A Probabilistic Function to Model the Relationship between Quality of Chest Compressions and the Physiological Response for Patients in Cardiac Arrest

Trygve Eftestøl1, Svein E Stokka1, Jan T Kvaløy1, Ali Bahrami Rad2, Unai Irusta3, Elisabete Aramendi3, Erik Alonso3, Trond Nordseth4, Eirik Skogvoll4, Lars Wik5, Jo Kramer-Johansen5

1 University of Stavanger, Stavanger, Norway
2 Emory University, Atlanta, Georgia, USA
3 University of the Basque Country, Bilbao, Spain
4 Norwegian University of Science and Technology, Trondheim, Norway
5 Oslo University Hospital, Oslo, Norway

Abstract

Cardiopulmonary resuscitation quality (CPRQ) parameters can be derived from electric signals obtained during resuscitation. We propose to model a probabilistic relationship between CPRQ parameters and the physiological response as judged by ECG-features, to guide therapy in a clinical context.

A total of 821 compression sequences were extracted from 394 out-of-hospital resuscitation episodes. Sequences were categorized as effective if the post sequence cardiac rhythm had better prognosis than the pre-sequence rhythm by a positive difference, otherwise as non effective if the difference was negative. CPRQ parameters related to depth and rate were calculated. Three alternative approaches were designed for the binary classifier based on the CPRQ parameters: quadratic discriminant analysis (QDA), logistic regression (LR) and artificial neural networks (ANN). The positive class discriminant function defined the probability of effective compressions (Pec).

The classification accuracies were around 0.6 for all three models. The highest probability estimates of effective chest compressions corresponded to the depth (5-6 cm) and rate (100-120 min⁻¹) currently recommended in the CPR guidelines.

We have proposed a novel method to relate the quality of chest compressions to the physiological response to CPR.

1. Introduction

To survive cardiac arrest, early and good cardiopulmonary resuscitation (CPR) is important [1]. By monitoring the quality of chest compressions it will be possible to give feedback to the rescuer. This feedback can be used to improve the quality of the compressions, thus increasing the likelihood of survival [2]. The purpose of this study is to develop a probabilistic model of the relationship between quality parameters for CPR and physiological response.

2. Materials and methods

The data set consists of 394 episodes of out-of-hospital cardiac arrest. Data were collected in connection with a study that examined the quality of cardiopulmonary resuscitation performed by ambulance personnel [2, 3]. The study was conducted between March 2002 and September 2004 in 3 cities: Akershus (Norway), Stockholm (Sweden) and London (England).

Our group has conducted previous work on the data set, where methods for automatic classification of heart rhythm [4] and detection of chest compressions [5] have been developed. A method has also been developed to estimate the probability of change to a cardiac rhythm with a better prognosis, Pdes[6]. In the current study, Pdes, will be used to measure the patient’s physiological response to therapy.

Figure 1 shows an episode where Pdes is given in the y-axis, against time in minutes on the x-axis. The cardiac rhythms are color-coded. Red color corresponds to the cardiac arrest rhythms ventricular fibrillation or tachycardia (VF/VT), yellow indicates pulseless electrical activity (PEA), gray is asystole (AS) and green indicates pulse generating rhythm (PR). Intervals with chest compressions are indicated by a black horizontal field at the bottom of each plot. During periods of chest compressions an adaptive filter is applied to reduce the effects of the chest compression artifact on the rhythm annotator [7]. Black vertical bars indicate electric shocks.

Criteria related to changes in Pdes are used to extract data points that indicate the quality of chest compressions. Figure 1 visualizes how data points from a compression
interval are retrieved and sorted into the groups favorable/unfavorable treatment, depending on the change in the patient’s condition as indicated by \( P_{des} \). The blue box in Figure 1 shows a sequence with heart rhythm VF/VT, with chest compressions. To describe a change in the patient’s condition, \( \Delta P_{des} \) is defined as \( \Delta P_{des} = P_{des2} - P_{des1} \), where \( P_{des1} \) and \( P_{des2} \) are average values for the two compression free sequences preceeding and succeeding the compression sequence in question. A value of \( \Delta P_{des} \) larger than 0.1 is considered as an increase in probability for change to better condition. If this is the case during a compression interval, it is marked as an effective treatment.

