Iraia Isasi, Unai Irusta, Elisabete Aramendi, Jan-Age Olsen, Lars Wik, Shock decision algorithm for use during load distributing band cardiopulmonary resuscitation, Resuscitation, Volume 165, 2021, Pages 93-100, ISSN 0300-9572,https://doi.org/10.1016/j.resuscitation.2021.05.028.

(https://www.sciencedirect.com/science/article/pii/S0300957221002124?via%3Dihub)

Aim: Chest compressions delivered by a load distributing band (LDB) induce artefacts in the electrocardiogram. These artefacts alter shock decisions in defibrillators. The aim of this study was to demonstrate the first reliable shock decision algorithm during LDB compressions.
Methods: The study dataset comprised 5813 electrocardiogram segments from 896 cardiac arrest patients during LDB compressions. Electrocardiogram segments were annotated by consensus as shockable (1154, 303 patients) or nonshockable (4659, 841 patients). Segments during asystole were used to characterize the LDB artefact and to compare its characteristics to those of manual artefacts from other datasets. LDB artefacts were removed using adaptive filters. A machine learning algorithm was designed for the shock decision after filtering, and its performance was compared to that of a commercial defibrillator's algorithm.
Results: Median (90% confidence interval) compression frequencies were lower and more stable for the LDB than for the manual artefact, 80 min−1 (79.9-82.9) vs 104.4 min−1 (48.5-114.0). The amplitude and waveform regularity (Pearson's correlation coefficient) were larger for the LDB artefact, with 5.5 mV (0.8-23.4) vs 0.5 mV (0.1-2.2) ($p<0.001$) and 0.99 (0.78-1.0) vs 0.88 (0.55-0.98) ($p<0.001$). The shock decision accuracy was significantly higher for the machine learning algorithm than for the defibrillator algorithm, with sensitivity/specificity pairs of 92.1/96.8% (machine learning) vs 91.4/87.1% (defibrillator) ($p<0.001$).
Conclusion: Compared to other cardiopulmonary resuscitation artefacts, removing the LDB artefact was challenging due to larger amplitudes and lower compression frequencies. The machine learning algorithm achieved clinically reliable shock decisions during LDB compressions.

Keywords: Out-of-hospital cardiac arrest (OHCA), Cardiopulmonary Resuscitation (CPR), Shock decision algorithm, Load Distributing Band (LDB), Machine Learning (ML), shock/no-shock

Shock decision algorithm for use during load distributing band cardiopulmonary resuscitation

I Isasi*[a], U Irusta[a,b], E Aramendi[a,b], JA Olsen[c], L Wik[c]

Affiliation and addresses:

[a] Communications Engineering Department.
University of the Basque Country UPV/EHU.
Plaza Ingeniero Torres Quevedo S/N
48013 Bilbao, Bizkaia, Spain

[b] Biocruces Bizkaia Health Research Institute.
Cruces Plaza, 48903 Barakaldo, Bizkaia, Spain

[c] National Advisory Unit for Prehospital Emergency Medicine (NAKOS) and Department of Anaesthesiology.
Oslo University Hospital and University of Oslo.
Po Box 4956 Nydalen, N-0424 Oslo, Norway

Corresponding author:

[*] Iraia Isasi
email: iraia.isasi@ehu.eus
Telephone: +34 946017386
Fax: +34 946014259

Word counts:

Abstract: 250 words

Paper : 3000 words

Abstract

Aim: Chest compressions delivered by a load distributing band (LDB) induce artefacts in the electrocardiogram. These artefacts alter shock decisions in defibrillators. The aim of this study was to demonstrate the first reliable shock decision algorithm during LDB compressions.

Methods: The study dataset comprised 5813 electrocardiogram segments from 896 cardiac arrest patients during LDB compressions. Electrocardiogram segments were annotated by consensus as shockable (1154, 303 patients) or nonshockable (4659, 841 patients). Segments during asystole were used to characterize the LDB artefact and to compare its characteristics to those of manual artefacts from other datasets. LDB artefacts were removed using adaptive filters. A machine learning algorithm was designed for the shock decision after filtering, and its performance was compared to that of a commercial defibrillator's algorithm.

Results: Median (90% confidence interval) compression frequencies were lower and more stable for the LDB than for the manual artefact, $80\,\mathrm{min}^{-1}$ (79.9-82.9) vs $104.4\,\mathrm{min}^{-1}$ (48.5-114.0). The amplitude and waveform regularity (Pearson's correlation coefficient) were larger for the LDB artefact, with 5.5 mV (0.8-23.4) vs 0.5 mV (0.1-2.2) ($p<0.001$) and 0.99 (0.78-1.0) vs 0.88 (0.55-0.98) ($p<0.001$). The shock decision accuracy was significantly higher for the machine learning algorithm than for the defibrillator algorithm, with sensitivity/specificity pairs of 92.1/96.8% (machine learning) vs 91.4/87.1% (defibrillator) ($p<0.001$).

Conclusion: Compared to other cardiopulmonary resuscitation artefacts, removing the LDB artefact was challenging due to larger amplitudes and lower compression frequencies. The machine learning algorithm achieved clinically reliable shock decisions during LDB compressions.

Keywords

Out-of-hospital cardiac arrest (OHCA), Cardiopulmonary Resuscitation (CPR), Shock decision algorithm, Load Distributing Band (LDB), Machine Learning (ML), shock/no-shock decision.

1. INTRODUCTION

The cardiopulmonary resuscitation (CPR) guidelines stress the importance of high-quality CPR and early defibrillation for a successful outcome after out-of-hospital cardiac arrest (OHCA). [1] Chest compressions during CPR introduce artefacts in the electrocardiogram (ECG) that impede a reliable rhythm analysis by the defibrillator. [2] Therefore, the current practice mandates CPR interruption to analyse the ECG and provide a reliable shock decision. These "hands-off" intervals for rhythm analysis lead to intermittent lack of cerebral and myocardial blood flow, and may compromise the outcome of the patient. [3–7] Methods allowing a reliable shock decision without interrupting CPR would therefore be of great value.

