due primarily to the two-fold shorter total resuscitation time observed in the BVM-only group.

In the parent PART trial, 72-h survival was 2.9% higher for LT than ETI. It is unclear if the CC interruptions in this series influenced patient outcomes. ETI is a complex procedure and has been associated with numerous adverse events including failed insertion efforts, tube misplacement or dislodgement and multiple insertion attempts.[25,26] The presumption is that these deviations influence outcomes by causing CC interruptions. However, it is unclear if the additional 20

Table 4 – Differences in chest compression interruptions between airway strategies (LT vs. ETI). Full table in Appendix 3. Airway insertion defined as point of successful airway placement or abandonment of failed attempts. LT = Laryngeal tube. ETI = Endotracheal intubation. CPR = cardiopulmonary resuscitation.

Time period	LT				ETI				Difference (ETI-LT)			
	Number of interruptions Median (IQR)	Duration each interruption Mean Sec (SD)	Duration all interruptions Mean Sec (SD)	Analysis time Mean Sec (SD)	Number of interruptions Median (IQR)	Duration each interruption Mean Sec (SD)	Duration all interruptions Mean Sec (SD)	Analysis time Mean Sec (SD)	Δ Duration each interruption Mean Sec (95% CI)	*p*-value	Δ Duration all interruptions Sec (95% CI)	*p*-value
Entire resuscitation episode	11.0 (6.0–17.0)	12.6 (24.2)	160.0 (157.9)	1355.4 (650.3)	12.0 (8.0–19.0)	13.0 (18.0)	180.8 (141.1)	1514.3 (680.5)	0.1 (−0.7 to 1.0)	0.78	20.8 (7.7–34.0)	0.002
Before airway insertion	4.0 (2.0–6.0)	10.7 (16.5)	56.4 (63.6)	463.5 (296.2)	5.0 (3.0–9.0)	11.5 (14.6)	77.9(74.0)	659.4 (349.3)	0.2 (−1.1 to 1.5)	0.78	21.5 (15.1–27.9)	<0.001
After airway insertion	7.0 (3.0–11.0)	13.8(23.8)	113.3 (136.2)	974.4 (589.6)	6.0 (3.0–10.0)	13.9 (20.0)	109.9 (118.3)	953.2 (596.8)	−0.2 (−1.4 to 1.0)	0.75	−3.3 (−15.2 to 8.5)	0.58

−30 s of CC interruptions in this series are clinically important. While potentially linked with airway technique and OHCA outcomes, we did not study CC depth because accelerometer-based data were available for only a portion of cases.[9,21] In ongoing efforts we are using advanced processing of the thoracic impedance signal to assess ventilation differences between ETI and LT.

While we did not observe major differences in CC quality between ETI and LT, we still believe that maintenance of CC quality is important during OHCA advanced airway management efforts. Resuscitation from OHCA requires teamwork and coordination of multiple rescue interventions. Mitigation of distractions is important to facilitate other components of resuscitation. The current study suggests that compared with ETI, LT is associated with shorter total CC interruptions, a finding that is consistent with the simpler technique of LT insertion. We would expect other supraglottic airways such as the igel and laryngeal mask airway to yield similar advantages. For ETI-based efforts, video laryngoscopy and the gum elastic bougie have been associated with improved first-pass ETI success, but these impact on CC and patient outcomes are unknown.[27–30] Regardless of the approaches selected, EMS teams should devise approaches that best optimize coordinated resuscitation care.

Limitations

While we did not have CPR process files for all enrolled cases, our analysis is one of the largest series linking airway technique with CC quality. We did not assess chest compression depth due to the limited number of cases with accelerometer-based CPR process files; CC depth may plausibly be influenced by different airway techniques. We did not formally link CC quality measures with patient outcomes, but we observed few discernible differences in CC quality when stratified by airway strategy. We did not assess ventilations; analysis of ventilations is target of a future analysis. ETI success rates observed in the trial were lower than those from prior reports and the extent to which this difference may contribute to the improved survival noted in the parent trial is not known.[31] We did not have information on the airway management protocols, training protocols, or practice patterns across the participating agencies; variations in practice and preparation may have influenced the observed results. The intubation success rate reported in the parent trial was approximately 53%, a figure lower than prior studies; intubation performance may have influenced CC quality. We relied upon EMS personnel reports to define the start and end of airway management efforts.

Conclusion

Compared with ETI, LT was associated with shorter total CC interruption but not other CC quality measures in the PART trial. Single-attempt airway insertion and BVM-only were associated with shorter total CC interruptions. OHCA airway management may influence CC quality.

Conflict of interest

This study received financial support from the Spanish Ministerio de Ciencia, Innovación y Universidades through project RTI2018-101475-BI00 jointly with the Fondo Europeo de Desarrollo Regional (FEDER); and from the Basque Government through grants IT1229-19 and grant PRE-2019-1-0209.

Authors' contribution

All authors have made substantial contributions to all of the following: (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, (3) final approval of the version to be submitted.

There is no overlap with previous publications other than the parent PART study and we confirm that the manuscript, including related data, figures and tables, has not been published previously and that the manuscript is not under consideration elsewhere at this time.

Acknowledgements

This work is supported by award UH3-HL125163 from the National Heart, Lung and Blood Institute.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.resuscitation.2021.01.043.

REFERENCES

1. Go AS, Mozaffarian D, Roger VL, et al. Heart disease and stroke statistics – 2014 update: a report from the American Heart Association. Circulation 2014;129:e28–e292.
2. Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. JAMA 2018;320:769–78.
3. Christenson J, Andrusiek D, Everson-Stewart S, et al. Chest compression fraction determines survival in patients with out-of-hospital ventricular fibrillation. Circulation 2009;120:1241–7.
4. Aufderheide TP, Lurie KG. Death by hyperventilation: a common and life-threatening problem during cardiopulmonary resuscitation. Crit Care Med 2004;32:S345–51.
5. Berg RA, Sanders AB, Kern KB, et al. Adverse hemodynamic effects of interrupting chest compressions for rescue breathing during cardiopulmonary resuscitation for ventricular fibrillation cardiac arrest. Circulation 2001;104:2465–70.
6. Idris AH, Guffey D, Aufderheide TP, et al. Relationship between chest compression rates and outcomes from cardiac arrest. Circulation 2012;125:3004–12.
7. Rea T, Olsufka M, Yin L, Maynard C, Cobb L. The relationship between chest compression fraction and outcome from ventricular fibrillation arrests in prolonged resuscitations. Resuscitation 2014;85:879–84.
8. Cheskes S, Schmicker RH, Christenson J, et al. Perishock pause: an independent predictor of survival from out-of-hospital shockable cardiac arrest. Circulation 2011;124:58–66.
9. Idris AH, Guffey D, Pepe PP, et al. Chest compression rates and survival following out-of-hospital cardiac arrest. Crit Care Med 2015.
10. Wang HE, Kupas DF, Greenwood MJ, et al. An algorithmic approach to prehospital airway management. Prehosp Emerg Care 2005;9:145–55.
11. Wang HE, Simeone SJ, Weaver MD, Callaway CW. Interruptions in cardiopulmonary resuscitation from paramedic endotracheal intubation. Ann Emerg Med 2009;54:645–52 e641.
12. Kurz MC, Prince D, Christenson J, et al. Supraglottic airway use is associated with higher chest compression fraction than endotracheal intubation during out-of-hospital cardiopulmonary arrest (abstract). Circulation 2014;130:A277.
13. Kurz MC, Prince DK, Christenson J, et al. Association of advanced airway device with chest compression fraction during out-of-hospital cardiopulmonary arrest. Resuscitation 2016;98:35–40.
14. Aramendi E, Lu Y, Chang MP, et al. A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation. Resuscitation 2018;132:41–6.
15. Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics. Resuscitation 2015;88:28–34.
16. Aramendi E, Ayala U, Irusta U, Alonso E, Eftestol T, Kramer-Johansen J. Suppression of the cardiopulmonary resuscitation artefacts using the instantaneous chest compression rate extracted from the thoracic impedance. Resuscitation 2012;83:692–8.
17. Alonso E, Gonzalez-Otero D, Aramendi E, et al. Can thoracic impedance monitor the depth of chest compressions during out-of-hospital cardiopulmonary resuscitation? Resuscitation 2014;85:637–43.
18. Ayala U, Eftestol T, Alonso E, et al. Automatic detection of chest compressions for the assessment of CPR-quality parameters. Resuscitation 2014;85:957–63.
19. Kramer-Johansen J, Edelson DP, Losert H, Kohler K, Abella BS. Uniform reporting of measured quality of cardiopulmonary resuscitation (CPR). Resuscitation 2007;74:406–17.
20. Vaillancourt C, Everson-Stewart S, Christenson J, et al. The impact of increased chest compression fraction on return of spontaneous circulation for out-of-hospital cardiac arrest patients not in ventricular fibrillation. Resuscitation 2011;82:1501–7.
21. Stiell IG, Brown SP, Nichol G, et al. What is the optimal chest compression depth during out-of-hospital cardiac arrest resuscitation of adult patients? Circulation 2014;130:1962–70.
22. Chamberlain D, Cummins RO. Recommended guidelines for uniform reporting of data from out-of-hospital cardiac arrest: the 'Utstein style'. European Resuscitation Council, American Heart Association, Heart and Stroke Foundation of Canada and Australian Resuscitation Council. Eur J Anaesthesiol 1992;9:245–56.
23. Deakin CD, Nolan JP, Ji C, et al. The effect of airway management on CPR quality in the PARAMEDIC2 randomised controlled trial. Resuscitation 2020;158:8–13.
24. Deakin CD, Nolan JP, Ji C, et al. The effect of airway management on CPR quality in the PARAMEDIC2 randomised controlled trial. Resuscitation.
25. Wang HE, Kupas DF, Paris PM, Bates RR, Yealy DM. Preliminary experience with a prospective, multi-centered evaluation of out-of-hospital endotracheal intubation. Resuscitation 2003;58:49–58.
26. Wang HE, Yealy DM. How many attempts are required to accomplish out-of-hospital endotracheal intubation? Acad Emerg Med 2006;13:372–7.
27. Driver BE, Prekker ME, Klein LR, et al. Effect of use of a bougie vs endotracheal tube and stylet on first-attempt intubation success among patients with difficult airways undergoing emergency intubation: a randomized clinical trial. JAMA 2018;319:2179–89.
28. Jabre P, Combes X, Leroux B, et al. Use of gum elastic bougie for prehospital difficult intubation. Am J Emerg Med 2005;23:552–5.
29. Okamoto H, Goto T, Wong ZSY, et al. Comparison of video laryngoscopy versus direct laryngoscopy for intubation in emergency department patients with cardiac arrest: a multicentre study. Resuscitation 2019;136:70–7.
30. Mort TC, Braffett BH. Conventional versus video laryngoscopy for tracheal tube exchange: glottic visualization, success rates, complications, and rescue alternatives in the high-risk difficult airway patient. Anesth Analg 2015;121:440–8.
31. Hubble MW, Brown L, Wilfong DA, Hertelendy A, Benner RW, Richards ME. A meta-analysis of prehospital airway control techniques: Part I. Orotracheal and nasotracheal intubation success rates. Prehosp Emerg Care 2010;14:377–401.

