Characterization of the respiratory pattern in patients with chronic heart failure applying joint symbolic dynamics

Caracterización del patrón respiratorio de pacientes con insuficiencia cardíaca crónica mediante análisis de dinámica simbólica conjunta

Molina-Ramírez, T. F.; Symonds-Torres, M. L.; Méndez-Carranza, D. F.; López-Barreras, J. D.


ABSTRACT
This work proposes a method to characterize the respiratory pattern of patients with chronic heart failure (CHF) to determine non-periodic breathing (nPB), periodic breathing (PB) and Cheyne-Stokes respiration (CSR) through non-linear, symbolic analysis of biological signals. A total of 43 patients were examined for their cardiorespiratory profiles, their ECG and respiratory pattern signals were processed, analyzed and studied for parameters that could be of potential use in clinical decision making, specifically in patient classification. Patients in the study were characterized through their cardiorespiratory signals, applying joint symbolic dynamics (JSD) analysis to cardiac beat and respiratory interval durations. The most statistically significant parameters across all groups were identified through a Kruskal-Wallis two tailed test ($\alpha = 0.05$) and a linear discriminant analysis (LDA) classification method based on such parameters was developed. The best result achieved with this classification method uses 10 features to discriminate patients with a 97.67% Accuracy (Acc). The best features to discriminate among groups are related to cardiorespiratory interaction rather than just respiration patterns alone. Results further support the idea that abnormal breathing patterns derive from physiological abnormalities in chronic heart failure.

KEYWORDS: Cheyne-Stokes respiration, Respiratory pattern, Joint Symbolic Dynamics, LDA.
RESUMEN

El trabajo propone un método para identificar el patrón respiratorio de pacientes con insuficiencia cardiaca crónica (CHF) con la finalidad de determinar respiración no periódica (NPB), respiración periódica (PB) y respiración de Cheyne-Stokes (CSR) a través de análisis simbólico no lineal de señales biológicas. Se examinaron los perfiles cardiorrespiratorios de 43 pacientes; sus señales de ECG y patrón respiratorio fueron procesados, analizados y estudiados en busca de parámetros que pudieran ser de utilidad, específicamente en la clasificación de pacientes. Estos pacientes se caracterizaron por medio de sus señales cardiorrespiratorias, aplicando un análisis de dinámica simbólica de conjuntos al ritmo cardíaco y a la duración de los intervalos respiratorios. Los parámetros de mayor significancia estadística entre todos los grupos se identificaron a través de una prueba Kruskal-Wallis de dos colas ($\alpha = 0,05$) y mediante un método de clasificación por análisis discriminante lineal (LDA). El mejor resultado conseguido con este método utiliza 10 características para discriminar a los pacientes con una precisión de 97,67% (Acc). Las características para discriminar entre grupos estuvieron relacionados con la interacción cardiorrespiratoria más que con solo los patrones de respiración, respaldando así la idea de que los patrones de respiración anormales derivan de anomalías fisiológicas presentes en la insuficiencia cardiaca crónica.

PALABRAS CLAVE: Respiración de Cheyne-Stokes, Patrón respiratorio, dinámica conjunta de símbolos, LDA.

Correspondencia

DESTINATARIO: Molina-Ramírez, T. F.
INSTITUCIÓN: Departamento de Ciencias Básicas,
Instituto Tecnológico de Hermosillo
DIRECCIÓN: Ave. Tecnológico y Periférico Poniente S/N,
Col. Sahuaro, C.P. 83170, Hermosillo, Son., México
CORREO ELECTRÓNICO: tomas.molina.r@gmail.com

Fecha de recepción:
14 de octubre de 2016

Fecha de aceptación:
19 de diciembre de 2016
INTRODUCTION

Periodic breathing (PB) is a breathing abnormality associated with various oscillatory forms characterized by rises and falls in ventilation, and Cheyne-Stokes respiration (CSR) is a more severe form of a PB pattern in which apneas and hypopneas alternate with repetitive gradual increases and subsequent gradual decreases in ventilation. Cheyne-Stokes respiration is one of several types of unusual breathing with recurrent apneas (dysrhythmias). This abnormal pattern of breathing can be seen in patients with hyponatremia, traumatic brain injuries and brain tumors [1].