Shock decision algorithms during manual CPR rely on artifact suppression filters. [8] These filters use additional signals like the thorax impedance [9–13] or the compression depth [9–12] to adapt to the time-changing characteristics of manual chest compressions. However, shock decision algorithms in defibrillators were designed to analyse artefact-free ECGs, and thus show a degraded performance even after filtering the artefact. [12,13] To address this limitation, research has focused on machine learning (ML) algorithms that analyse the ECG after filtering the CPR artefact. [14–17] These algorithms learn the characteristics of the filtered ECG, inluding those of the filtering residuals, and have been shown to meet the American Heart Association's (AHA) sensitivity and specificity requirements for shock decision algorithms.

The use of mechanical compression devices is rising in the prehospital setting, although none of the more recent clinical trials have demonstrated improved survival. [18–20] The benefits of mechanical CPR include guaranteed compression rate and depth, and the possibility for high-quality chest compressions even during transport and invasive procedures. [21–23] At present there are two main mechanical chest compression technologies: piston-driven and load distributing band (LDB) devices. ECG analysis algorithms for a reliable shock decision during mechanical CPR have only been demonstrated for piston-driven devices. [24–27] These studies have shown that a fixed compression rate and depth results in a much more periodic artefact than that of manual CPR. [25] Therefore, simpler adaptive filters that exploit the periodic nature of the artefact have been successfuly used to remove mechanical artefacts without the need of additional reference signals. These adaptive filters in combination with ML-algorithms for the shock decision produced solutions compliant with the AHA requirements. [26]

This study demonstrates the first method for a reliable shock decision during LDB compressions using a combination of a CPR artifact filter and a ML-algorithm. The analysis also includes a characterization of the LDB artefact in the time and frequency domain, a preliminary step needed to properly design the CPR artefact suppression filters.

2. MATERIALS AND METHODS

2.1. Data collection

The data used in this study were extracted from the randomized controlled Circulation Improving Resuscitation Care (CIRC) [18] trial conducted between March 2009 and January 2011 by emergency services in the United States (three services) and Europe (two services). The CIRC trial was designed to compare the effectiveness of an LDB device (AutoPulse, ZOLL, Chelmsford, Massachusetts, USA) against manual CPR in terms of survival to hospital discharge. [18] Data from 969 OHCA patients from one US (Hillsborough) and two European (Nijmegen and Vienna) services were used to conduct this study. The LDB device provides chest compressions in a fixed position, a constant rate of $80\,\mathrm{min}^{-1}$ ($f_{\mathrm{LDB}} = 1.33\,\mathrm{Hz}$) and a depth of 20% of the patient's anterior posterior diameter of the chest. Anonymized waveform data from the Lifepak 12 and 15 monitor-defibrillators (Physio-Control, Redmond, WA, USA) was exported to Matlab (MathWorks Inc., Natick, MA) using Physio-Control's Code Stat data review software, and resampled for processing with a sampling period of $T_s = 4\,\mathrm{ms}$. The data included the ECG and thorax impedance signals together with the compression instants detected by the Code Stat software. The ECG was band limited to 0.5–40 Hz (4th order Butterworth filter), a typical ECG monitor bandwidth used in defibrillators.

The use of the LDB device was identified when the compression rate stabilized at the device's fixed rate of $80\,\mathrm{min}^{-1}$ for at least 20 s (see Fig. 1). Then, 22 s signal segments were automatically extracted following these criteria: unique rhythm type in the entire segment, and an interval of 16 s with LDB compressions followed or preceded by a 6 s interval without compressions (see highlighted segment in Fig. 1). The intervals during compressions were used to develop the shock decision algorithms, whereas the artefact-free intervals were used to annotate the underlying rhythm. Rhythms were annotated as shockable (coarse ventricular fibrillation (VF), i.e. with amplitude larger than $200\,\mu\mathrm{V}$, pulseless ventricular tachycardia (VT)) or nonshockable (asystole, organized rhythm) by consensus of an experienced anesthesiologist (LW) and two biomedical engineers (UI and EA), all specialized in cardiac arrest data management.

2.2. LDB artefact characterization

Segments during asystole were used to characterize the artefact. First, spectral characteristics were analysed using the normalized power spectrum density, computed as the average of the power

spectrum of each segment normalized to unit area (Fig. 2). Multiple segments from the same patient were weighted to avoid patient-bias. Second, parameters like the peak-to-peak amplitude and waveform regularity were estimated to characterize the artefact in the time domain. Waveform regularity was quantified by Pearson's correlation coefficient (PCC) between each compression cycle and the segment's compression cycle template (the average of all cycles). [25] PCC values close to ± 1 correspond to periodic waveforms, whereas values close to 0 indicate irregularity. Finally, the signal to noise ratio was calculated as the power ratio between the clean ECG and the artefact for shockable and organized rhythms in a logarithmic scale (in decibel, dB). [9,28] The power of the artefact was calculated as the power difference between the artefacted ECG (16 s interval) and the clean ECG (6 s interval). All characteristics were compared between the LDB artefact, the piston-driven artefact and the manual artefact. The datasets used in Aramendi et al [25] (piston-driven CPR), and Isasi et al [16] (manual CPR) were used for this purpose.

2.3. LDB artefact suppression method

The LDB artefact presents an almost periodic waveform due to the fixed rate and depth of mechanical compressions. Consequently, the compression rate (f_{LDB}) can be used to model the artefact accurately. [26,27] The spectral characterization of the artefact (Fig. 2) shows that the power content is concentrated around the harmonics of f_{LDB}. Therefore, we propose a model of the artefact composed of N harmonically related sines and cosines with amplitudes $a_k(n)$ and $b_k(n)$: [12,26]

$$s_{\mathrm{cpr}}(n) = \sum_{k=1}^{N} a_k(n)\cos(2\pi k f_{\mathrm{LDB}} n T_s) + b_k(n)\sin(2\pi k f_{\mathrm{LDB}} n T_s) \tag{1}$$

where n is the discrete time index ($t = nT_s$) and $f_{\mathrm{LDB}} = 1.33\,\mathrm{Hz}$ ($80\,\mathrm{min}^{-1}$). The artefact was estimated by calculating the amplitudes $a_k(n)$ and $b_k(n)$ using the recursive least squares (RLS) algorithm. [26,29] The CPR artefact was then subtracted from the corrupt ECG to obtain the filtered ECG. Filter design consists in the selection of the number of harmonics, N, and the adaptability of the filter, λ. Consult [27,30] for further details on the adaptive RLS filter.