A.3.2 Second journal paper

	Publication in international journal
Reference	Henry E. Wang, Xabier Jaureguibeitia, Elisabete Aramendi, Graham Nichol, Mohamud R. Daya, Matthew Hansen, Michelle Nassal, Ashish R. Panchal, Dhimitri A. Nikolla, Erik Alonso, Jestin Carlson, Robert H. Schmicker, Shannon W. Stephens, Unai Irusta, Ahamed Idris, "Airway strategy and ventilation rates in the Pragmatic Airway Resuscitation Trial", *Resuscitation 2022*, vol. 176, pp. 80-87.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (2/32) based on Web of Science Rank 2022 • **Impact factor**: 6.500

RESUSCITATION 176 (2022) 80–87

Available online at ScienceDirect

Resuscitation

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

Clinical paper

Airway strategy and ventilation rates in the pragmatic airway resuscitation trial

Henry E. Wang[a,*], *Xabier Jaureguibeitia*[b], *Elisabete Aramendi*[b], *Graham Nichol*[c], *Tom Aufderheide*[d], *Mohamud R. Daya*[e], *Matthew Hansen*[e], *Michelle Nassal*[a], *Ashish R. Panchal*[a], *Dhimitri A. Nikolla*[f], *Erik Alonso*[b], *Jestin Carlson*[g], *Robert H. Schmicker*[c], *Shannon W. Stephens*[h], *Unai Irusta*[b,†], *Ahamed Idris*[i]

Abstract

Background: We sought to describe ventilation rates during out-of-hospital cardiac arrest (OHCA) resuscitation and their associations with airway management strategy and outcomes.

Methods: We analyzed continuous end-tidal carbon dioxide capnography data from adult OHCA enrolled in the Pragmatic Airway Resuscitation Trial (PART). Using automated signal processing techniques, we determined continuous ventilation rates for consecutive 10-second epochs after airway insertion. We defined hypoventilation as a ventilation rate < 6 breaths/min. We defined hyperventilation as a ventilation rate > 12 breaths/min. We compared differences in total and percentage post-airway hyper- and hypoventilation between airway interventions (laryngeal tube (LT) vs. endotracheal intubation (ETI)). We also determined associations between hypo-/hyperventilation and OHCA outcomes (ROSC, 72-hour survival, hospital survival, hospital survival with favorable neurologic status).

Results: Adequate post-airway capnography were available for 1,010 (LT n = 714, ETI n = 296) of 3,004 patients. Median ventilation rates were: LT 8.0 (IQR 6.5–9.6) breaths/min, ETI 7.9 (6.5–9.7) breaths/min. Total duration and percentage of post-airway time with hypoventilation were similar between LT and ETI: median 1.8 vs. 1.7 minutes, p = 0.94; median 10.5% vs. 11.5%, p = 0.60. Total duration and percentage of post-airway time with hyperventilation were similar between LT and ETI: median 0.4 vs. 0.4 minutes, p = 0.91; median 2.1% vs. 1.9%, p = 0.99. Hypo- and hyperventilation exhibited limited associations with OHCA outcomes.

Conclusion: In the PART Trial, EMS personnel delivered post-airway ventilations at rates satisfying international guidelines, with only limited hypo- or hyperventilation. Hypo- and hyperventilation durations did not differ between airway management strategy and exhibited uncertain associations with OCHA outcomes.

Keywords: Ventilation, Cardiopulmonary arrest, Airway management, Intubation, Emergency medical services

* Corresponding author at: Department of Emergency Medicine, The Ohio State University, 376 W. 10th Ave, 760 Prior Hall, Columbus, OH 43210, United States.

E-mail addresses: henry.wang@osumc.edu (H.E. Wang), xabier.jaureguibeitia@ehu.eus (X. Jaureguibeitia), elisabete.aramendi@ehu.eus (E. Aramendi), nichol@uw.edu (G. Nichol), TAufderh@mcw.edu (T. Aufderheide), dayam@ohsu.edu (M.R. Daya), hansemat@ohsu.edu (M. Hansen), Michelle.Nassal@osumc.edu (M. Nassal), Ashish.Panchal@osumc.edu (A.R. Panchal), erik.alonso@ehu.es (E. Alonso), jcarlson@ahn-emp.com (J. Carlson), rschmick@uw.edu (R.H. Schmicker), swstephens@uabmc.edu (S.W. Stephens), Ahamed.Idris@utsouthwestern.edu (A. Idris).

† This manuscript is dedicated to the memory of our dear colleague Unai Irusta.

https://doi.org/10.1016/j.resuscitation.2022.05.008

Received 25 March 2022; Received in Revised form 10 May 2022; Accepted 12 May 2022

Introduction

Each year over 350,000 persons in the United States experience sudden out-of-hospital cardiac arrest (OHCA).[1] Optimal delivery of oxygen and control of ventilation are important components of OHCA resuscitation and the main motivations for airway management interventions such as bag-valve-mask ventilation, intubation and supraglottic airway insertion.[2] Inadequate ventilation can result in hypoxemia, hypercapnia, acidemia, alveolar atelectasis and pulmonary shunting.[2–3] However, excessive ventilation can also be harmful in OHCA, increasing intrathoracic pressure, and decreasing venous return, cardiac output and coronary perfusion.[4–6].

Prior studies characterizing ventilation parameters such as ventilation rate and end-tidal carbon dioxide (ETCO2) levels have relied upon discrete measurements or manual assessment of resuscitation records.[4] Modern portable cardiac monitors enable continuous real-time recordings of resuscitation parameters such as thoracic impedance, chest compression depth and ETCO2 levels.[7–13] We previously demonstrated the utility of advanced automated signal processing techniques in characterizing chest compressions delivered to OHCA patients.[14–16] There have been few prior studies using similar automated methods to characterize ventilations delivered during OHCA.

The Pragmatic Airway Resuscitation Trial (PART) found improved OHCA outcomes with an airway strategy of initial laryngeal tube (LT) insertion compared with endotracheal intubation (ETI).[17] Different airway management techniques may potentially influence ventilation performance, including ventilation rate control. In this study, we sought to determine the association between airway strategy and ventilation rates in the PART trial. We also sought to determine the association between ventilation rates and OHCA outcomes.

Methods

Study design

We conducted a secondary analysis of data from the PART trial.[17] The Institutional Review Boards of participating institutions approved the parent study under federal regulations for Exception from Informed Consent for Emergency Research (21 CFR 50.24). This post hoc analysis was approved by the Ohio State University Office of Responsible Research Practices.

Setting

The objective of the PART trial was to compare the effect of airway management strategies (initial LT vs. initial ETI) upon adult OHCA outcomes.[17] The PART trial involved 27 emergency medical services (EMS) agencies from the Birmingham (Alabama), Dallas-Fort Worth (Texas), Milwaukee (Wisconsin), Pittsburgh (Pennsylvania) and Portland (Oregon) communities of the Resuscitation Outcomes Consortium (ROC). The trial enrolled subjects over a 2-year period and found improved 72-hour survival, hospital survival and hospital survival with favorable neurologic outcome with the initial LT strategy.

Selection of participants

The parent trial included adult OHCAs $\geq$ 18 years (or per local interpretation) requiring advanced airway management or bag-valve-mask (BVM) ventilation. Key exclusion criteria included patients < 18 years, pregnant women, prisoners, traumatic cardiac arrest and the initial presence of a non-study advanced EMS unit. The trial enrolled patients from December 1, 2015 through November 4, 2017. For this *post hoc* analysis we included only subjects with continuous capnography data.

Interventions

The interventions of the parent trial were initial airway management strategies using either LT or ETI, assigned in cluster-crossover fashion with each EMS agency alternating between interventions at 3–5 month intervals. Per intention-to-treat principles, patients receiving bag-valve-mask ventilation only were retained in their assigned treatment groups. However, because of our interest in ventilation performance after airway insertion, we assessed the results on an as-treated basis in this secondary analysis, classifying patients according to the final deployed airway device.

Methods of measurement – Analysis of capnography data

We identified and characterized ventilations using the capnography signal from CPR process files, including patients enrolled in the trial. As part of standard clinical care, the participating EMS agencies recorded CPR process data using portable cardiac monitors manufactured by Physio-Control, Inc. (Life-Pak 15 series, Physio-Control, Redmond Washington), Zoll, Inc. (X-series, Zoll, Inc., Chelmsford, Massachusetts) and Philips, Inc. (MRx series, Philips Healthcare, Andover, Massachusetts). In addition to identification of chest compressions through accelerometers (Zoll and Philips) or changes in electrical impedance (Physio-Control and Philips), the monitors also incorporate side-stream sensors for continuously recording end-tidal capnography.[15,18–20] EMS personnel used ETCO2 sensors for confirming airway placement and monitoring resuscitation according to local EMS protocols.

We used a previously validated algorithm for automated import and analysis of capnography files and detection of ventilations delivered during chest compresions.[20] The algorithm detects ventilations based upon comparisons of features of the four phases of the capnogram signal, including the a) duration of inspiration, b) mean baseline inspiration CO2, c) mean expiratory plateau CO2, d) area of the first section of the expiratory plateau, and e) relative CO2 increase. Import and analysis of capnography signals were accomplished using MATLAB (Mathworks, Inc., Natick, MA) and a custom graphical user interface (GUI).[21] The GUI facilitated identification of chest compressions and ventilations, and integration of clinical information such as return of spontaneous circulation and advanced airway type.