PB has a prevalence as high as 70% in congestive heart failure (CHF) patients [2], and is associated with increased mortality [3], especially in CSR patients. Clinical studies show that elderly patients often have an altered breathing pattern, with PB and CSR, coinciding simultaneously with the presence or absence of CHF.

CSR is present in up to 40% of patients with CHF and several studies have shown an increased mortality in patients with both conditions [4-6]. Based upon small case series, patients with congestive heart failure and Cheyne-Stokes respiration have a significantly greater mortality [6, 7], particularly if present during wakefulness [8] than those without Cheyne-Stokes respiration.

An effective respiratory pattern classification method as a clinical decision tool could provide patients undergoing treatment with a benefit in prognostic outcome and provide us with better tools to help patient recovery.

Efforts to characterize breathing patterns and classify patients using biological signals, mainly ECG and breathing rate, have already been developed using traditional time-domain analysis techniques. Garde et al [9] proposed two highly effective methods for distinguishing healthy patients from those with abnormal breathing patterns using both linear and non-linear classifiers with an accuracy level of 93% and 100% respectively. Another method to discern obstructive sleep apnea (OSA) from Cheyne-Stokes respiration (CSR) using frequency domain analysis from overnight electrocardiography (ECG) was described by Suhas, Vijendra, Burk, Lucas and Behbehani [10] yielding an 87.5% sensitivity and a specificity of 75%. Neural networks classifiers have also been applied to the detection of irregular breathing patterns [11].

While traditional methods might prove useful in patient classification, the characterizations of the cardiorespiratory interactions of the patients might help us better understand the mechanisms of respiratory regulation and the effect of cardiac regulation on abnormal breathing patterns. To better examine these interactions and their effect on the status of the patients a joint symbolic dynamics (JSD) analysis was applied to the cardiac (ECG) and respiratory (RESP) signals of the patients in this study.

The main idea behind the concept of JSD is the elimination of detailed information in order to keep the robust properties of the dynamics by a coarse graining of the measurement [12, 13].

After patients’ cardiorespiratory profile characterization, a classification model based on statistical analysis was designed to discriminate between study groups.

METHODOLOGY

The cardiorespiratory profiles (electrocardiographic and respiratory flow signal) of 45 elderly patients (age 71+) with congestive heart failure were recorded for 15 minutes under no stress at the Departments of Intensive Care Unit at the Santa Creu i Sant Pau Hospital, Barcelona, Spain, and the Getafe Hospital, Getafe, Spain. The patients of the study (age 71-93) gave their informed consent to participate in the study and the recording was done according to protocols established by local ethics committees. Using clinical criteria based on the patients’ respiratory profile and medical observation, patients included in this study were classified into one of three study groups:
Group 0, 20 patients (9 male, 11 female, aged 81.7 ± 8 years) that presented a non-periodic breathing pattern (nPB).

Group 1, 17 patients (9 male, 8 female, aged 81.5 ± 7 years) that presented a periodic breathing pattern (PB),

Group 2, 6 patients (2 male, 4 female, aged 81.3 ± 7 years) with a Cheyne-Stokes breathing pattern (CSR).

Patients with PB could or could not exhibit this behavior during the recordings.

Of the total 45 patients, two (patients ‘p0027’ and ‘p0042’) were removed from the study due to high noise in the signal recordings.

Both signals, ECG and respiratory flow signals (ECG and RESP respectively), were recorded synchronously with a sampling frequency of 250 Hz for 15 min. The patients’ ECG and respiratory flow signals were preprocessed to remove signal recording noise. Some of the signals showed some small drift and a simple correction algorithm based on drift compensation was used on the signals to minimize these deviations.

**Signal processing and time series extraction**

The cardiorespiratory signals for all 43 patients of the study were processed and analyzed using MATLAB® (v R2015) signal and data analysis software. Custom written computer software developed under MATLAB® was used to analyze both signals in order to extract the time series using an algorithm based on wavelet transform and analysis [14].