2.4. Shock decision algorithm

The ML-algorithm for shock decision consisted of three stages. First, ECG features were computed using 8 s segments of the filtered ECG ($4\,\mathrm{s} - 12\,\mathrm{s}$ interval). Then the most discriminative

subset of features was selected using state-of-the-art feature selection techniques. Finally, features were combined in a Support Vector Machine (SVM) classifier for the shock/no-shock decision.

In total 38 ECG features were computed using methods recently developed for ECG rhythm analysis during piston-driven chest compressions. [26] These features were based on the Stationary Wavelet Transform decomposition and include time domain, frequency domain and signal complexity characterizations of the ECG.

In the feature selection procedure four features were sequentially added to the SVM classifier and then three were sequentially removed at each interation. [26,31] The criterion to add/remove features and to set the soft margin (C) and the width of the gaussian kernel (γ) in the SVM was to maximize the balanced accuracy. The balanced accuracy is the arithmetic average of sensitivity and specificity and thus gives equal importance to shock and no-shock decisions. Similar ML approaches and ML parameter optimization procedures have been used in the design of shock decision algorithms during piston-driven CPR. [26]

2.5. Diagnostic evaluation

Data were randomly partitioned patient-wise into training (70%) and test (30%) sets ensuring that the prevalences of shockable and nonshockable rhythms in both sets resembled those of the whole dataset. The training set was split into 5-fold cross validation (CV) partitions for feature selection and SVM parameter optimization. The feature selection algorithm was run until the optimal 15 feature set was obtained. These features were then used to determine the values of the SVM parameters that maximized the balanced accuracy in the ranges: $10^{-1} \leq C \leq 10^{2}$ and $10^{-3} \geq \gamma \geq 10^{1}$. The optimized ML-algorithm was then used to obtain the shock/no-shock decisions in the test set, which were compared with the ground truth labels to obtain the sensitivity, specificity, balanced accuracy and total accuracy. The process was repeated 100 times with random 70/30 train/test partitions to obtain statistically meaningful results. Two additional experiments were conducted to obtain baseline accuracy results. First, the ML-algorithm was optimized and evaluated before and after filtering, to quantify the benefits of filtering the LDB artefact. Second, the accuracy of a commercial defibrillator (Reanibex R series, BexenCardio, Ermua, Spain) [12] was obtained before and after filtering, to evaluate the increase in accuracy of the ML-algorithm over the classical diagnostic approach.

3. RESULTS

Shockable rhythms comprised lethal ventricular arrhythmia, predominantly ventricular fibrillation VF (19.2%), but also VT (0.7%). Nonshockable rhythms included asystole (27.8%) and organized rhythms (%52.3). In total 5813 segments (from 896 patients) were extracted, whereof 1154 (from 303 patients) were shockable and 4659 (from 841 patients) nonshockable.

3.1. Characteristics of the artefact

Fig. 2 shows the spectral analysis of the LDB, the piston-driven and the manual CPR artefacts. Mechanical CPR artefacts presented very stable compression frequencies, with 90% of the power concentrated around the first 7 harmonics (see errorbars) of the compression frequency. Manual CPR artefacts showed great variability in compression frequency that resulted in a spread spectrum with 90% of the power concentrated around the first 3 harmonics. So, mechanical CPR artefacts have more harmonics and higher bandwidths than manual CPR artefacts. The median compression frequencies (90% confidence interval, CI) were $80.1\,\mathrm{min}^{-1}$ (79.7 – 82.9), $101.7\,\mathrm{min}^{-1}$ (101.1 – 102.1) and $104.4\,\mathrm{min}^{-1}$ (48.5 – 114.0) for the LDB, the piston-driven, and the manual compressions respectively. The median artefact amplitudes were 5.5 mV (0.8 – 23.4), 1.2 mV (0.2 – 4.8) and 0.5 mV (0.1 – 2.2) for the LDB, piston-driven and manual artefacts, respectively. The amplitude was significantly larger ($p<0.001$ for the Mann-Whitney U test) for the LDB artefact. This was consistent with the median signal to noise ratios observed, which were −7.4 dB (−21.1 – 5.3), −2.6 dB (−12.3 – 6.5) and −0.2 dB (−13.6 – 11.7), respectively. The median PCC values obtained for LDB, piston-driven and manual compressions were 0.991 (0.776 – 0.999), 0.979 (0.729 – 0.998) and 0.883 (0.550 – 0.979), respectively. PCC values showed significantly larger regularity for LDB devices ($p<0.001$), when compared to other modes of CPR, but also larger for any mode of mechanical CPR when compared to manual CPR ($p<0.001$).

3.2. Diagnostic accuracy of the shock decision

Fig. 3 shows the sensitivity, specificity and balanced accuracy values (averaged over the 100 test partitions) for the shock decision during LDB use as a function of the filter's adjustable parameters, for both the ML-algorithm and the algorithm of a defibrillator. The defibrillator algorithm showed near-optimal performance with a balanced accuracy exceeding 88.0% for $N > 30$ and $\lambda \sim 0.985$–0.990. The ML-algorithm achieved a balanced accuracy above 93.5% for $N > 30$

and a wider range of $\lambda \sim 0.975 - 0.990$. Table I compares the performance of the defibrillator algorithm and the ML-algorithm before and after filtering. The optimal working point of the RLS filter, $N = 35$ and $\lambda = 0.989$, was used to remove the artefacts from the ECG. Filtering increased the balanced accuracy of the commercial and ML-algorithm by over 26 and 8 points, respectively. The increased balanced accuracy was due mainly to an increase of over 56 (commercial algorithm) and 13 (ML-algorithm) points in sensitivity. After filtering the sensitivity of the defibrillator's algorithm was above 90%, the minimum value recommended by the AHA, but the specificity was almost 8 points below the 95% AHA recommendation. The ML-algorithm increased the balanced accuracy by over 5-points, but most importantly the specificity by over 8-points, obtaining sensitivity/specificity pairs that met the values recommended by the AHA.