Outcomes – Determination of ventilation rates

The primary outcome of this analysis was the ventilation rate delivered a) after advanced airway insertion, and b) during cardiac arrest. We included the time period from airway insertion to the earliest of return of spontaneous circulation (ROSC – inferred by cessation of chest compressions), arrival at hospital or field termination of resuscitation. We excluded the time periods before advanced airway insertion and after attainment of ROSC. We further constrained the analysis to cases with a) $\geq$ 3 minutes of available post-advanced airway capnography data and b) $\geq$ 50% interpretable capnography data.

We determined ventilation rates in consecutive 10-second time epochs. (Fig. 1) We defined ventilation rate as the number of ventilations delivered during the prior 60 seconds. If a full 60-second per-

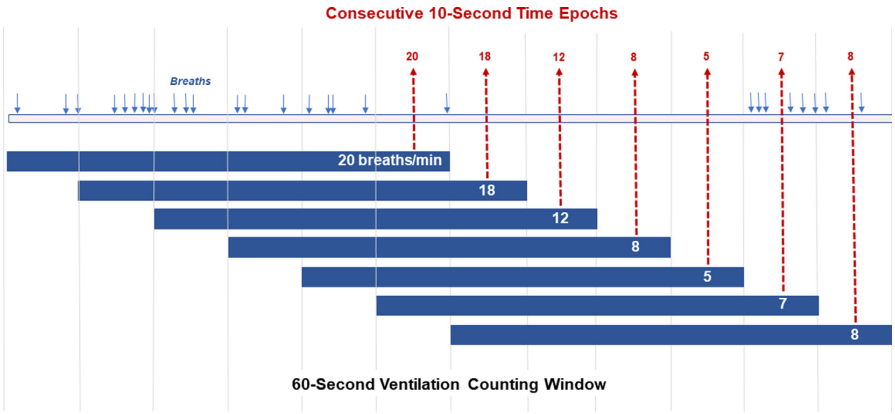


Fig. 1 – Ventilation counting strategy. Ventilation rate for each consecutive 10-second epoch determined by the number of breaths delivered in the preceding 60-seconds.

iod was not available before the time epoch (for example, at the point of airway insertion), we determined ventilation rate from the available data period.

Current resuscitation guidelines recommend delivery of ventilations at 8–10 breaths/min.[22] While the terms "hypo-" and "hyperventilation" formally refer to insufficient or excessive minute ventilation, these terms commonly connote low and high ventilation rates during attempted resuscitation.[2,23] Therefore, for this analysis we defined "hypoventilation" as a low ventilation rate < 6 breaths/min, and "hyperventilation" as a high ventilation rate > 12 breaths/min. We further classified hyperventilation as mild (>12 to 16 breaths/min), moderate (>16 to 20 breaths/min) and severe (>20 breaths/min). We emphasize that the capnography-based analysis of this study only estimated the delivery of individual breaths, not the tidal or minute volume.

For the outcomes analysis, we used the endpoints of the parent trial which were ROSC, 72-hour survival, hospital survival and hospital survival with favorable neurologic outcome.

Data analysis

We compared the median ventilatory rates between LT and ETI groups. We determined the total duration of hypoventilation and hyperventilation (including mild, moderate and severe) by summing the 10-second epochs with hypo- or hyperventilation. We also determined the percentage of post-airway resuscitation period with hypo- and hyperventilation. We compared the duration and percentage and hypo- and hyperventilation exposure between LT and ETI groups. We verified the non-normality of the distributions of ventilation rate, hypo- and hyperventilation duration, and hypo- and hyperventilation percentage using histograms and the Shapiro-Wilk test. We compared differences between the LT and ETI groups using the Wilcoxon Rank Sum test.

For the outcomes analysis, we fit a series of Generalized Estimating Equation (GEE) models.[24] We first fit a model with ROSC as the dependent variable, and duration of hypoventilation and mild, moderate and severe hyperventilation as the primary independent variables. We modeled the duration of hypo- and hyperventilation as continuous variables. We adjusted the models for age, sex, witnessed arrest (bystander or EMS), bystander CPR, initial ECG rhythm (shockable vs. non-shockable), public location, treatment group (LT vs. ETI) and duration of post-airway resuscitation. We implemented the GEE models with robust standard errors and exchangeable correlation structure to account for grouping by randomization cluster.[17] We repeated the analysis for the outcomes 72-hour survival, hospital survival and hospital survival with favorable neurologic outcome.

In a sensitivity analysis we repeated the outcomes analysis modeling duration of hypo- and hyperventilation as categorical variables (0.00–1.99, 2–3.99 and $\geq$ 4.00 minutes). We conducted the analyses using MATLAB (MathWorks Inc., Natick, MA) and Stata v.17.0 (Stata, Inc., College Station, TX).

Results

Of 3,004 patients enrolled in the parent trial, CPR process data were available for 2,020 patients, and capnography data of sufficient quality were available for n = 1,010, including 538 (53%) from Philips monitors, 436 (43.2%) from Zoll monitors, and 36 (3.6%) from PhysioControl monitors. (Fig. 2) Of the 13 randomization clusters, 1 cluster had only 3 cases with suitable capnography data, and 3 clusters had no cases with suitable capnography data. Compared with those excluded, cases included in the analysis were more likely to include white race, unwitnessed arrests, bystander automated external defibrillator use, and initial non-shockable rhythms (Appendix 1). Compared with excluded cases, patients included in the analysis exhibited lower ROSC (17.5% vs. 30.5%), 72-hour survival (8.3% vs. 21.2%), hospital survival (3.7% vs. 12.4%), and hospital survival with favorable neurologic status (1.6% vs. 8.1%).

Stratified by last airway inserted, ventilation data were available for n = 714 LT and n = 296 ETI. Mean age was slightly higher for ETI than LT (Table 1). There were more witnessed arrests and bystander AED use in the ETI than LT group. Sex, race, bystander chest compressions, initial rhythm and clinical outcomes were similar between treatment groups.

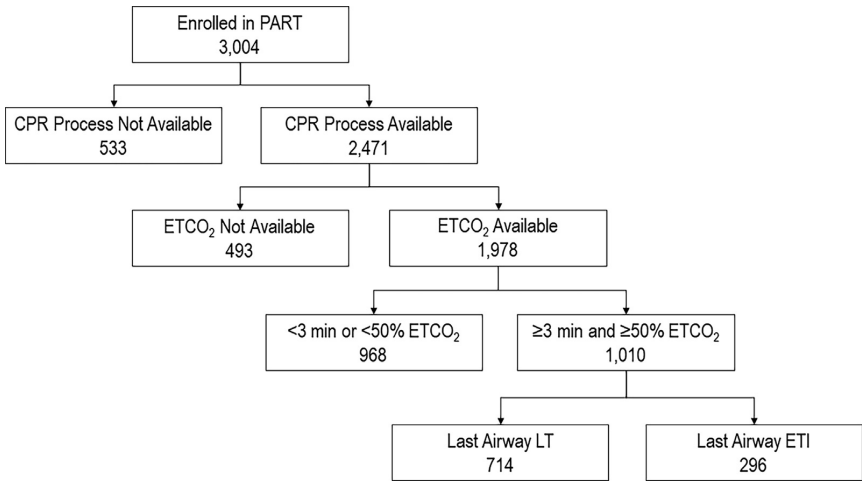


Fig. 2 – Patients included in the analysis. Analysis limited to patient with available ≥ 3 minutes and ≥ 50% interpretable capnography data. CPR=cardiopulmonary resuscitation, LT=laryngeal tube, ETI=endotracheal intubation, ETCO2=end-tidal cabon dioxide, PART=Pragmatic Airway Resuscitation Trial.

The duration of available post-airway capnography data was similar between groups (Table 2, Appendix 2). Ventilatory rate was similar between LT and ETI (Appendix 3). Of the 1,010 patients included in the analysis cohort, 74.1% experienced at least one episode of hypoventilation, and 55.3% experienced at least one episode of hyperventilation. Hypoventilation occurred in approximately 11% of the post-airway period. Median duration of hypoventilation was similar between LT and ETI. (Appendix 4). Hyperventilation occurred in approximately 2% of the post-airway period. Median duration of total, mild, moderate and severe hyperventilation were similar between LT and ETI (Appendices 5–6).

Associations between the durations of hypoventilation, and mild, moderate and severe hyperventilation varied across OHCA outcomes. The duration of hypoventilation was negatively associated with adjusted odds of ROSC, 72-hour survival and hospital survival but not hospital survival with favorable neurologic status (Table 3). Duration of mild hyperventilation exhibited positive associations with ROSC, hospital survival and hospital survival with favorable neurologic status. Duration of moderate hyperventilation was positively associated with 72-hour survival but not other outcomes. Duration of severe hyperventilation was positively associated with hospital survival but not other outcomes. In the sensitivity analysis hypoventilation exhibited negative associations with 72-hour survival and hospital survival (Appendix 7). Associations between hyperventilation and OHCA outcomes were less certain.

Discussion

Ventilation control plays an important role in OHCA resuscitation, facilitating oxygen delivery and the prevention and treatment of hypoxemia, hypercapnia and acidosis.[2] Our study offers important new perspectives of ventilation rates delivered during OHCA. In this subset from the PART trial, EMS personnel delivered ventilations at 6–12 breaths per minute (a range consistent with international care guidelines) in almost 90% of the post-advanced-airway time epochs.[22] The limited durations of hypo- and hyperventilation did not differ between patients treated with LT and ETI, and exhibited small associations with select OHCA outcomes.

Few studies have characterized ventilation rates delivered during clinical OHCA resuscitation. Aufderheide, et al. first observed the common clinical practice of hyperventilating throughout the entire duration of out-of-hospital CPR resuscitation and highlighted the detrimental effects in pig models of cardiac arrest, finding that increased mean intrathoracic pressure resulting from hyperventilation significantly reduced coronary perfusion pressure and survival.[4] In a concurrent study of 13 humans with OHCA, the authors observed that the mean ventilation rate was 30 ± 3.2 breaths/min, raising awareness that unrecognized and inadvertent, nearly continuous hyperventilation may have detrimental hemodynamic and survival consequences during low flow states such as CPR. In a subsequent cohort of 337 OHCA in Belgium, Vissers, et al. found that the mean ventilation rate was 15.3 breaths/min; a ventilation rate > 10 breaths/min was not associated with ROSC.[25–26].