The original signals of the study were decomposed into a lower frequency signal using a discrete wavelet transform in order to remove noise and reduce signal details while preserving overall signal pattern to allow features to be more easily extracted. A biorthogonal Daubechies 4 (db4) wavelet was selected for signal decomposition.

The decomposed signals were automatically inspected for QRS complexes. By detecting these peaks in the down-sampled signal, the location of these peaks in the reconstructed signal had to be cross validated in the original ECG wave. The signal was then visually inspected and edited, if necessary. Ectopic beats were determined, removed and interpolated using an algorithm based on local variance estimation.

Respiratory flow signal was examined using an algorithm based on the zero-crossing of the respiratory flow signal. The respiratory flow signal was visually inspected and edited, if necessary.

After R peak and breathing episode detection on the original signals (ECG, RESP), a time series vector was then created from the intervals between R peaks and breathing episodes, as follows:

\[
RR(k_1) = (ECG_j_{1+1} - ECG_j_1) \quad (1)
\]

\[
T_{tot}(k_2) = (RESP_j_{2+1} - RESP_j_2) \quad (2)
\]

Where RR\( (k_1) \) and T\( tot(k_2) \) represent the time series of the ECG and respiratory signals, respectively; and \( j_1 \) and \( j_2 \) represent the location of an R peak and a breathing episode in the original ECG and RESP signals, respectively.

Since heart rate and breathing rate greatly vary in both frequency and variability, most of the time series values do not correlate and both time series values had to be synchronized in order to correctly apply joint symbolic dynamics [12]. An interpolation of the original time series was made using an estimation method based on local values at a 250 Hz frequency (the sampling frequency for the original ECG and RESP signals).
This new $RR(n)$ and $T_{tot}(n)$ time series allows us to compare and apply JSD to study the interactions between the two signals and their effect on the patient’s status.

A new bivariate vector $x$ is obtained from the sampling of each value of the resampled time series, as follows:

\[
x = \left\{ [x_n^c, x_n^r] \right\}_{n=0,1,...} \quad x \in \mathbb{R}^2
\]

\[
x_n^c = RR_n
\]

\[
x_n^r = T_{tot,n}
\]

From this synchronized sampled time series, $x$, a symbol sequence vector, named $s$, is obtained to look for symbols that might help us differentiate study groups.

**Joint symbolic dynamics: feature extraction**

After signal processing and analysis, the bivariate sampling vector for the time series was then transformed into a series of discrete symbols in order to extract information about the cardiorespiratory system and their relation to the patients’ respiratory pattern and clinical classification.
From the bivariate vector \( x \) (\( c, r \)), which represents the sample vector of the cardiac (\( c \)) time series \( RR(n) \) and the respiratory (\( r \)) time series \( T_{ra}(n) \), a bivariate symbol vector, \( s \), is obtained, as shown in Eq. (7). This vector \( s \) obtained by transforming \( x \) using a symbol alphabet. When the \( x \) time series vector showed an increase above 50% of the local standard deviation (measured in a 5 sample window), a 1 symbol was used to characterize an increase in the \( s \) time series vector. When the \( x \) time series vector showed a decrease or an increase below the 50% of the local standard deviation, a 0 was added to the \( s \) time series vector.

\[
x = \{(x^c_n, x^r_n)^T\}_{n=0,1,...} \quad x \in \mathbb{R} \quad (6)
\]
\[
s = \{(s^c_n, s^r_n)^T\}_{n=0,1,...} \quad s \in 0,1 \quad (7)
\]

This new symbol vector contains information about the behavior of the system and its dynamics. Each of the vector components, cardiac (\( c \)) and respiratory (\( r \)) vectors, contains a total of \( n \) symbols, which are divided into words \( (w_k) \) of length 3, with an overlap of \( \tau \) (0, +1, +2) symbols. A raw 4 dimensional matrix for word occurrence was then calculated using the bivariate sampling vector \( s \) for all patients (third matrix dimension) across groups for each value of symbol overlapping (fourth matrix dimension, \( \tau = 0, +1, +2 \)). The first two dimensions of the matrix span an 8x8 matrix \( W_{raw} \) (rows: \( c \)-cardiac word types; columns: \( r \)-respiratory word types), which ranges from word type \([000,000]^T\) to \([111,111]^T\). The method for obtaining the occurrence matrix \( W \) and the symbol vector \( s \) from the time series \( x \) is illustrated in Figure 2.