The left panels on Fig. 5 show examples in which the defibrillator algorithm and the ML-algorithm correctly classify the ECG after filtering (panels a, b and c). In the examples on the right, filtering residuals confound the defibrillator's algorithm which was designed to analyse clean ECGs. These examples are correctly classified by the ML-algorithm since it was trained with filtered ECGs, and thus learnt the characteristics of filtering residuals. Two possible causes of error are shown in Fig. 5: disorganized filtering residuals that resemble VF during asystole and organized rhythms (panels e and d, respectively), and spiky artefacts for each compression during VF that can be confounded as QRS complexes (panel f).

4. Discussion

The use of mechanical CPR devices has grown considerably in the last years through two main technologies, LDB and piston-driven devices. Mechanical CPR guarantees high-quality chest compressions when manual compressions cannot be delivered safely. Examples of such situations include ambulance transportation, [32,33] primary percutaneous coronary intervention, [34,35] as a bridge to extracorporeal CPR, [36] and to facilitate uncontrolled organ donation after circulatory death. [37]

This study describes the first method to give an AHA compliant shock/no-shock decision during chest compressions provided by a LDB device. The solution consists of an adaptive filter that removes the CPR artefact, followed by a ML-algorithm that obtains an accurate diagnosis based on multiple ECG features. Similar approaches have been used to design shock decision algorithms during manual [15] and piston-driven chest compressions. [26] The sensitivity/specificity values obtained in previous studies for manual (95.2/95.6%) and piston-driven (97.5/98.2%) CPR, and those obtained in this study for LDB devices (92.1/96.8%) provide clinically reliable solutions that cover rhythm analysis during chest compressions in any situation.

Chest compression artefacts in the ECG act as major confounders for shock decision algorithms. The sensitivity of a commercial shock decision algorithm designed to work on clean ECGs dropped from a nominal 99% [38] to 35% in the presence of LDB artefacts. This is mainly because artefacted ECGs show an organized activity locked to the compression frequency ($80\,\mathrm{min}^{-1}$) which resembles an organized rhythm even during VF. The ML-algorithm substantially increased the sensitivity of the commercial algorithm due to its ability to select features that are minimally affected by the artefact. Still, the sensitivity of the ML-algorithm without filtering was 79%, far from the 90% recommended by the AHA.

The addition of the adaptive filtering stage substantially increased the performance of both the ML and the commercial algorithm. AHA requirements were only met by the use of the ML-algorithm, which showed an improvement of 9.8-points in specificity and 0.7-points in sensitivity over the commercial algorithm for the filtered ECG. [12] The defibrillator algorithm was very affected by the configuration of the artefact suppresion filter (see variability in Fig. 3), while the ML-algorithm learnt to classify ECGs in the presence of filtering residuals.

The performance of the ML-algorithm in this study is similar to that of a recently demonstrated

ML-algorithm during manual CPR. [15] The characteristics of manual compressions are rescuer dependant, so the variability of the resulting artefact anticipates a more challenging filtering problem. However, manual artefacts showed significantly smaller artefact amplitudes and less harmonic components (smaller bandwidths) than LDB artefacts and both effects compensated yielding similar accuracies. That was not the case for piston-driven chest compressions, as recently shown in Isasi et al. [26], for which the solution showed 5/2 points of sensitivity/specificity above the LDB solution. Three factors could explain the lower performance during LDB compressions. First, the LDB compresses a larger portion of the chest, producing artefact with amplitudes 5 times larger than for the piston driven device (5.5 mV vs 1.2 mV, $p<0.001$). Second, the compression rate of the LDB device is lower (80 min^{-1} vs 100 min^{-1}), and consequently more harmonics of the chest compression frequency fall in the frequencies of interest of the ECG (see Fig. 2). Third, the LDB compression pattern is trapezoidal, compared to the more sinusoidal piston-driven and manual artefact patterns, which causes more spectral broadening.

The SVM algorithm could be replaced by other machine learning algorithms. A Logistic Regression and a Random Forest algorithm were also evaluated in this study following a similar procedure. The differences in balanced accuracy between the three ML-algorithms were not statistically significant ($p>0.3$ for the McNemar test, in all comparisons) although the SVM achieved the best overall accuracy and specificity values. In the future deep neural network algorithms could be investigated. Deep learning algorithms outperform classical ML-algorithms when large volumes of annotated data are available. [16] The accuracy for shock advice during LDB use should increase if deep learning solutions are trained with sufficiently large datasets.

4.1. Limitations

The main limitations of this study are related to the data. Although the dataset came from two different monitor-defibrillators, both models are from the same vendor. So, the methods may need adjusting if ECG data from other vendors is used. Moreover, data was gathered by three emergency services and there may be differences in resuscitation protocols and LDB device usage across services. These differences in treatment may alter the characteristics of the CPR artefacts and thus the algorithms may need fine-tuning. However, the mechanical CPR is more uniform than manual CPR, so we do not anticipate major changes in our algorithms to accomodate data from different services.

5. Conclusions

The first method for an automatic and clinically safe shock decision during LDB CPR was demonstrated, meeting AHA requirements. Filtering the LDB artefact was challenging because of the larger amplitudes and a ML-algorithm was required to provide a reliable shock decision on the filtered ECG.

The proposed solution together with solutions already available for piston-driven and manual chest compressions would cover rhyhm analysis in every CPR scenario. This may open the possibility of bringing uninterrupted mechanical and manual CPR into real practice, eliminating the no-flow intervals for rhythm analysis, and thus contributing to improve OHCA outcomes.

Conflict of interest

Dr. Wik is member of the medical advisor board at Stryker and has patents licensed to Zoll and Stryker.

Acknowledgements

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 grant IT1229-19, and by the university of the Basque Country (UPV/EHU) under grant COLAB20/01.

References

[1] Perkins G, Handley A, Koster R, 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] Fitzgibbon E, Berger R, Tsitlik J, Halperin HR. Determination of the noise source in the electrocardiogram during cardiopulmonary resuscitation. Crit Care Med 2002;30(4 Suppl):S148–S153.