Our analysis adds to existing knowledge, offering the first study to precisely quantify the duration and proportion of resuscitation time ventilation compliant with international guidelines. Our observations were made possible in part by the application of advanced signal processing techniques, which enabled us to systematically identify and characterize ventilation patterns from the capnography signal.[20] These findings are bolstered by our use of clinical trial data from multiple EMS agencies and a range of cardiac monitor manufacturers.[27] The current effort demonstrates the feasibility of our approach and sets the stage for more complex analyses such as examination of time-varying trends in chest compressions and ventilation rates.

In contrast to prior studies, we observed only limited episodes of hyperventilation representing approximately 2% to assessed resuscitation time.[4] We do not know if these observations would generalize

Table 1 – Characteristics of study population included in the ventilation analysis stratified according to last airway device inserted. *Select patients included in more than one race category.

Characteristics	Laryngeal Tube (n = 714) N (%)	Endotracheal Intubation (n = 296) N (%)	p-value
Age, Mean (SD)	61.4 (17.0)	66.2 (16.1)	<0.001
Sex			
Male	447 (62.6)	178 (60.1)	0.61
Female	266 (37.3)	118 (39.9)	
Unknown	1 (0.1)	0 (0.0)	
Race*			
White	391 (54.8)	159 (53.7)	0.77
Black	181 (25.4)	83 (28.0)	0.38
Asian	15 (2.1)	3 (1.0)	0.24
Pacific Islander	1 (0.1)	0 (0.0)	0.52
Native American	3 (0.4)	1 (0.3)	0.85
Hispanic	45 (6.3)	20 (6.8)	0.79
Other	3 (0.4)	5 (1.7)	0.04
Unknown	79 (11.1)	25 (8.5)	0.21
Witnessed Arrest			
Bystander Witnessed	235 (32.9)	107 (36.2)	0.04
EMS Witnessed	61 (8.5)	35 (11.8)	
Unwitnessed	346 (48.5)	138 (46.6)	
Unknown	72 (10.1)	16 (5.4)	
Bystander chest compressions			
Yes	342 (47.9)	161 (54.4)	0.06
No	372 (52.1)	135 (45.6)	
Bystander automated external defibrillator			
Yes	69 (9.7)	46 (15.5)	0.007
No	645 (90.3)	250 (84.5)	
Initial rhythm			
Shockable	110 (15.4)	48 (16.2)	0.75
Non-shockable	604 (84.6)	248 (83.8)	
Outcomes			
ROSC	125 (17.5)	52 (17.6)	0.98
72-hour survival	63 (8.8)	21 (7.1)	0.37
Hospital survival	28 (3.9)	9 (3.0)	0.50
Hospital survival with favorable neurologic status	13 (1.8)	3 (1.0)	0.35

to other EMS agencies in the US or elsewhere. The reasons for this excellent ventilation performance are unclear; EMS personnel in the PART trial may have gained considerable experience from prior OHCA trials (ROC PRIMED and CCC trials) and may have been attentive to ventilation rate control or had access to real time ventilation feedback.[28–30] The EMS agencies in the trial varied in the use of 30:2 and continuous chest compression ratios, but these practices would be expected to influence ventilation patterns before (not after) advanced airway placement.

While we observed select associations between ventilation rates and OHCA outcomes, we believe that these results should be considered hypothesis generating only. First, the characteristics of cases with available and sufficient ventilation data differed from excluded cases; most notably exhibiting lower rates of ROSC, 72-hour survival, hospital survival and hospital survival with neurologic outcome. Ventilation rates and the frequency of hypo- and hyperventilation did not different between airway interventions and thus may not explain the differences in outcomes seen in the parent trial; however, the current analysis contains data on only one third of patients enrolled in the PART trial. While the varying availability of capnography may have been due to resuscitation time bias (-i.e., ETCO2 may not have been available for cases with early ROSC), it may also have resulted from variations in clinical practices. For example, this analysis largely excluded cases from Physio-Control monitors. The observed associations were inconsistent across OHCA outcomes and sensitivity analyses and in many cases exhibited wide confidence intervals. We measured ventilation rates only; even with strong compliance with recommended ventilation rates, minute volumes may have varied. More studies are needed to verify the veracity of these results.

We note that low ventilation rates (<6 breaths/min) occurred more frequently (10% of time epochs) and seemed consistently associated with poorer OHCA outcomes. In an analysis of adult OHCA assigned to 30:2 chest compressions before advanced airway insertion in the ROC Continuous Chest Compressions Trial, Chang, et al. found that 424 of 560 patients exhibited ventilations in less than 50% of the chest compression pauses, and that ROSC and hospital survival were lower in this subset.[27] While our observations potentially align with the Chang et al. findings, given their low incidence and duration of hypoventilation in the present study, it is difficult to make definitive conclusions linking hypoventilation and outcomes.

Clinicians should exercise care in the interpretating of these results. Our analysis was able to characterize ventilation rate only. There are several other dimensions of ventilation such as tidal vol-

Table 2 – Differences in ventilation between airway strategies (LT vs. ETI). Hypo- and hyperventilation refer to ventilation rate only; the analysis did not measure tidal or minute volume. Based upon last airway device inserted. Includes 714 LT and 296 ETI. *Determined by Wilcoxon Rank-Sum test. LT = Laryngeal tube. ETI = Endotracheal intubation. Distributions depicted in Appendices 2–6.

Time period	LT	ETI	TOTAL	p-value ETI vs. LT*
Post-airway available ETCO2 data – mins, median (IQR)	16.4 (11.0–23.3)	16.7 (10.2–22.7)	16.5 (10.8–23.1)	0.78
Post-airway ventilatory rate - breaths/min, median (IQR)	8.0 (6.5–9.6)	7.9 (6.5–9.7)	8.0 (6.5–9.7)	0.81
Post-airway hypoventilation (<6 breaths/min)				
Total duration – min, median (IQR)	1.8 (0.0–5.6)	1.7 (0.0–6.1)	1.7 (0.0–5.6)	0.94
Percentage of available post-airway time (%)	10.5 (0.0–32.1)	11.5 (0.0–36.6)	10.7 (0.0–33.5)	0.60
Post-airway hyperventilation (>12 breaths/min)				
Total duration – min, median (IQR)	0.4 (0.0–2.4)	0.4 (0.0–2.2)	0.4 (0.0–2.3)	0.91
Percentage of available post-airway time (%)	2.1 (0.0–15.0)	1.9 (0.0–13.1)	2.0 (0.0–14.5)	0.99
Mild post-airway hyperventilation (>12 to 16 breaths/min)				
Total duration – min, median (IQR)	0.3 (0.0–1.8)	0.4 (0.0–1.8)	0.3 (0.0–1.8)	0.57
Percentage of available post-airway time (%)	1.6 (0.0–11.4)	1.7 (0.0–9.3)	1.6 (0.0–11.0)	0.96
Moderate post-airway hyperventilation (>16 to 20 breaths/min)				
Total duration – min, median (IQR) [range]	0.0 (0.0–0.0) [0.0, 18.0]	0.0 (0.0–0.2) [0.0, 9.3]	0.0 (0.0–0.0) [0.0, 18.0]	0.86
Percentage of available post-airway time (%)	0.0 (0.0–0.0) [0.0, 80.1]	0.0 (0.0–1.0) [0.0, 38.9]	0.0 (0.0–0.0) [0.0, 80.1]	0.38
Severe post-airway hyperventilation (>20 breaths/min)				
Total duration – min, median (IQR)	0.0 (0.0–0.0) [0.0, 20.4]	0.0 (0.0–0.0) [0.0, 25.8]	0.0 (0.0–0.0) [0.0, 25.8]	0.37
Percentage of available post-airway time (%)	0.0 (0.0–0.0) [0.0, 100.0]	0.0 (0.0–0.0) [0.0, 80.0]	0.0 (0.0–0.0) [0.0, 100.0]	0.96

Table 3 – Association of duration of post-airway hypo- and hyperventilation with out-of-hospital cardiac arrest outcomes. Analyses based upon Generalized Estimating Equations (GEE) models including all four exposures (duration of hypoventilation, duration of mild hyperventilation, duration of moderate hyperventilation, and duration of severe hyperventilation) as independent variables. Models adjusted for age, sex, witnessed arrest (bystander or EMS), bystander CPR, initial ECG rhythm (shockable vs. non-shockable), public location, treatment group (LT vs. ETI) and duration of post-airway resuscitation. Models implemented with robust standard errors and exchangeable correlation structure accounting for randomization cluster. ORs reflect association of each additional minute of hypo- or hyperventilation with each OHCA outcome. Bolded entries indicate statistically significant associations. Hypo- and hyperventilation refer to ventilation rate only; the analysis did not measure tidal or minute volume. OR = Odds Ratio. CI = Confidence Interval. ROSC = Return of Spontaneous Circulation. MRS = Modified Rankin Scale.

Duration of Hypo- or Hyperventilation	ROSC Adjusted OR (95% CI)	72-Hr Survival Adjusted OR (95% CI)	Hospital Survival Adjusted OR (95% CI)	Favorable Neurologic Status (MRS $\leq$ 3) Adjusted OR (95% CI)
Duration of Hypoventilation (<6 breaths/min) – min	**0.96 (0.94–0.99)**	**0.90 (0.85–0.96)**	**0.79 (0.72–0.87)**	0.83 (0.66–1.05)
Duration of Mild Hyperventilation (>12 to 16 breaths/min) – min	**1.09 (1.04–1.15)**	1.06 (0.90–1.24)	**1.23 (1.07–1.40)**	**1.36 (1.01–1.84)**
Duration of Moderate Hyperventilation (>16 to 20 breaths/min) – min	1.05 (0.92–1.20)	**1.11 (1.002–1.23)**	1.12 (0.98–1.27)	1.08 (0.79–1.47)
Duration of Severe Hyperventilation (>20 breaths/min) – min	1.05 (0.89–1.23)	1.11 (0.9998–1.24)	**1.22 (1.06–1.41)**	1.30 (0.92–1.84)

ume (and resulting minute ventilation) and airway pressure. In fact, hypo- and hyperventilation are more properly characterized by insufficient or excessive minute volume – not just ventilation rate. Measurement of ventilation quality during bag-valve-mask ventilation (prior to advanced airway insertion) is also likely relevant but more difficult given the absence of a closed ventilatory circuit. Technological developments may enable characterization of these parameters; for example, identification of ventilations may be possible from

changes in thoracic impedance.[31] We also did not incorporate ETCO2 values into the current analysis; while potentially useful for gauging the efficacy of resuscitation, ETCO2 values are confounded by chest compression quality, drug administration and pre-existing conditions.[32] Further developments are needed to assess the contribution of these factors to OHCA outcomes. Most importantly, we observed marked differences between cases included in versus excluded from the analysis; future efforts with more detailed collection of data are needed to build on the insights of our study.