![Figure 2. Process of joint symbolic dynamics (JSD).](image)

To compare word-type distributions between data sets of different length, the sum of all counted words for each group was normalized to 1. Word probability occurrence, \( p(W_{c,r}) \), was calculated separately for each symbol overlapping value (\( \tau = 0, +1, +2 \)). The sum of each row in \( W_{raw} \) is computed as \( p(W_c) \) and represents the occurrence probability of each word from the cardiac time series, as follows:

\[
p(w_c) = \sum_c p(W_{c,r}) \quad (8)
\]

Where \( c \) takes binary values from “000” to “111”.

The sum of each column in \( W_{raw} \) is computed as \( p(W_r) \) and represents the occurrence probability of each word from the respiratory time series.

\[
p(w_r) = \sum_r p(W_{c,r}) \quad (9)
\]

Where \( r \) takes binary values from “000” to “111”.

![Figure 3. Process of joint symbolic dynamics (JSD).](image)

Transformation of the bivariate sample vector \( x \) \((ECG = RR \; intervals \; [s]; \; RESP = respiratory \; episode \; interval \; [s])\) into the bivariate symbol vector \( s \) (0: equal, decreasing or increasing values below 50% of the local standard deviation; 1: increasing values above 50% of the local standard deviation).

Words share 2 symbols (overlapping value \( \tau = +2 \)).

**RESULTS**

After the word occurrence matrix \( W \) was obtained, word probability distribution was plotted for each group to see if any significant difference across patient groups could be observed. A simple statistical analysis was then conducted on this matrix to obtain symbols with statistical significance across the groups. Single word types (cardiorespiratory words), as well as individual cardiac and respiratory symbols were examined.
across all groups to look for features that might prove valuable in patient classification. Individual words were then compared across groups to find those that were of statistical significance.

A Kruskal-Wallis test was performed for each individual parameter for parameter comparison across all groups. For parameter comparison between paired groups a simple Mann-Whitney U test was performed for each of the groups against the others.

Considering the problem of multiple testing, the necessary local significance level of a single parameter from an observed 64 dimensional parameter space had to fulfill Bonferroni’s inequality to guarantee a global significance.

**Feature selection**

Different discrimination functions were constructed using the most significant parameters from the symbolic analysis of the signals. The initial LDA model was created using the most prominent features from the statistical analysis yielding a 93.02% accuracy in patient classification. Word occurrence probability significance across groups was used to select words for the initial LDA classification model. Each parameter was selected only if it was present in at least two of three overlapping value distributions in order to reduce the possibility of overlapping dependent symbols.

The best features to discriminate among groups are related to cardiorespiratory interactions or cardiac words alone (Table 1). While the cardiac word for increasing intervals ($W_c = ‘111’$) shows one of the highest probability significance across groups using a Kruskal-Wallis test, the most common cardiac words among the selected features for the classification models are those that represent constantly alternating intervals and further support the importance that cardiac variability plays in the poor respiratory regulation in patients with periodic breathing.

**Classification models** were validated using a leave-one-out (LOO) cross-validation method. The best discriminant model was constructed using a stepwise feature selection to optimize accuracy from the initial model based on significant symbols from the Kruskal-Wallis test (Table 2). This model used 10 features to discriminate among groups and yielded a 97.67% accuracy in patient classification. 9 of the 10 features selected are related to cardiorespiratory interaction, while the $f_{85}$ parameter is the number of cardiac words ($w$) whose probability of occurrence $p(w)$ is lower than 0.03.