[3] 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(12):1501–1507.

[4] Kern KB, Hilwig RW, Berg RA, Sanders AB, Ewy GA. Importance of continuous chest compressions during cardiopulmonary resuscitation: improved outcome during a simulated single lay-rescuer scenario. Circulation 2002;105(5):645–649.

[5] Eftestøl Trygve, Sunde Kjetil, Steen Petter Andreas. Effects of Interrupting Precordial Compressions on the Calculated Probability of Defibrillation Success During Out-of-Hospital Cardiac Arrest. Circulation 2002; 105(19):2270–2273.

[6] Sato Y, Weil MH, Sun S, et al. Adverse effects of interrupting precordial compression during cardiopulmonary resuscitation. Critical Care Medicine 1997;25(5):733–736.

[7] Yu T, Weil MH, Tang W, et al. Adverse outcomes of interrupted precordial compression during automated defibrillation. Circulation 2002;106(3):368–372.

[8] Gong Y, Chen B, Li Y. A review of the performance of artifact filtering algorithms for cardiopulmonary resuscitation. Journal of Healthcare Engineering 2013;4(2):185–202.

[9] Langhelle A, Eftestøl T, Myklebust H, Eriksen M, Holten BT, Steen PA. Reducing CPR artefacts in ventricular fibrillation in vitro. Resuscitation 2001;48(3):279–291.

[10] Aase SO, Eftestøl T, Husøy JH, Sunde K, Steen PA. CPR artifact removal from human ECG using optimal multichannel filtering. IEEE Trans Biomed Eng 2000;47(11):1440–1449.

[11] Eilevstjønn J, Eftestøl T, Aase SO, Myklebust H, Husøy JH, Steen PA. Feasibility of shock advice analysis during CPR through removal of CPR artefacts from the human ECG. Resuscitation 2004;61(2):131–141.

[12] Irusta U, Ruiz J, Ruiz de Gauna S, Eftestøl T, Kramer-Johansen J. A Least Mean-Square Filter for the Estimation of the Cardiopulmonary Resuscitation Artifact Based on the Frequency of the Compressions. IEEE Trans Biomed Eng 2009;56(4):1052–1062.

[13] 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(6):692–698.

[14] Ayala U, Irusta U, Ruiz J, et al. A Reliable Method for Rhythm Analysis during Cardiopulmonary Resuscitation. BioMed Res Int 2014;2014:Article ID 872470.

[15] Isasi I, Irusta U, Rad AB, et al. Automatic Cardiac Rhythm Classification With Concurrent Manual Chest Compressions. IEEE Access 2019;7:115147–115159.

[16] Isasi I, Irusta U, Aramendi E, Eftestøl T, Kramer-Johansen J, Wik L. Rhyrhm Analysis during Cardiopulmonary

Resuscitation Using Convolutional Neural Networks. Entropy 2020;2:595–611.

[17] Isasi I, Irusta U, Aramendi E, Idris A, Sörnmo L. Restoration of the electrocardiogram during mechanical cardiopulmonary resuscitation. Physiological Measurement 2020;41:105006.

[18] 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(6):741–748.

[19] Rubertsson S, Karlsten R. Increased cortical cerebral blood flow with LUCAS; a new device for mechanical chest compressions compared to standard external compressions during experimental cardiopulmonary resuscitation. Resuscitation 2005;65(3):357–363.

[20] Perkins GD, Quinn T, Deakin CD, Lall R, Gates S. Mechanical chest compression in the PARAMEDIC trial - Authors' reply. Lancet 2015;386(9988):26–27.

[21] Olasveengen TM, Wik L, Steen PA. Quality of cardiopulmonary resuscitation before and during transport in out-of-hospital cardiac arrest. Resuscitation 2008;76(2):185–190.

[22] Grogaard HK, Wik L, Eriksen M, Brekke M, Sunde K. Continuous mechanical chest compressions during cardiac arrest to facilitate restoration of coronary circulation with percutaneous coronary intervention. J Am Coll Cardiol 2007;50(11):1093–1094.

[23] Stub D, Bernard S, Pellegrino V, et al. Refractory cardiac arrest treated with mechanical CPR, hypothermia, ECMO and early reperfusion (the CHEER trial). Resuscitation 2015;86:88–94.

[24] Sullivan J, Walker R, Esibov A, Chapman F. A digital filter can effectively remove mechanical chest compression artifact. Resuscitation 2014;85:S41.

[25] 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;98:41–47.

[26] Isasi I, Irusta U, Elola A, et al. A machine learning shock decision algorithm for use during piston-driven chest compressions. IEEE Transactions on Biomedical Engineering 2018;66(6):1752–1760.

[27] Isasi I, Irusta U, Aramendi E, et al. A multistage algorithm for ECG rhythm analysis during piston-driven mechanical chest compressions. IEEE Transactions on Biomedical Engineering 2018;66(1):263–272.

[28] Aase SO, Eftestøl T, Husøy J, Sunde K, Steen PA. CPR artifact removal from human ECG using optimal multichannel filtering. IEEE Trans Biomed Eng 2000;47(11):1440–1449.

[29] Alonso E, Aramendi E, Daya M, et al. Circulation detection using the electrocardiogram and the thoracic impedance acquired by defibrillation pads. Resuscitation 2016;99:56–62.

[30] Xiao Y, Ma L, Ward RK. Fast RLS Fourier analyzers capable of accommodating frequency mismatch. Signal Processing 2007;87(9):2197–2212.

[31] Rad AB, Eftestøl T, Engan K, et al. ECG-based classification of resuscitation cardiac rhythms for retrospective data analysis. IEEE Transactions on Biomedical Engineering 2017;64(10):2411–2418.

[32] Gässler H, Ventzke MM, Lampl L, Helm M. Transport with ongoing resuscitation: a comparison between manual and mechanical compression. Emergency Medicine Journal 2013;30(7):589–592.

[33] Gässler H, Kümmerle S, Ventzke MM, Lampl L, Helm M. Mechanical chest compression: an alternative in helicopter emergency medical services? Internal and emergency medicine 2015;10(6):715–720.