Limitations

As discussed previously, this analysis focused on ventilation rate only and could not characterize other key ventilatory parameters such as tidal volume, minute volume and airway pressure. Capnography files of adequate length and quality were available for only one-third of cases enrolled in the PART trial. Few files were available from PhysioControl monitors; we do not know if this was due to practice variation at these EMS agencies or technical issues associated with capturing this information with this specific monitor. Differences in the course of ETI compared to LT placement may have influenced the capture of capnography. Characteristics differed between cases included in and excluded from the analysis; these differences may have resulted from resuscitation time bias. The parent trial did not protocolize ventilation rate or the application or use of capnography. This current analysis did not incorporate information on chest compression quality, which may have influenced ventilation quality.

ETI success rates observed in the trial were lower than those from prior reports; many of these cases were rescued by LT insertion, which may have influenced our observations. We did not have information on the airway management protocols, training protocols or practice patterns across the participating agencies; these variations may have influenced observed ventilation. We relied upon EMS personnel reports to define the start and end of airway management efforts. We focused on intra-arrest ventilation after advanced airway insertion; we did not assess the bag-valve-mask phase of ventilation nor post-ROSC ventilation. We did not adjust for FiO2 or oxygen delivery or consumption or medication administration. EMS providers may have altered care in reaction to measured ETCO2 values.

Conclusions

In this post-hoc analysis of the PART Trial, EMS personnel delivered the majority of post-airway ventilations at rates satisfying international treatment guidelines, with only limited episodes of hypo- and hyperventilation. Duration of hypo- and hyperventilation did not differ with airway management strategy. Duration of hypo- and hyperventilation exhibited uncertain associations with OCHA outcomes. Further research is needed to understand the impact of ventilation in OHCA resuscitation.

Sources of Funding

Research Supported by Grant UH2/UH3-HL125163 from National Heart Lung and Blood Institute. This work was partially supported by the Spanish Ministerio de Ciencia, Innovacion y Universidades through grant RTI2018-101475-BI00, jointly with the Fondo Europeo de Desarrollo Regional (FEDER) and by the Basque Government through grants IT1229-19 and PRE 2020 2 0182.

Conflicts of Interest

The authors declare no conflicts of interest.

CRediT authorship contribution statement

Henry E. Wang: Conceptualization, Data curation, Formal analysis, Investigation, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Writing – original draft, Writing – review & editing. **Xabier Jaureguibeitia:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. **Elisabete Aramendi:** Conceptualization, Data curation, Methodology, Resources, Software, Writing – review & editing, Investigation. **Graham Nichol:** Funding acquisition, Investigation, Writing – review & editing. **Tom Aufderheide:** Investigation, Writing – review & editing. **Mohamud R. Daya:** Investigation, Writing – review & editing. **Matthew Hansen:** Investigation, Writing – review & editing. **Michelle Nassal:** Investigation, Writing – review & editing. **Ashish R. Panchal:** Investigation, Writing – review & editing. **Dhimitri A. Nikolla:** Investigation, Writing – review & editing. **Erik Alonso:** Conceptualization, Data curation, Methodology, Resources, Software, Writing – review & editing. **Jestin Carlson:** Investigation, Writing – review & editing. **Robert H. Schmicker:** Investigation, Writing – review & editing. **Shannon W. Stephens:** Investigation, Writing – review & editing. **Unai Irusta:** Conceptualization, Funding acquisition, Methodology, Software, Writing – review & editing, Investigation. **Ahamed Idris:** Conceptualization, Investigation, Writing – review & editing.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resuscitation.2022.05.008.

Author details

[a]*The Ohio State University, United States* [b]*University of the Basque Country, Spain* [c]*University of Washington, United States* [d]*Medical College of Wisconsin, United States* [e]*Oregon Health and Science University, United States* [f]*Allegheny Health Network – Saint Vincent, United States* [g]*University of Pittsburgh, United States* [h]*University of Alabama at Birmingham, United States* [i]*University of Texas Southwestern Medical Center, United States*

REFERENCES

1. Benjamin EJ, Muntner P, Alonso A, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation 2019;139:e56–e528.

2. Neth MR, Idris A, McMullan J, Benoit JL, Daya MR. A review of ventilation in adult out-of-hospital cardiac arrest. J Am Coll Emerg Physicians Open 2020;1:190–201.
3. Scanlon 3rd TS, Benumof JL, Wahrenbrock EA, Nelson WL. Hypoxic pulmonary vasoconstriction and the ratio of hypoxic lung to perfused normoxic lung. Anesthesiology 1978;49:177–81.
4. Aufderheide TP, Lurie KG. Death by hyperventilation: a common and life-threatening problem during cardiopulmonary resuscitation. Crit Care Med 2004;32:S345–51.
5. Theres H, Binkau J, Laule M, et al. Phase-related changes in right ventricular cardiac output under volume-controlled mechanical ventilation with positive end-expiratory pressure. Crit Care Med 1999;27:953–8.
6. Cheifetz IM, Craig DM, Quick G, et al. Increasing tidal volumes and pulmonary overdistention adversely affect pulmonary vascular mechanics and cardiac output in a pediatric swine model. Crit Care Med 1998;26:710–6.
7. Christenson J, Andrusiek D, Everson-Stewart S, et al. Chest compression fraction determines survival in patients with out-of-hospital ventricular fibrillation. Circulation 2009;120:1241–7.
8. Idris AH, Guffey D, Aufderheide TP, et al. Relationship between chest compression rates and outcomes from cardiac arrest. Circulation 2012;125:3004–12.
9. Stiell IG, Brown SP, Nichol G, et al. What is the optimal chest compression depth during out-of-hospital cardiac arrest resuscitation of adult patients?. Circulation 2014;130:1962–70.
10. Vaillancourt C, Petersen A, Meier EN, et al. The impact of increased chest compression fraction on survival for out-of-hospital cardiac arrest patients with a non-shockable initial rhythm. Resuscitation 2020;154:93–100.
11. Cheskes S, Schmicker RH, Rea T, et al. Chest compression fraction: A time dependent variable of survival in shockable out-of-hospital cardiac arrest. Resuscitation 2015;97:129–35.
12. Cheskes S, Common MR, Byers PA, Zhan C, Morrison LJ. Compressions during defibrillator charging shortens shock pause duration and improves chest compression fraction during shockable out of hospital cardiac arrest. Resuscitation 2014;85:1007–11.
13. Vaillancourt C, Everson-Stewart S, Christenson J, et al. The impact of increased chest compression fraction on return of spontaneous circulation for out-of-hospital cardiac arrest patients not in ventricular fibrillation. Resuscitation 2011;82:1501–7.
14. Wang HE, Jaureguibeitia X, Aramendi E, et al. Airway strategy and chest compression quality in the Pragmatic Airway Resuscitation Trial. Resuscitation 2021;162:93–8.
15. Aramendi E, Ayala U, Irusta U, Alonso E, Eftestol T, Kramer-Johansen J. Suppression of the cardiopulmonary resuscitation artefacts using the instantaneous chest compression rate extracted from the thoracic impedance. Resuscitation 2012;83:692–8.
16. Elola A, Aramendi E, Irusta U, Del Ser J, Alonso E, Daya M. ECG-based pulse detection during cardiac arrest using random forest classifier. Med Biol Eng Comput 2018.
17. Wang HE, Schmicker RH, Daya MR, et al. Effect of a Strategy of Initial Laryngeal Tube Insertion vs Endotracheal Intubation on 72-Hour Survival in Adults With Out-of-Hospital Cardiac Arrest: A Randomized Clinical Trial. JAMA, J Am Med Assoc 2018;320:769–78.
18. Aramendi E, Lu Y, Chang MP, et al. A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation. Resuscitation 2018;132:41–6.
19. Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics. Resuscitation 2014;88C:28–34.
20. Aramendi E, Elola A, Alonso E, et al. Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation. Resuscitation 2017;110:162–8.
21. Jaureguibeitia X, Aramendi E, Irusta U, et al. Methodology and framework for the analysis of cardiopulmonary resuscitation quality in large and heterogeneous cardiac arrest datasets. Resuscitation 2021;168:44–51.
22. Kleinman ME, Goldberger ZD, Rea T, et al. 2017 American Heart Association Focused Update on Adult Basic Life Support and Cardiopulmonary Resuscitation Quality: An Update to the American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2018;137:e7–e13.
23. Panchal AR, Berg KM, Hirsch KG, et al. 2019 American Heart Association Focused Update on Advanced Cardiovascular Life Support: Use of Advanced Airways, Vasopressors, and Extracorporeal Cardiopulmonary Resuscitation During Cardiac Arrest: An Update to the American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2019;140:e881–94.
24. Hardin JW, Hilbe JM. Generalized estimating equations. second ed. Boca Raton, FL: CRC Press; 2013.
25. Vissers G, Duchatelet C, Huybrechts SA, Wouters K, Hachimi-Idrissi S, Monsieurs KG. The effect of ventilation rate on outcome in adults receiving cardiopulmonary resuscitation. Resuscitation 2019;138:243–9.
26. Kalmar AF, Absalom A, Monsieurs KG. A novel method to detect accidental oesophageal intubation based on ventilation pressure waveforms. Resuscitation 2012;83:177–82.
27. Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. Resuscitation 2019;141:174–81.
28. Nichol G, Leroux B, Wang H, et al. Trial of Continuous or Interrupted Chest Compressions during CPR. New Engl J Med 2015;373:2203–14.
29. Stiell IG, Nichol G, Leroux BG, et al. Early versus later rhythm analysis in patients with out-of-hospital cardiac arrest. New Engl J Med 2011;365:787–97.
30. Aufderheide TP, Nichol G, Rea TD, et al. A trial of an impedance threshold device in out-of-hospital cardiac arrest. New Engl J Med 2011;365:798–806.
31. Jaureguibeitia X, Irusta U, Aramendi E, et al. Impedance Based Automatic Detection of Ventilations During Mechanical Cardiopulmonary Resuscitation. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:19–23.
32. Sandroni C, De Santis P, D'Arrigo S. Capnography during cardiac arrest. Resuscitation 2018;132:73–7.