### Table I. Words with significant probability occurrence across all three groups at an overlapping level $\tau = +2$

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$p(n(w))$</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0046</td>
<td>$p_n(w_{000}, r_{000})$</td>
<td></td>
</tr>
<tr>
<td>0.0092</td>
<td>$p_n(w_{001}, r_{010})$</td>
<td></td>
</tr>
<tr>
<td>0.0094</td>
<td>$p_n(w_{111}, r_{111})$</td>
<td></td>
</tr>
<tr>
<td>0.0175</td>
<td>$p_n(w_{101}, r_{110})$</td>
<td></td>
</tr>
<tr>
<td>0.0344</td>
<td>$p_n(w_{100}, r_{110})$</td>
<td></td>
</tr>
<tr>
<td>0.0445</td>
<td>$p_n(w_{010}, r_{110})$</td>
<td></td>
</tr>
</tbody>
</table>

### Table II. Best feature subset for patient group discrimination using LDA

<table>
<thead>
<tr>
<th>Features</th>
<th>$p$</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>0.0052</td>
<td>$p_n(w_{000}, r_{000})$</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>0.0100</td>
<td>$p_n(w_{001}, r_{001})$</td>
</tr>
<tr>
<td>$f_{12}$</td>
<td>0.0175</td>
<td>$p_n(w_{011}, r_{001})$</td>
</tr>
<tr>
<td>$f_{18}$</td>
<td>0.0384</td>
<td>$p_n(w_{001}, r_{010})$</td>
</tr>
<tr>
<td>$f_{26}$</td>
<td>0.0098</td>
<td>$p_n(w_{001}, r_{011})$</td>
</tr>
<tr>
<td>$f_{51}$</td>
<td>0.0089</td>
<td>$p_n(w_{010}, r_{110})$</td>
</tr>
<tr>
<td>$f_{53}$</td>
<td>0.0061</td>
<td>$p_n(w_{100}, r_{110})$</td>
</tr>
<tr>
<td>$f_{54}$</td>
<td>0.0364</td>
<td>$p_n(w_{101}, r_{110})$</td>
</tr>
<tr>
<td>$f_{55}$</td>
<td>0.0070</td>
<td>$p_n(w_{110}, r_{110})$</td>
</tr>
<tr>
<td>$f_{85}$</td>
<td>0.0177</td>
<td>$\Sigma(p_n(w)) &lt;0.03$</td>
</tr>
</tbody>
</table>
DISCUSSION

JSD analysis revealed significant cardiorespiratory coupling patterns across the 3 study groups (non-periodic breathing, periodic breathing and Cheyne-Stokes breathing). Of these parameters, only one \( P_n(w_{111}) = 0.0031 \) fulfilled a conservative Bonferroni-Holm adjustment, which suggests a true significance across patient groups. This parameter proved a good discriminant between nPB and the other groups (PB and CSR), but showed poor discrimination power between PB and CSR. Individual cardiorespiratory coupling words including the respiratory word \( w_{110} \) showed good discrimination power between PB and CSR groups, particularly \( w_{100}, r_{110} \) and \( w_{010}, r_{110} \).

The coupling of other physiologically relevant signals could potentially improve not only the classification model but also provide more information regarding the mechanisms involved in the poor respiratory regulation in PB and CSR patients. Patients with congestive heart failure and Cheyne-Stokes respiration have increased pulmonary vascular pressures and a JSD analysis using a blood pressure signal could potentially improve patient characterization. Patient characterization using JSD coupling ECG and blood pressure signals have already been developed providing good results [15-17].

CONCLUSIONS

In conclusion, by applying JSD we were able to provide detailed information about short-term regulatory mechanisms of respiratory patterns. From the results we can observe that the most relevant features selected prioritize cardiorespiratory variability rather than cardiac or respiratory variability alone, which supports the idea that abnormal breathing patterns originate from physiological abnormalities in chronic heart failure [18]. An adequate control of the cardiac status of a patient could prove a beneficial influence in CSR with improved sleep quality and quality of life for patients [19], and might help prevent the development of CSR in patients with normal breathing patterns.

ACKNOWLEDGEMENTS

The author T. F. Molina-Ramírez would like to thank Beatriz Giraldo of the Institute for Bioengineering of Catalonia (IBEC) for her support and invaluable guidance during the course of this research, although any possible error would be our own.
REFERENCES