[34] Wagner H, Terkelsen CJ, Friberg H, et al. Cardiac arrest in the catheterisation laboratory: a 5-year experience

of using mechanical chest compressions to facilitate PCI during prolonged resuscitation efforts. Resuscitation 2010;81(4):383–387.

[35] Zhang ZP, Su X, Liu CW, Song D, Peng J, Yan H. Continuous mechanical chest compression-assisted percutaneous coronary intervention in a patient with cardiac arrest complicating acute myocardial infarction. Chinese Medical Journal 2015;128(6):846.

[36] Menegazzi JJ, Salcido DD, Housler GJ, Logue ES. Feasibility of initiating extracorporeal life support during mechanical chest compression CPR: a porcine pilot study. Resuscitation 2012;83(1):130–133.

[37] Morozumi J, Sakurai E, Matsuno N, et al. Successful kidney transplantation from donation after cardiac death using a load-distributing-band chest compression device during long warm ischemic time. Resuscitation 2009; 80(2):278–280.

[38] Irusta U. New signal processing algorithms for automated external defibrillators. Ph.D. thesis, University of the Basque Country 2010.

Figure Legends

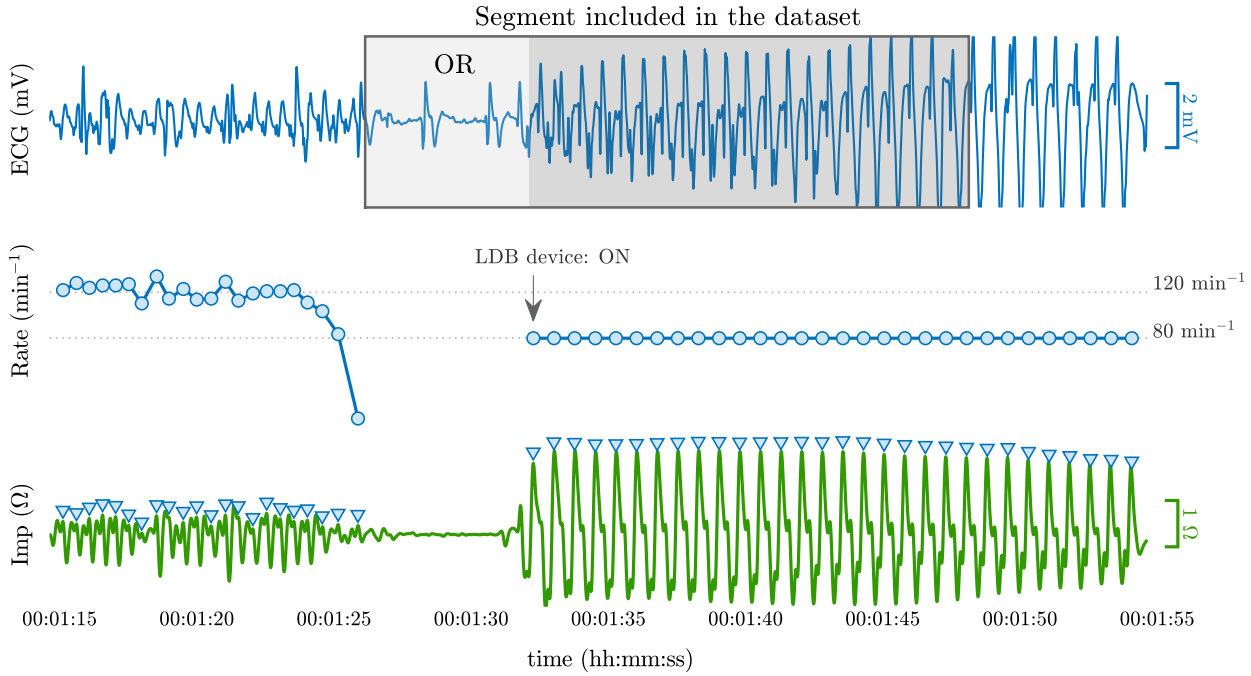

Figure 1: A 40 s interval from an OHCA episode showing the ECG and the thorax impedance (Imp) signals. Activity shows manual CPR followed by a pause for the LDB device application and immediate resumption of mechanical CPR. The compression rate for manual CPR was variable and fluctuated around 120 min^{-1}, but when the LDB device was applied the chest compression frequency stabilized around 80 min^{-1}. Note the clear change in the impedance pattern during mechanical compressions, with much larger amplitude, quicker compression/decompression cycles and a duty cycle around 50%. The interval highlighted in grey corresponds to a 22 s segment included in the dataset. The first 6 s were used to annotate the ground truth rhythm (organized) of the patient in a compression-free interval, and the last 16 s with artefact were used to develop the algorithms.