A.3.3 Third journal paper

	Publication in international journal
Reference	Michelle M.J. Nassal, Xabier Jaureguibeitia, Elisabete Aramendi, Unai Irusta, Ashish R. Panchal, Henry E. Wang, Ahamed H. Idris, "Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the Pragmatic Airway Resuscitation Trial", *Resuscitation 2021*, vol. 168, pp. 58-64.
Quality indices	• **Type of publication**: Journal paper indexed in JCR • **Quartile**: Q1 (3/32) based on Web of Science Rank 2021 • **Impact factor**: 6.251

RESUSCITATION 168 (2021) 58–64

Available online at ScienceDirect

Resuscitation

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

Clinical paper

Novel application of thoracic impedance to characterize ventilations during cardiopulmonary resuscitation in the pragmatic airway resuscitation trial

Michelle M.J. Nassal[a,*], *Xabier Jaureguibeitia*[b], *Elisabete Aramendi*[b], *Unai Irusta*[b], *Ashish R. Panchal*[a], *Henry E. Wang*[a], *Ahamed Idris*[c]

[a] Department of Emergency Medicine, The Ohio State University, Columbus, OH, USA
[b] Department of Communication Engineering, BioRes Group, University of the Basque Country, Bilbao, Spain
[c] Department of Emergency Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA

Abstract

Background: Significant challenges exist in measuring ventilation quality during out-of-hospital cardiopulmonary arrest (OHCA) outcomes. Since ventilation is associated with outcomes in cardiac arrest, tools that objectively describe ventilation dynamics are needed. We sought to characterize thoracic impedance (TI) oscillations associated with ventilation waveforms in the Pragmatic Airway Resuscitation Trial (PART).

Methods: We analyzed CPR process files collected from adult OHCA enrolled in PART. We limited the analysis to cases with simultaneous capnography ventilation recordings at the Dallas-Fort Worth site. We identified ventilation waveforms in the thoracic impedance signal by applying automated signal processing with adaptive filtering techniques to remove overlying artifacts from chest compressions. We correlated detected ventilations with the end-tidal capnography signals. We determined the amplitudes (Ai, Ae) and durations (Di, De) of both insuation and exhalation phases. We compared dierences between laryngeal tube (LT) and endotracheal intubation (ETI) airway management during mechanical or manual chest compressions using Mann-Whitney U-test.

Results: We included 303 CPR process cases in the analysis; 209 manual (77 ETI, 132 LT), 94 mechanical (41 ETI, 53 LT). Ventilation Ai and Ae were higher for ETI than LT in both manual (ETI: Ai 0.71 Ω, Ae 0.70 Ω vs LT: Ai 0.46 Ω, Ae 0.45 Ω; $p < 0.01$ respectively) and mechanical chest compressions (ETI: Ai 1.22 Ω, Ae 1.14 Ω VS LT: Ai 0.74 Ω, Ae 0.68 Ω; $p < 0.01$ respectively). Ventilations per minute, duration of TI amplitude insufflation and exhalation did not dier among groups.

Conclusion: Compared with LT, ETI thoracic impedance ventilation insuation and exhalation amplitude were higher while duration did not dier. TI may provide a novel approach to characterizing ventilation during OHCA.

Keywords: Ventilation, Cardiac arrest, Resuscitation, Thoracic impedance

Introduction

Sudden out-of-hospital cardiopulmonary arrest (OHCA) is a major public health problem annually affecting up to 450,000 adults in the United States.[1] Successful OHCA resuscitation requires quality chest compressions and controlled ventilation for effective oxygen delivery.[2] Recognizing that poor ventilation during resuscitation results in adverse outcomes,[3] data that objectively describes ventilation dynamics during OHCA resuscitation are needed.

Currently there is limited portable technology to accurately measure key aspects of ventilation in the prehospital environment, for example, tidal volume, inspiration and expiratory dynamics. End-tidal capnography (ETCO2) is the current available method for characterizing ventilation during resuscitation. However, ETCO2 is lim-

* Corresponding author at: Department of Emergency Medicine, The Ohio State University, 376 W 10th St, 760B Prior Hall, Columbus, OH 43210, USA.

https://doi.org/10.1016/j.resuscitation.2021.08.045
Received 1 July 2021; Received in Revised form 26 August 2021; Accepted 29 August 2021

ited by several factors including recording generally begins after advanced airway placement, it is most reliable after endotracheal intubation (ETI) and it can be effected by perfusion. Alternative methods are necessary to evaluate ventilation dynamics during the entire resuscitation.

Thoracic impedance (TI) is the measured electrical resistance across the thorax. It is typically recorded through chest defibrillator pads during chest wall expansion and contraction. Changes in TI are commonly used to characterize chest compression quality in OHCA but may also potentially reflect ventilation quality.[4,5] Early TI pneumography was developed to record electrocardiogram and ventilations simultaneously.[6,7] Using low frequency signal filtering on modern defibrillators, TI oscillations associated with ventilation waveforms can also be visualized.[8] This offers advantages over ETCO2 as TI recording begins as soon as defibrillator electrodes are placed on the chest. Analysis of TI could provide valuable information on ventilation dynamics.

The Pragmatic Airway Resuscitation Trial (PART) tested different advanced airway techniques in the resuscitation of OHCA. In this evaluation, our objective was to characterize thoracic impedance ventilation waveforms measured during OHCA resuscitation in the PART trial.

Methods

Study design and setting

This study was a secondary analysis of chest compression waveforms from the Resuscitation Outcomes Consortium (ROC) Pragmatic Airway Resuscitation Trial.[9] The PART trial enrolled adults (age $\geq$ 18 years) with nontraumatic out-of-hospital cardiac arrest treated by emergency medical services paramedics. The trial randomized EMS agencies to either of two initial advanced airway management strategies: laryngeal tube (LT) insertion or endotracheal intubation (ETI). The primary outcome was 72-hour survival. The Institutional Review Board of the Ohio State University approved this post-hoc analysis of the parent trial data.

CPR process files

EMS personnel recorded CPR process measures on all OHCA per local protocols. Chest compression measurements occurred using standard commercial portable cardiac monitors using accelerometer or electrical impedance-based detection systems (Physiocontrol, Inc., Redmond, Washington; Philips, Inc, Andover, Massachusetts; and Zoll, Inc., Chelmsford, Massachusetts). All electronic files were downloaded from defibrillators and stored in the central registry.

Selection of subjects

The objective of this analysis was to characterize ventilations using TI. Furthermore, to validate the TI waveforms, we sought cases with simultaneous capnography ventilation recordings. In the parent trial, the only defibrillators able to provide simultaneously recorded continuous TI and capnography were the Phillips MRx defibrillators. Additionally, we also wanted to evaluate differences in TI waveforms during manual and mechanical chest compressions. To satisfy all these needs, and avoid potential cross site confounders, we limited the analysis to cases enrolled at the Dallas-Fort Worth ROC site, where EMS providers used both manual and mechanical chest compressions (LUCAS device, Stryker-Physio Control, Kalamazoo, MI), and Phillips HeartStart MRx defibrillators (Philips Healthcare, Eindhoven).

Phillips HeartStart MRX defibrillator CPR process file collection

The HeartStart MRX defibrillator provides four modes of operation including monitor, manual defibrillation, automatic external defibrillation and pacing. Monitoring mode includes simultaneous recording of pulse oximetry, blood pressure, temperature, ETCO2 through designated input ports. Multifunction electrode pads can simultaneously record ECG waveforms and TI or provide defibrillation/pacing as needed. Multifunction electrode pads are rectangular shaped with rounded corners covering 102 cm^2 each that are placed on the anterior chest (Fig. 1).

Preparation of the CPR process files

Defibrillator files were converted to an open MATLAB (MathWorks Inc., Natick, MA) format using ad-hoc data conversion software and then signals were processed in MATLAB. TI was recorded with a sampling rate of 200 Hz and a resolution of 0.74 mΩ per least significant bit. The capnogram was acquired using Microstream (sidestream) technology, with a sampling rate of 40 Hz and a resolution of 0.004 mmHg. When available, the acceleration, force and compression depth signals from the Q-CPR assist pad were also included.

We used custom tools developed in MATLAB (MathWorks Inc., Natick, MA) to explore defibrillator data and to extract segments, according to the following inclusion criteria: at least 1 min of CPR after airway insertion, containing a minimum of 10 ventilations, and good quality TI waveform and capnography recordings. We defined good quality for intervals where the automated algorithm for ventilation detection provided an accuracy above 80%. We identified delivered chest compressions using a combination of impedance changes and compression depth detected from the accelerometer.[10] We differentiated manual from mechanical CPR by identifying the unique fixed compression rate delivered by the Lucas devices (frequency 101.7 compressions per min). We defined airway insertion time as the time point of the last successful advanced airway insertion, as indicated in the Utstein-style annotations by the EMS personnel.

Fig. 1 – Philips HeartStart MRX Defibrillator Pads Multipurpose defibrillator pads are placed on the anterior chest of adults as shown.

