Journal of Pharmacological and Toxicological Methods 84 (2017) 78–85

Contents lists available at ScienceDirect

Journal of Pharmacological and Toxicological Methods

journal homepage: www.elsevier.com/locate/jpharmtox

Research article

An intelligent prognostic system for analyzing patients with paraquat poisoning using arterial blood gas indexes

Lufeng Hu [a,1], Feiyan Lin [b,1], Huaizhong Li [c], Changfei Tong [d], Zhifang Pan [e], Jun Li [d], Huiling Chen [d,*]

[a] Department of Pharmacy, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China
[b] Centre Laboratory, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China
[c] Department of Computing, Lishui University, Lishui 323000, Zhejiang, China
[d] College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China
[e] Information Technology Center, Wenzhou Medical University, 325035, China

ARTICLE INFO

Article history:
Received 18 August 2016
Received in revised form 26 October 2016
Accepted 19 November 2016
Available online 21 November 2016

Keywords:
Support vector machine
Particle swarm optimization
Arterial blood gas
Paraquat
Prognostic system
Outcome prediction

ABSTRACT

The arterial blood gas (ABG) test is used to assess gas exchange in the lung, and the acid-base level in the blood. However, it is still unclear whether or not ABG test indexes correlate with paraquat (PQ) poisoning. This study investigates the predictive value of ABG tests in prognosing patients with PQ poisoning; it also identifies the most significant indexes of the ABG test. An intelligent machine learning-based system was established to effectively give prognostic analysis of patients with PQ poisoning based on ABG indexes. In the proposed system, an enhanced support vector machine combined with a feature selection strategy was developed to predict the risk status from a pool of 103 patients (56 males and 47 females); of these, 52 subjects were deceased and 51 patients were alive. The proposed method was rigorously evaluated against the real-life dataset in terms of accuracy, sensitivity, and specificity. Additionally, the feature selection was investigated to identify correlating factors for the risk status. The results demonstrated that there were significant differences in ABG indexes between deceased and alive subjects (p-value <0.01). According to the feature selection, we found that the most important correlated indexes were associated with partial pressure of carbon dioxide (PCO_2). This study discovered the relationship between ABG test and poisoning degree to provide a new avenue for prognosing PQ poisoning.

1. Introduction

Arterial blood gas (ABG) tests are used to evaluate gas exchange in humans by determining the partial pressure of carbon dioxide and oxygen, blood pH, bicarbonate level, and so on (Pishbin et al., 2015). It is normally performed by taking blood from an artery, such as a radial artery or femoral artery to assess the subject's oxygen delivery ability, ventilatory function integrity, and acid-base equilibrium systems. The information it provides is very useful in respiratory therapy and in judging disease severity (Rahimi, Bidabadi, Mashouf, Seyed Saadat, & Rahimi, 2014, Park et al., 2015).

Paraquat (1,1′-dimethyl-4,4′-bipyridium dichloride, PQ), the most toxic herbicides for humans, has caused numerous deaths around the world (Soloukides, Moutzouris, Kassimatis, Metaxatos, & Hadjiconstantinou, 2007). Severely poisoned patients usually die from lung damage characterized by pulmonary fibrosis (Yoon, 2009). Therefore, ABG tests must be performed on patients with PQ poisoning as a routine test to assess gas exchange in the lungs.

To date, only Huang et al. (Huang & Zhang, 2011) have used Cox Regression analysis to identify the correlation between ABG indexes and patients' survival time. However, they only included partial ABG indexes, and they did not thoroughly investigate the correlation between ABG indexes and PQ poisoning. It is still unclear which indexes are significant in prognosing PQ poisoning.

In this study, we developed an intelligent system based on ABG indexes to evaluate PQ poisoning prognoses. Prognosing PQ poisoning can be formulated as a classification problem that predicts an alive or deceased status. Obviously, due to the nature of the original prognosis, this classification problem is multivariate and has complex relationships. Therefore, the support vector machine (SVM) is a good classification method candidate in this scenario. SVM (Boser, Guyon, & Vapnik, 1992, Vapnik, 1995) is rooted in the Vapnik-Chervonenkis theory and the structural risk minimization principle. Specifically, by exploring the tradeoff between minimizing the training set error and maximizing the margin, SVM seeks the best generalization ability while avoiding over fitting. Due to its success in this task, SVM has been applied in a wide variety of classification tasks (Bystritsky, Craske, Maidenberg, Vapnik, & Shapiro, 1995, Shin, Lee, & Kim, 2005, Chen, Liu et al., 2011, Chen, Yang et al., 2011). In particular, it has been reported that SVM is effective at many medical diagnostic tasks (Ubeyli, 2007, Akay, 2009,

* Corresponding author.
E-mail address: chenhuiling.jlu@gmail.com (H. Chen).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.vascn.2016.11.004
1056-8719/

Maglogiannis, Zafiropoulos, & Anagnostopoulos, 2009, Chen, Liu et al., 2011, Chen, Yang et al., 2011, Chen et al., 2012a, Chen et al., 2012b). It has been demonstrated that SVM's performance can be substantially improved by establishing proper model parameter settings (Keerthi & Lin, 2003, Chen et al., 2014). The two key parameters of SVM, such as the penalty parameter and the kernel bandwidth of the kernel function, should be properly determined prior to its application to practical problems. The first parameter, the penalty parameter C, determines the trade-off between the fitting error minimization and the model complexity. The second parameter, the kernel bandwidth γ, defines the non-linear mapping from the input space to some high-dimensional feature space.

