G Model
JAB 163 No. of Pages 11
ARTICLE IN PRESS
Journal of Applied Biomedicine xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Journal of Applied Biomedicine
journal homepage: www.elsevier.com/locate/jab

Original research article

Nonlinear Heart Rate Variability based artificial intelligence in lung cancer prediction

Reema Shyamsunder Shukla, Yogender Aggarwal*

Birla Institute of Technology, Department of Bio-Engineering, Mesra, Ranchi, India

ARTICLE INFO

Article history:
Received 14 July 2017
Received in revised form 13 November 2017
Accepted 11 December 2017
Available online xxx

Keywords:
Artificial intelligence
Eastern cooperative oncology group
Heart rate variability
Lung cancer
Nonlinear analysis

Introduction

Cancer is the development of an irregular pattern of cells resulting into abnormal tissue growth called neoplasia, which can further be categorized into potentially malignant or malignant cancer (Weinberg, 1996). Lung cancer arises mainly in the epithelial lining of the bronchial tree and originates in the main and lobar bronchi (Garland et al., 1968). Squamous Cell carcinoma, adenocarcinoma and small cell carcinoma are different histological types of lung cancer. Malignancy due to lung cancer is the utmost cause of mortality amongst patients. The incidences and deaths due to lung cancer are increasing sharply among men and women with air pollution, smoking and tobacco consumption is the leading factor (De Couck et al., 2013; Kim et al., 2010). Age and gender are assumed as independent factors which may affect the severity of illness based on several symptoms (Walsh et al., 2000).

Chronic stress has been suggested to be common in cancer patients (Li et al., 2013) and associated with autonomic nervous system (ANS) dysfunction (Guo et al., 2015; Walsh and Nelson, 2002). Review of literature suggested reduced heart rate variability (HRV) (Kim et al., 2010; Schlenker et al., 2014) in cancer subjects. HRV has been evaluated as beat-to-beat interval derived from electrocardiogram (ECG). HRV had previously been described as the technique to assess the interaction among sympathetic (SNS) and parasympathetic nervous system (PNS) (Acharya et al., 2002; Aggarwal et al., 2014). ANS works on feedback loop mechanism to control the physiological variables like heart rate, which exhibits the nonlinear dynamics (Aggarwal et al., 2014; Signorini et al., 2001). ANS tone varies instantly under cancer condition to meet undue metabolic demand with higher SNS and lower PNS activity (Aggarwal et al., 2014; Schlenker et al., 2014). The alteration in ANS tone reflects pathological status of patients in critical state (Chiang et al., 2013).

Literature review demonstrated the application of nonlinear HRV in depicting the diseases at its early stages (Acharya et al., 2002; De Souza et al., 2014; Mohebbi et al., 2011; Roy and Ghatak, 2013; Schlenker et al., 2014; Yeh et al., 2006; 2010). The cardiac signals are nonlinear time series and require nonlinear analysis to study even minor fluctuation in signal. It has suggested to be robust method thus preferred over time and frequency domain analysis (De Souza et al., 2014; Roy and Ghatak, 2013; Yeh et al., 2010). Few studies have been reported in identification of cancer using time and frequency domain analysis of HRV (De Couck et al., 2016; Guo et al., 2015; Kim et al., 2010; Walsh and Nelson, 2002). However, literature on nonlinear analysis is still obscure. Review of literature also suggested the application of artificial intelligence (AI) in prognosis of different types of cancer (Cosma et al., 2016; Utomo et al., 2014). However, to the best of authors' knowledge, none of the study has utilised HRV indices as input to machine learning. Thus, the objectives of present work were to identify alterations in nonlinear HRV features with performance status (PS), gender and age in lung cancer, if any and to discriminate their clinical data with the help of artificial neural

* Corresponding author at: Birla Institute of Technology, Department of Bio-Engineering, Mesra, Ranchi, 835215, India.
E-mail address: yogender.aggarwal@gmail.com (Y. Aggarwal).

https://doi.org/10.1016/j.jab.2017.12.002

network (ANN) and support vector machine (SVM) for design of an automated diagnostic tool for lung cancer.