To identify ventilations, we used adaptive filtering techniques to remove impedance waveform artifacts due to chest compressions (Fig. 2). For manual CPR, the filter used the compression frequency information acquired through the accelerometer.[11] For mechanical CPR the chest compression frequency was fixed to that of the Lucas device.[12] We applied a harmonic Least Mean Squares Filter tuned to the frequency of the chest compressions to remove chest compression artifacts. Then, we applied a multi-stage algorithm combining peak detection and machine learning techniques to identify impedance fluctuations associated with each ventilation.[11,12]

We used capnography data to validate the detection by the impedance-based algorithm. Capnography ventilations were automatically annotated using a state-of-the-art algorithm[13] but manually reviewed. The time-delay of the capnogram due to air transport was also manually corrected in each case, so that ventilations in the capnogram would be time-aligned with the impedance ventilation waveform (Fig. 2). Annotations served as the gold standard to audit the ventilation peaks and the insufflation/exhalation phases. After quality-audited impedance waveforms were obtained, we computed the amplitudes (Ai, Ae) and durations (Di, De) of both the insufflation and exhalation phases (Fig. 2).

Data analysis

We determined median values of the amplitudes (Ai, Ae) and durations (Di, De) for each analysis interval and assigned them to either the ETI group or the LT group. Distributions are represented as medians (w/ IQR). We assessed differences between TI in ETI vs LT using a Mann-Whitney U-test. Demographic differences between groups were assessed using Chi squared tests. Tests were separately analyzed for manual and mechanical chest compressions. P-values < 0.01 were considered statistically significant.

Results

The PART trial enrolled 3004 patients from December 1st 2015 through November 4 2017. CPR process data files were available for 1996 participants. Dallas Fort-Worth included 1018 CPR process files available for analysis. Only CPR process files collected by Phillips MRx defibrillators with successful advanced airway placement were selected for analysis resulting in 692 CPR process data files. After quality assessment, we included 209 manual and 94 mechanical chest compression CPR process files (Fig. 3).

Manual chest compression CPR process files included 77 ETI and 132 LT cases. Mechanical chest compression CPR process files included 41 ETI and 53 LT cases. Subjects who had endotracheal tube devices were significantly older than subjects with laryngeal tube devices in the manual chest compression group; but not different in the mechanical chest compression group. White race was not different between airway devices during manual chest compression but was different between airway devices during mechanical chest compressions. Sex, minority race, witnessed arrest, and initial rhythm were not different among airway devices in either manual or mechanical chest compressions (Table 1).

CPR defibrillator files were analyzed using filters to remove chest compression artifacts, resulting in automatically identified TI ventilation waveforms (Fig. 2). Filtering techniques differed for manual and mechanical chest compressions. TI ventilations appropriately aligned with corresponding capnography signals. The amplitude and duration of insufflation and exhalation phases were also automatically annotated but manually verified (Fig. 2).

TI amplitude of insufflation (Ai) and exhalation (Ae) were higher for subjects with placement of ETI compared to LT in both manual and mechanical chest compressions (Fig. 4). Summary values for TI waveform dynamics in manual compressions (ETI: Ai 0.71, Ae 0.70 vs LT: Ai 0.46, Ae 0.45; $p < 0.01$ respectively) and mechanical compression (ETI: Ai 1.22, Ae 1.14 VS LT: Ai 0.74, Ae 0.68; $p < 0.01$ respectively) are shown in Table 2. Ventilations per minute, duration of TI amplitude insufflation and exhalation did not differ among groups.

Discussion

This is one of the first efforts to characterize pre-hospital ventilations using TI during cardiac arrest.[14] We found there are important contrasts in TI that vary between airway devices and mode of chest compression. ETI consistently had increased ventilation waveform amplitude during both insufflation and exhalation compared with LT under both mechanical and manual chest compressions. In contrast,

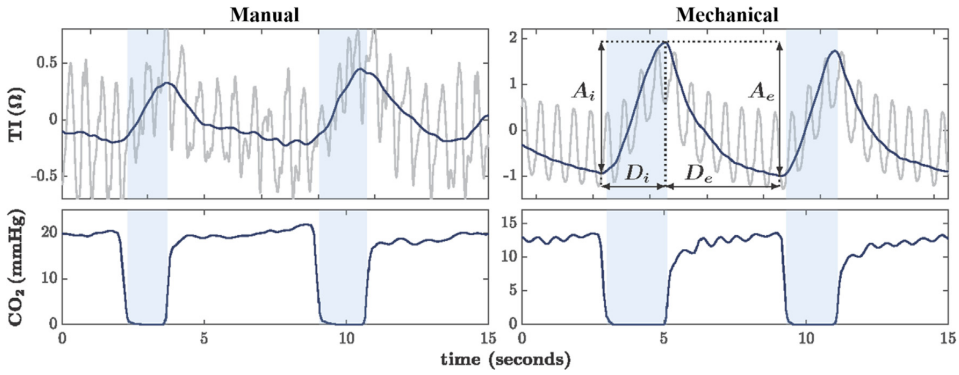


Fig. 2 – Bioimpedance waveform fluctuations during ventilations in manual (left) and mechanical (right) chest compressions with corresponding end-tidal capnography waveforms. Blue line represents filtered thoracic bioimpedance (TI) signal. Grey line represents raw TI signal. Blue shaded area denotes insufflation phase. Ai = amplitude of insufflation Ae = amplitude of exhalation Di = duration of insufflation De = duration of exhalation.

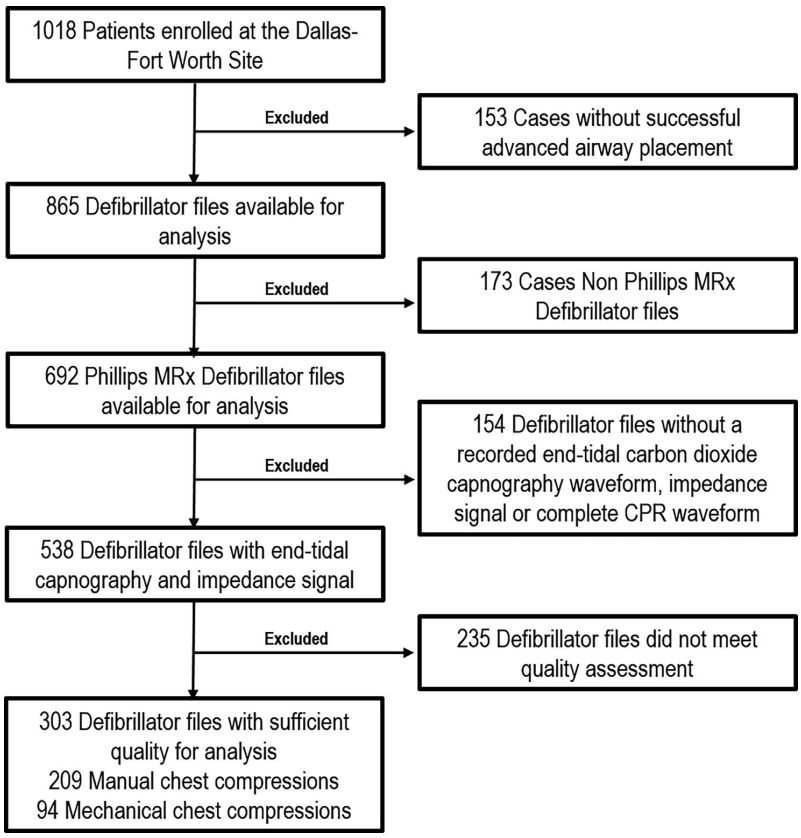


Fig. 3 – Study cohort and exclusions.

Table 1 – Baseline characteristics for manual and mechanical chest compressions. ETI = endotracheal intubation, LT = laryngeal tube, IQR = interquartile range.

	Manual chest compressions			Mechanical chest compressions		
	ETI	LT		ETI	LT	
	N (%) 77 (36.8%)	N (%) 132 (63.2%)	p	N (%) 41 (43.6%)	N (%) 53 (56.4%)	p
Age (years with IQR)	75 (61-85)	66 (54-77)	<0.01	65 (53-75)	66 (52-75)	0.84
Gender (male)	40 (52.0%)	69 (52.3%)	0.92	28 (68.3%)	33 (62.2%)	0.70
Race						
White	38 (49.4%)	75 (56.8%)	0.37	15 (36.6%)	34 (64.2%)	0.01
Black	28 (36.4%)	38 (28.8%)	0.33	16 (39.0%)	11 (20.8%)	0.09
Hispanic	5 (6.5%)	9 (6.8%)	0.84	6 (14.6%)	2 (3.8%)	0.13
Asian	0 (0%)	3 (2.3%)	0.47	1 (2.4%)	0 (0%)	0.90
Other	2 (2.6%)	0 (0%)	0.26	2 (4.9%)	0 (0%)	0.37
Unspecified	4 (5.2%)	7 (5.3%)	0.77	1 (2.4%)	6 (11.3%)	0.22
Witnessed Arrest						
Bystander	33 (42.9%)	46 (34.9%)	0.32	16 (39%)	24 (45.3%)	0.69
EMS	8 (10.4%)	11 (8.3%)	0.8	3 (7.3%)	4 (7.6%)	0.72
Not witnessed	36 (46.7%)	75 (56.8%)	0.16	22 (53.7%)	25 (47.2%)	0.53
Initial Rhythm						
Shockable	6 (7.8%)	14 (10.6%)	0.50	7 (17.1%)	8 (15.2%)	0.80
Nonshockable	71 (88.3%)	118 (87.1%)		34 (82.9%)	45 (83.0%)	

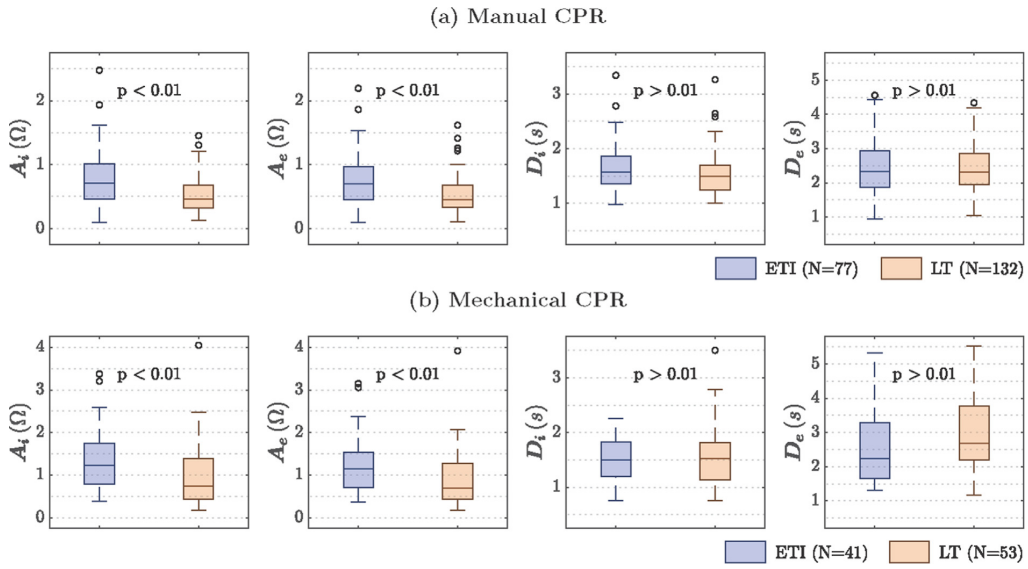


Fig. 4 – Bioimpedance waveform fluctuations during ventilation in manual chest compressions (top) and mechanical chest compressions (bottom). Ai = amplitude of insufflation Ae = amplitude of exhalation Di = duration of insufflation De = duration of exhalation. Values are depicted using standard box and whisker plots with median with interquartile ranges. ETI, Purple boxes = endotracheal tube devices. LT, Orange boxes = laryngeal tube airway devices. Outliers are represented by open circles. P-values are denoted.