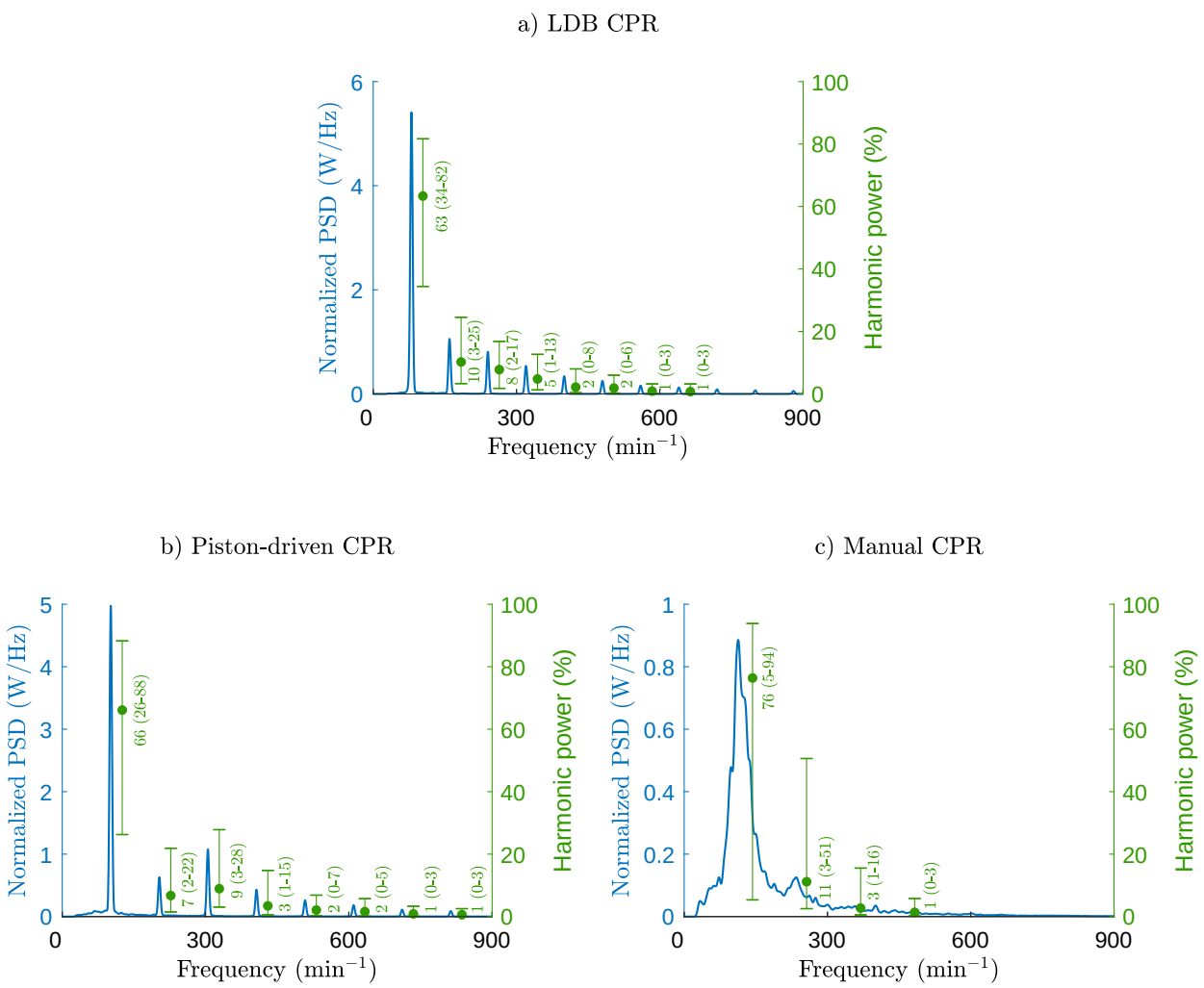

Figure 2: Normalized power spectral density (PSD) recorded during asysole for (a) the LDB device, (b) the piston-driven device, and (c) manual CPR. The errorbars represent median (80% confidence interval) of the power concentrated around each harmonic of the compression frequency.

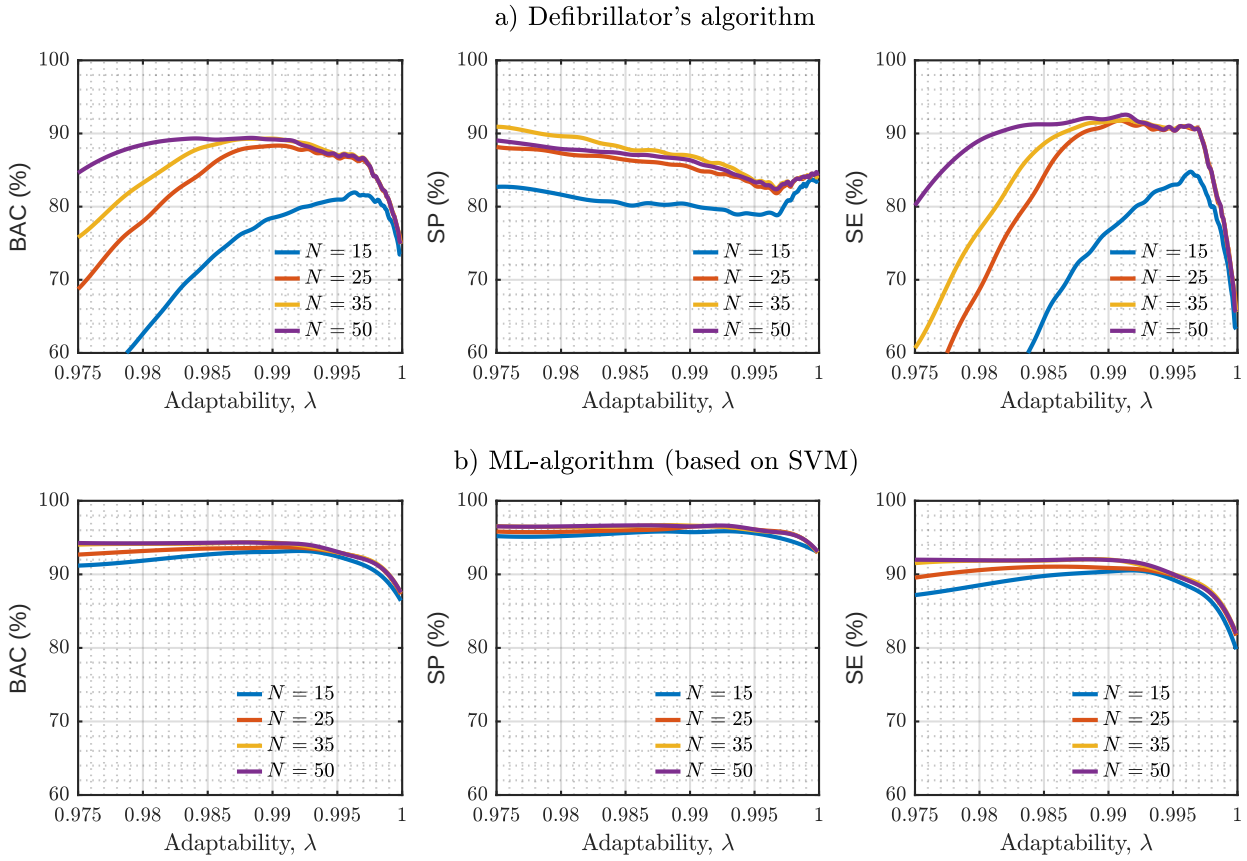

Figure 3: The mean values of balanced accuracy (BAC), specificity (Sp) and sensitivity (Se) obtained in the 100 test repetitions of the 70/30 partitions for the diagnosis of the filtered ECG using (a) a defibrillator's algorithm and (b) a ML-algorithm. The performance is given as a function of the adaptability (λ) and the number of harmonics (N) of the RLS filter.