We attempted to adopt the famous swarm intelligence based method, the particle swarm optimization algorithm (PSO), to optimize the two parameters of SVM, since the background of PSO is intuitive, and it is simple and easy to realize. Furthermore, we have also adopted three other benchmark methods including SVM trained by the grid search (Grid-SVM), back propagation neural network (BPNN) and k-nearest neighbor methods (KNN) to perform the same prognostic task for comparison. In addition, we explored the possibility of adopting feature selection in pre-processing before the SVM model was constructed to potentially identify the significant correlating factors in prognosing PQ poisoning.

2. Individuals

2.1. Ethics statement

This study involved patients with PQ poisoning who had acute contact with PQ and were hospitalized in the emergency intensive care unit (EICU) of the First Affiliated Hospital of Wenzhou Medical University from January 1, 2013 to October 31, 2015. Our experiment was approved by the Medical Ethics Committee of The First Affiliated Hospital of Wenzhou Medical University, and conducted in accordance with the Declaration of Helsinki. All data relating to the PQ patients were anonymously recorded in a Microsoft Excel spreadsheet recorded by two emergency physicians.

2.2. Data preparation

This study included 103 PQ poisoning patients, including 56 male and 47 female patients aged 14 to 78 years, with a mean age of 35.74 ± 14.26 years, and a median age of 33 years. The patients' initial plasma PQ concentration and ABG were immediately determined when they arrived in the hospital emergency room (ER). Patients were excluded from the study if their plasma PQ concentration was not detected within the first 24 h of PQ poisoning. The study also excluded patients with ventilatory and/or metabolic diseases, such as emphysema and pulmonary heart disease.

All patients included in this study received the same treatment methods. They were given hemoperfusion (HP) treatment, immunesuppressing drugs, a large dose vitamin C, stomach and liver-protective drugs, and antimicrobials. They were separated into two different groups, the survival group and the deceased group, according to their treatment outcome, which was assessed by vital signs including respiratory rate, heart rate, and blood pressure.

The plasma PQ concentrations were determined by an Agilent 1260 Infinity High Performance Liquid Chromatography (HPLC) system, which was equipped with an automatic sampler, quaternary pump, column oven, and diode array detector (DAD). Firstly, 20 μL 5-bromopyrimidine (100 μg/mL, internal standard) was added into 1.5 mL centrifuge tube which contained 200 μL plasma samples. Then, 100 μL trichloroacetic acid-methanol (9:1, v/v) was added and vortex mixed for 0.5 min. After the mixture was centrifuged at 14,900g for 10 min, 20 μL supernatant was injected into the HPLC system for analysis. The area of PQ was calculated by calibration curves to obtain the plasma concentration. The validation of this method included intra-day and inter-day precision, recovery, stability. For more detailed information, please refer to our previous published article (Hu et al., 2015). ABG was determined by an ABL800 Fully Automatic Blood-gas Analyzer.

This study used 21 ABG indexes (feathers), as illustrated in Table 1. Statistical analysis was performed using SPSS 17 software. The ABG indexes in the survival and deceased groups were listed as mean values (mean) and standard deviations (SD), listed in Table 2. The differences in ABG indexes between the two groups were analyzed by variance analysis. If the data showed a normal distribution, independent sample tests were used, or the nonparametric test was applied to analyze the difference. $p < 0.05$ was considered a statistical difference.

3. Methodology introduction

The proposed system provides an efficient and accurate prognostic tool to predict whether patients belonged to the decreased or survive group using ABG indexes. Fig. 1 illustrates the flowchart of the proposed system, which is comprised of two main stages. In the first stage, feature selection was used to identify the informative indexes from the pool of ABG indexes. Then, different feature sets were fed into the optimal SVM model trained by PSO strategy to discover the most significant related indexes associated with the risk status of the patients with PQ poisoning. Finally, the obtained predictive model conducted the prognostic task using the most discriminating features.

3.1. The feature selection phase

Feature selection can improve classifier model performance and provide much more effective models with a simpler structure; it also provides deeper insights into the problem's underlying mechanism (Guyon & Elisseeff, 2003). In our study, the feature selection process was performed through the Fisher Score method (Duda, Hart, & Stork, 2012). As a commonly used supervised feature ranking method, Fisher Score employs the Fisher criterion to determine the most discriminative features. The score for each feature in a two-class problem can be defined as follows:

$$F = \frac{\sum_{i=1}^{2} n_i(\mu_i - \mu)^2}{\sum_{i=1}^{2} n_i(\sigma_i)^2}. \quad (1)$$

Table 1
List of ABG features used in this study and their abbreviation.

Number	Features	Abbreviation
F_1	Blood pH	pH
F_2	Blood pH at patient temperature	pHT
F_3	Partial pressure of carbon dioxide	PCO_2
F_4	Partial pressure of carbon dioxide at patient temperature	PCO_2T
F_5	Oxygen partial pressure	PO_2
F_6	Oxygen partial pressure at patient temperature	PO_2T
F_7	Base excess	BE
F_8	Base excess extracellular fluid	BEecf
F_9	Buffer base	BB
F_{10}	Bicarbonate	HCO_3^-
F_{11}	Total carbon dioxide	TCO_2
F_{12}	Standard bicarbonate	$st.HCO_3^-$
F_{13}	Standard pH value	st.PH
F_{14}	Oxygen content	O_2cont
F_{15}	Oxygen saturation	O_2sat
F_{16}	Pulmonary arterial oxygen partial pressure	$paPO_2$
F_{17}	Pulmonary arterial PO_2 at patient temperature	$paPO_2T$
F_{18}	Alveolar-arterial oxygen tension difference	$A\text{-}aDO_2$
F_{19}	Alveolar-arterial oxygen tension difference at patient temperature	$A\text{-}aDO_2T$
F_{20}	Fraction of inspiration O_2	FIO_2
F_{21}	Hydrogen ion concentration	cH^+