Table 2 – Summary data for patients whom received manual and mechanical chest compression. ETI = endotracheal intubation, LT = laryngeal tube, IQR = interquartile ranges, Ai = amplitude of insufflation, Ae = amplitude of exhalation, Di = duration of insufflation, De = duration of exhalation. VR = ventilation rates per minute.

	Manual chest compressions			Mechanical chest compressions		
	ETI (n = 77) Median (IQR)	LT (n = 132) Median (IQR)	p	ETI (n = 41) Median (IQR)	LT (n = 53) Median (IQR)	p
Ai (Ω)	0.71 (0.47–1.01)	0.46 (0.32–0.68)	<0.01	1.22 (0.78–1.74)	0.74 (0.44–1.39)	<0.01
Ae (Ω)	0.70 (0.45–0.97)	0.45 (0.34–0.67)	<0.01	1.14 (0.70–1.53)	0.68 (0.43–1.27)	<0.01
Di (s)	1.57 (1.36–1.86)	1.50 (1.24–1.70)	0.03	1.5 (1.19–1.82)	1.52 (1.14–1.82)	0.96
De (s)	2.33 (1.87–2.94)	2.32 (1.95–2.86)	0.78	2.23 (1.66–3.29)	2.68 (2.21–3.77)	0.15
VR (vpm)	8.3 (6.9–9.3)	8.5 (6.9–10.2)	0.35	6.4 (5.2–8.0)	7.3 (5.3–8.7)	0.31

duration of insufflation and exhalation were not significantly different. Our analysis illustrates the potential usefulness of TI in characterizing ventilations during cardiac arrest.

Understanding ventilations in cardiac arrest is important because ventilation affects hemodynamics during CPR,[15,16] poor ventilations result in worse outcomes,[3] and ventilation metrics can guide resuscitation decisions.[17] Objectively measuring the rate, tidal volume and pattern of inspiratory and expiratory ventilation during OHCA can provide necessary quality metrics for resuscitation improvement. We have not yet validated TI as an accurate surrogate of these measurements. Our next efforts are to understand these real-world measurements and compare them to ventilations in controlled conditions.

Few studies have evaluated ventilation during OHCA.[5,18,19] Active chest compressions further reduce the accuracy of measurements with prior methods.[18] Previously, we developed a novel method of characterizing TI ventilation quality metrics during chest compression pauses.[14] Using chest compression filtering techniques, we now further developed a method to include ventilation quality metrics during active chest compressions. This resulted in filtered TI waveform amplitudes that are not comparable to prior studies or between manual and mechanical chest compressions. However, the proportional relationships between airway devices remain true. We also simultaneously included capnography for ventilation measurements which allowed for ventilations with amplitudes < 0.5 Ω to be included in our analysis.

The translational effect of our findings has yet to be determined. This study raises specific new hypotheses that need to be evaluated further defining if our observed differences in ventilation amplitude are affected by tidal volume, airway leak, or dead space. Our observed differences in ventilation amplitude may contribute to the observed survival difference reported in the PART trial.[9] Prior studies did not find differences in time to epinephrine,[20] chest compression fraction or rate.[21] Although differences in chest compression interruptions[21] and first pass success[9] have been reported, it is unclear what defining characteristics contributed to the PART outcomes.

Limitations

Our study only included Phillips MRx defibrillators as these are the only defibrillators with simultaneous TI and capnography in PART. We were able to leverage the use of simultaneous capnography recordings to allow for inclusion of lower amplitude TI waveforms. Defibrillators made by other manufacturers may record TI waveforms with different amplitudes that are not comparable among devices. For example, Lifepak defibrillators have pre-filtered TI waveforms, resulting in amplitudes that differ from those in the present study. Further, we describe the two most common advanced airway interventions, ETI and LT, but it is unknown if bag-valve-mask ventilation results in differences in TI. Finally, our studies are exploratory, future investigations must define the clinical meaning and applicability of our results. Further work must develop fully automatic TI analysis for use in real-time where more potential noise to signal may exist in an OHCA resuscitation.

Conclusions

In this study, we characterized ventilation TI waveforms during OHCA resuscitation. TI characterization may provide valuable ventilation dynamics during resuscitation efforts.

CRediT authorship contribution statement

Michelle M.J. Nassal: Conceptualization, Visualization. **Xabier Jaureguibeitia:** Methodology, Software, Formal analysis, Validation. **Elisabete Aramendi:** Methodology, Software, Validation, Formal analysis. **Unai Irusta:** . **Ashish Panchal:** Conceptualization, Supervision. **Henry E. Wang:** Conceptualization, Supervision, Project administration. **Ahamed Idris:** Conceptualization, Supervision, Project administration.

Acknowledgements

Supported by grant UH2/UH3-HL125163 from the National Heart Lung and Blood Institute.

This work was partially supported by the Spanish Ministerio de Ciencia, Innovacion y Universidades through grant RTI2018-101475-BI00, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), by the Basque Government through grants IT1229-19 and PRE 2020 2 0182, and by the University of the Basque Country (UPV/EHU) under grant COLAB20/01.

REFERENCES

1. Kong MH, Fonarow GC, Peterson ED, et al. Systematic review of the incidence of sudden cardiac death in the United States. J Am Coll Cardiol 2011;57:794–801.
2. Idris AH, Guffey D, Pepe PE, et al. Chest compression rates and survival following out-of-hospital cardiac arrest. Crit Care Med 2015;43:840–8.
3. Aufderheide TP, Lurie KG. Death by hyperventilation: a common and life-threatening problem during cardiopulmonary resuscitation. Crit Care Med 2004;32:S345–51.
4. Ayala U, Eftestøl T, Alonso E, et al. Automatic detection of chest compressions for the assessment of CPR-quality parameters. Resuscitation 2014;85:957–63.
5. Losert H, Risdal M, Sterz F, et al. Thoracic impedance changes measured via defibrillator pads can monitor ventilation in critically ill patients and during cardiopulmonary resuscitation. Crit Care Med 2006;34:2399–405.
6. Geddes LA et al. Recording respiration and the electrocardiogram with common electrodes. Aerosp Med 1962;33:791–3.
7. Geddes LA et al. The impedance pneumography. Aerosp Med 1962;33:28–33.
8. Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. Resuscitation 2019;141:174–81.
9. Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. JAMA 2018;320:769. https://doi.org/10.1001/jama.2018.7044.
10. Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics. Resuscitation 2015;88:28–34.
11. Jaureguibeitia X, U.I., Aramendi E, Wang H, Idris A. An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation. 2020 Computing Cardiology 2020;47:1–4.
12. Jaureguibeitia Xabier, Irusta Unai, Aramendi Elisabete, Owens Pamela C, Wang Henry E, Idris Ahamed H. Automatic detection of ventilations during mechanical cardiopulmonary resuscitation. IEEE J Biomed Health Inf 2020;24:2580–8.
13. Aramendi Elisabete, Elola Andoni, Alonso Erik, et al. Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation. Resuscitation 2017;110:162–8.
14. Aramendi Elisabete, Lu Yuanzheng, Chang Mary P, et al. A novel technique to assess the quality of ventilation during pre-hospital cardiopulmonary resuscitation. Resuscitation 2018;132:41–6.
15. Aufderheide Tom P, Frascone Ralph J, Wayne Marvin A, et al. Standard cardiopulmonary resuscitation versus active compression-decompression cardiopulmonary resuscitation with augmentation of negative intrathoracic pressure for out-of-hospital cardiac arrest: a randomised trial. Lancet 2011;377:301–11.
16. Yannopoulos Demetris, Nadkarni Vinay M, McKnite Scott H, et al. Intrathoracic pressure regulator during continuous-chest-compression advanced cardiac resuscitation improves vital organ perfusion pressures in a porcine model of cardiac arrest. Circulation 2005;112:803–11.
17. Paiva Edison F, Paxton James H, O'Neil Brian J. The use of end-tidal carbon dioxide (ETCO2) measurement to guide management of cardiac arrest: A systematic review. Resuscitation 2018;123:1–7.

18. Edelson Dana P, Eilevstjønn Joar, Weidman Elizabeth K, Retzer Elizabeth, Hoek Terry L Vanden, Abella Benjamin S. Capnography and chest-wall impedance algorithms for ventilation detection during cardiopulmonary resuscitation. Resuscitation 2010;81:317–22.
19. Stecher Frederik S, Olsen Jan-Aage, Stickney Ronald E, Wik Lars. Transthoracic impedance used to evaluate performance of cardiopulmonary resuscitation during out of hospital cardiac arrest. Resuscitation 2008;79:432–7.
20. Lupton Joshua R, Schmicker Robert, Daya Mohamud R, et al. Effect of initial airway strategy on time to epinephrine administration in patients with out-of-hospital cardiac arrest. Resuscitation 2019;139:314–20.
21. Wang Henry E, Jaureguibeitia Xabier, Aramendi Elisabete, et al. Airway strategy and chest compression quality in the pragmatic airway resuscitation trial. Resuscitation 2021;162:93–8.