The categories \( \omega_{ec} \) and \( \omega_{nec} \) for effective and non-effective compressions are defined for cases of \( \Delta P_{des} \geq 0.1 \) and \( \Delta P_{des} \leq -0.1 \) respectively.

Each compression sequence is characterised through 3-second analysis windows, with 2 seconds overlap. For each analysis window all chest compressions are detected in the chest compression depth signal. From the detections compression rate (\( CR \)), compression depth (\( CD \)), and the respective standard deviations, (\( SDCR \) and \( SDCD \)), are calculated. Similarly, compression rate features are calculated from the thoracic impedance signal, \( CRTI \) and \( SDCRTI \).

The features are optionally subjected to log transformation, normalisation by subtraction of mean and division by standard deviation, and finally decorrelation through principal component analysis (PCA). There is also the option of reducing the training data set of class \( \omega_{ec} \) by limiting the data to be within \( \pm 2 \) cm and \( \pm 20 \) compressions/min to fit the model to the recommendations for compression depth and rate.

From these, or a subset of these, feature vectors, \( x \) are constructed as characterisations of the therapeutic performance in the analysis windows. Our aim is to indicate whether a given \( x \), corresponds to efficient compressions or not. A probabilistic function, \( P(\omega_{ec}|x) \) is defined to express the probability that the compressions within the analysis window characterised by \( x \) are efficient (belonging to category \( \omega_{ec} \)). We denote this function as

\[
P_{ec} = \hat{P}(\omega_{ec}|x).
\]

We estimate this function by defining a classification problem where the aim is to discriminate \( X_{ec} = \{x|P(\omega_{ec}|x) \geq P(\omega_{nec}|x)\} \) from \( X_{nec} = \{x|P(\omega_{nec}|x) > P(\omega_{ec}|x)\} \). By Bayes law gives

\[
P(\omega_i|x) = \frac{P(\omega_i)p(x|\omega_i)}{\sum_j P(\omega_j)p(x|\omega_j)}
\]

where \( \omega_i, \omega_j \in \{\omega_{ec},\omega_{nec}\} \).

Thus, different strategies can be used to estimate \( P_{ec} \):

- **QDA** : Quadratic discriminant analysis, where \( P_{ec} \) is estimated through parametric estimation of \( p(x|\omega_i) \) and \( \hat{P}(\omega_i) = \eta_i/(n_{ec} + n_{nec}) \).
- **LR** : Logistic regression where \( P_{ec} \) is estimated through the parameters in \( \theta \) related by

\[
\ln \left( \frac{P_{ec}}{1 - P_{ec}} \right) = \theta^t x.
\]

ANN : \( P_{ec} \) and \( \hat{P}(\omega_{nec}|x) \) can be estimated by training the weights of a multilayer artificial neural network with two outputs, one for each of the functions, and \( x \) will be the inputs.

For each of these methods a series of 16 experiments were conducted where the various combinations of data handling methods (1) fitting model to guidelines by reducing the training data, (2) log transformation, (3) normalisation, and (4) PCA transform were applied.

For the QDA method, the mean vector \( \mu_i \) and covariance matrix \( \Sigma_i \) were estimated for \( i \in \{\omega_{ec},\omega_{nec}\} \). For the LR method quadratic versions of \( CD, CR \) and \( CRTI \) were added to provide the necessary nonlinearity. For the ANN the following hyperparameter space was explored: learning rates (0.00008, 0.00004, 0.00001), nodes (8, 12, 16), batch sizes (64, 128) and number of hidden layers (1,2,3,4).

A total of 821 compression sequences were extracted from the 394 episodes. Table 1 shows the number of compression sequences and corresponding number of 3-second windows found for the two categories of efficient and non-efficient performance.