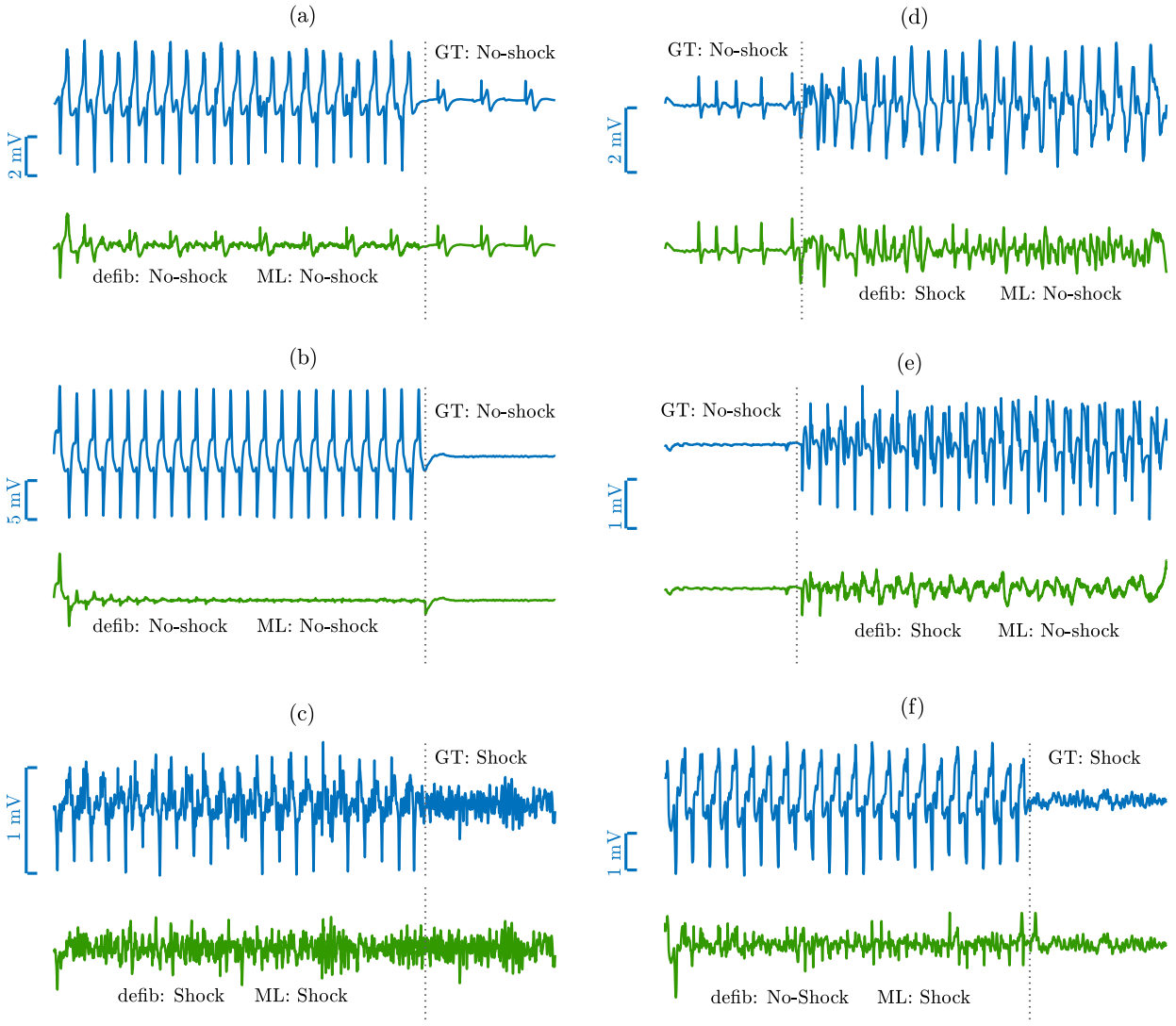

Figure 4: Examples of segments correctly classified by the defibrillator's algorithm and the ML-algorithm (left), and incorrectly classified by the defibrillator's algorithm but correctly classified by the ML-algorithm (right). The ground truth (GT) classification is shown on the top right corner in the interval without compressions, and the diagnoses of the defibrillator (defib) and the ML-algorithm. The wrong diagnoses of the defibrillator were caused by the inability of the filter to properly remove artefacts. This leads to very disorganized filtering residuals that resemble a VF during asystole or organized rhythms (d,e), or to spiky filtering residuals that look like QRS complexes during VF (f).

Table Legends

1 Sensitivity (Se), specificity (Sp), balanced accuracy (BAC) and total accuracy (Acc) for the commercial and the ML-algorithm before and after filtering. The performance of the ML-algorithm is given as median (95% CI) for the 100 test data repetitions of the 70/30 partitions. The commercial shock decision algorithm is a previously designed and trained algorithm, so its performance corresponds to the entire dataset.

Table 1: Sensitivity (Se), specificity (Sp), balanced accuracy (BAC) and total accuracy (Acc) for the commercial and the ML-algorithm before and after filtering. The performance of the ML-algorithm is given as median (95% CI) for the 100 test data repetitions of the 70/30 partitions. The commercial shock decision algorithm is a previously designed and trained algorithm, so its performance corresponds to the entire dataset.

Method	Se (%)	Sp (%)	BAC (%)	Acc (%)
Defibrillator				
Before filt.	35.4	90.2	62.8	79.3
After filt.	91.4	87.1	89.3	88.0
ML-algorithm				
Before filt.	78.9 (72.4 – 86.6)	93.6 (91.3 – 95.4)	86.3 (82.5 – 89.3)	90.7 (88.7 – 92.5)
After filt.	92.1 (91.0 – 93.0)	96.8 (96.1 – 97.2)	94.5 (93.8 – 95.0)	96.0 (95.4 – 96.2)