Materials and methods

Subjects and biosignal acquisition

A total of 104 lung cancer subjects and 30 healthy controls have participated in this study. The patient consent was signed before their demographics and clinical measures and history was recorded with the help of expert clinicians. Lead II electrocardiogram was recorded for 5 min for each subject with standard electrode placements in supine position from 11 a.m. to 12 at 24 °C at controlled breathing rate. The signal was recorded with the help of MP45 bioamplifier (Biopac System Inc., USA). The duration of recording has been opted as suggested by Voss et al. (2015). The recorded signal was band pass filtered with cut-off frequency of 0.5 to 35 Hz and digitized at 200 samples per second. Ag/AgCl disposable electrodes were used for surface recording using SS2LA lead that connects to the amplifier. The ECG recording was performed in accordance with the ethical standards of Declaration of Helsinki and a written consent has been obtained from the patients and volunteers. The tachogram was derived from the recorded ECG using Acknowledge 4.0 (Biopac Systems Inc, USA) with R-wave threshold level set to 0.5 and interpolated with cubic-spline resampling frequency of 8 samples/s (Aggarwal et al., 2012; Shukla and Aggarwal, 2017). The cases were composed of lung cancer stages as follows Stage II 2.73%, Stage III 15.45%, Stage IV 81.82%. Histopathologically, adenocarcinoma comprised 42.73%, Squamous cell carcinoma (SCC) 30.91%, small cell lung cancer (SCLC) 12.73%, poorly differentiated carcinoma 13.64%. There were 59 males and 51 females. The oxygen saturation of most of the patients was approximately an average of 95%. The average left ventricular ejection fraction (LVEF) was 55%. The segregation was done on the basis of gender and age. The age group evaluated was 0 to 40, 41 to 64 and more than or equal to 65 years. Cardiac disorder, diabetes, hypertensive and mental illness patients were excluded from the study. The obtained tachogram was used to evaluate nonlinear parameters of HRV at different PS scales using Kubios HRV 2.0 (University of Kuopio, Finland). The details of HRV analysis have been discussed earlier, which is also used for this study (Tarvainen et al., 2014).

Eastern cooperative oncology group performance status scale

Two PS scales are routinely used in oncology: Karnofsky (KPS) and Eastern Cooperative Oncology Group (ECOG). The KPS ranges from 0 to 100 to define 11 different PS levels from dead (0) to fully normal functioning (100) with increment of 10. ECOG ranges from 0 (fully ambulatory without symptoms) to 5 (dead) with six levels (Lilenbaum et al., 2008). Level 0–4 refers to normal day-to-day activity, exhibits symptoms, difficulty in day-to-day activities with less than 50% bedridden, bedridden (more than 50%) and completely bedridden, respectively (Sorensen et al., 1993).

Feature extraction for nonlinear HRV analysis

The parameters of nonlinear HRV measurement such as Poincare plot (PP), approximate (ApEn) and sample entropy (SampEn), detrended fluctuation analysis (DFA), correlation dimension (CD), and recurrence plot (RP) have been investigated.

PP estimates the similarity of successive RR intervals by fitting an ellipse to obtain graphical plot. Each RR interval is plotted as a point. The point lying on the line of identity have equal RR interval. However, points above or below the line represent high or low RR interval, respectively. The standard deviation of points perpendicular and along the line is discussed as short term variability (SD_1) and long term variability (SD_2), respectively. Entropy (ApEn and SampEn) has been used to study the signal complexity, which represents randomness of heart activity. Larger the value higher will be irregularity and smaller value suggest more regular signal. Shannon Entropy (ShanEn) in similar lines quantified the persistence of short binary symbols (length N 5) measured in 10 min intervals. DFA prevents wrong detection of long range correlations and eliminates constant or linear trends from the time series. It correlates to identify the similarity in non-stationary signal. The fluctuation (α) is root-mean square of detrended time series, measured at different length. The short (α_1) and long (α_2) fluctuation were calculated at $4 \leq n \leq 16$ and $16 \leq n \leq 64$ detrended time series segment, respectively. CD value quantitatively measured the obtained line patterns from plotting of HR with delayed HR. Point fixes to a point for steady HR otherwise different line patterns will be observed. The slope of line pattern saturates at finite value with increased embedding value. The value of 10 was found suitable for estimating the embedding dimension from phase space plot. RP signifies short line

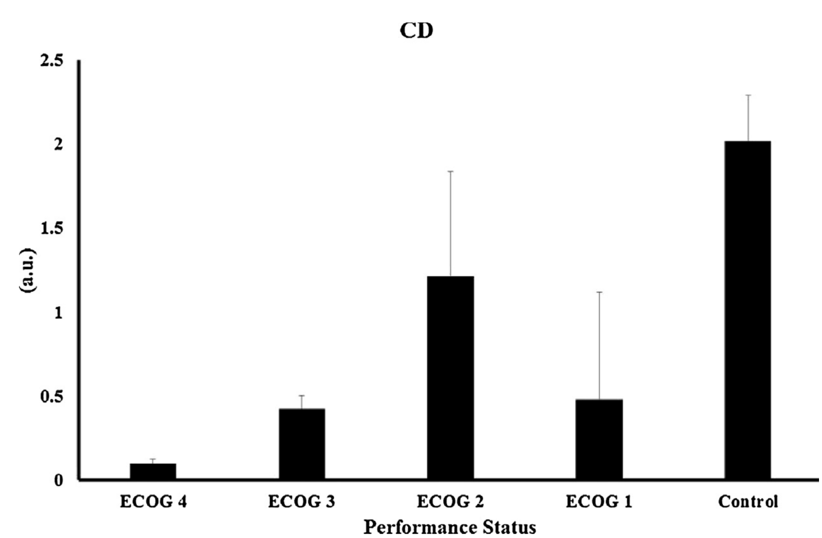

Fig. 1. Correlation dimension (CD) analysis with progression in lung cancer.