In addition, an extended data set of 44123 analysis windows was compiled from all available compression sequences, including \( |\Delta P_{des}| < 0.1 \). Figure 2 illustrates the extended data set, the QDA model and corresponding \( P_{ec} \) values. The distribution of the model specific probabilities for \( CD \) and \( CR \) were computed for the extended data set.
\[ \Delta P_{des} \geq 0.1 \leq -0.1 \]

<table>
<thead>
<tr>
<th>Sequences</th>
<th>Train</th>
<th>Test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>467</td>
<td>354</td>
<td>821</td>
</tr>
<tr>
<td>Train</td>
<td>338</td>
<td>254</td>
<td>592</td>
</tr>
<tr>
<td>Test</td>
<td>129</td>
<td>100</td>
<td>229</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3-sec windows</th>
<th>Train</th>
<th>Test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>8702</td>
<td>5051</td>
<td>13753</td>
<td></td>
</tr>
<tr>
<td>5856</td>
<td>3423</td>
<td>9279</td>
<td></td>
</tr>
<tr>
<td>2846</td>
<td>1628</td>
<td>4474</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Number of compression sequences and analysis windows.

**Table 2. Model discrimination results.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Sens</th>
<th>Spec</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>QDA</td>
<td>0.59</td>
<td>0.36</td>
<td>0.72</td>
<td>0.42</td>
<td>0.66</td>
</tr>
<tr>
<td>LR</td>
<td>0.63</td>
<td>0.50</td>
<td>0.70</td>
<td>0.49</td>
<td>0.71</td>
</tr>
<tr>
<td>ANN</td>
<td>0.63</td>
<td>0.53</td>
<td>0.69</td>
<td>0.50</td>
<td>0.72</td>
</tr>
</tbody>
</table>

3. **Results**

The highest performance for both QDA and LR were achieved by applying all data handling techniques. Therefore all data handling techniques were applied in the ANN experiments to reduce the number of experiments. For the ANN a learning rate of 0.0001, 15 nodes, a batch size of 128 and one hidden layer gave the best performance.

Table 2 shows the performance metrics for the highest performing classifier for each model. The methods are fairly comparable, with slightly lower sensitivity and PPV for QDA.

Figure 3 shows a visualisation which was partly shown in Figure 1, the difference being that the compression sequences have been color-coded according to the \( P_{ec} \) values. Figure 4 shows a scatter plot of the extended data color-coded according to the computed \( P_{ec} \) values and contour lines for each model.

4. **Discussion**

The models behave nicely in the sense that the maximum \( P_{ec} \) values are in accordance with the guidelines as can be seen from both the \( P_{ec} \) distribution data in Table 3. Interestingly, Figure 3 shows that the compression sequences are shorter with more interruptions prior to the three first shocks. The quality in terms of \( P_{ec} \) is generally low prior to these shocks with the exception of some sequences prior to shock no 2. In contrast to this, the compression sequences prior to shock 4 are longer and with a higher value of \( P_{ec} \). Only shock 4 is successful in reestablishing a spontaneous circulation (PR rhythm). This is just
Figure 4. Distributions of $P_{ec}$ for CD and CR for models QDA (a), LR (b) and ANN (c).

Studying the contour lines of the models, Figure 4, the ANN based model is strikingly different from the other two models. A reason might be the larger number of parameters to be learned as well as generating highly nonlinear decision boundaries. The other models seem more reasonable considering the guidelines, but the ANN model seems to describe the data distribution better. This needs further investigation.

The rationale for using the discriminatory power to select the parametric configuration for each of the models might be questioned. We do not desire high discriminatory power itself, but see it as an indication that the probability function will have some degree of soft transition from high to low probability.

It is important to consider the effect of limiting the training data set which has been a contributing factor to the nice fit to the guideline recommended values.

5. Conclusion

We have proposed a novel method to relate the quality of chest compressions to the physiological response to CPR. The highest probability estimates of effective chest compressions corresponded to values currently recommended in the CPR guidelines.

References


Address for correspondence:
Trygve Eftestøl
University of Stavanger, N-4036 Stavanger, Norway
tel/fax: +47-5183-2035/1750
trygve.etestol@uis.no