Journal of Applied Biomedicine 16 (2018) 145–155

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 21 December 2017

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

ABSTRACT

Lung cancer is uncontrolled growth of cells that occurs due to exposure to smoke, radiation and chemicals, which causes chronic stress and associated with impaired autonomic nervous system. Nonlinear heart rate variability (HRV) analysis has been suggested to uncover the performance status of lung cancer subjects and distinguish them from healthy controls. The present work obtained tachogram from recorded electrocardiogram of 104 lung cancer subjects and 30 healthy controls to extract HRV indices. The obtained results suggested lowered HRV (altered autonomic nervous system tone) values from Eastern Cooperative Oncology Group (ECOG) 1 to ECOG4. Subject males had higher HRV measures than their female counterparts. The HRV parameters decreased from ECOG PS of 1 to 4. Control females had higher HRV measures than control males. There was no association between age and HRV measures. Statistically, nonlinear HRV features were observed significant. ANN exhibited ECOG1 83.3%, ECOG2 50%, ECOG3 90%, ECOG4 95% and Controls 86.7%. The prediction analysis using artificial neural network (ANN) and support vector machine (SVM) scoring an accuracy of 93.09% and 100% with nonlinear HRV indices as input thus has been suggested to be a tool of prognostic importance.
© 2017 Faculty of Health and Social Sciences, University of South Bohemia in Ceske Budejovice. Published by Elsevier Sp. z o.o. All rights reserved.

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.,

* Author for correspondence: 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
1214-021X/© 2017 Faculty of Health and Social Sciences, University of South Bohemia in Ceske Budejovice. Published by Elsevier Sp. z o.o. All rights reserved.

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 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 segments parallel to the main diagonal line that represents the time for which two points were close to each other with length of diagonals. RP analysis considers 'lengths of the diagonal lines'. The maximum diagonal length is known as L_{max}. The tangential motion of the trajectory also forms diagonal lengths that have been removed with a L_{min} threshold of 2. L_{mean} is the mean diagonal line length. The value of embedding dimension, embedding lag and Euclidean distance were fixed to 10, 1 and $\sqrt{}$mSD (standard deviation of RR time series). The most important quantitative measure of RP is the recurrence rate (REC) defined as the ratio of ones to zeros in the RP matrix. Determinism (DET) is the ratio of vertical lines which is formed by the recurrence points to the total recurrence points (Marwan et al., 2002; Tarvainen et al., 2014).

Classification

Artificial neural network (ANN)

A three-layered feed-forward backpropagation ANN has been used to classify different PS of lung cancer (ECOG1–4 and controls) based on features extracted from nonlinear HRV analysis. The neurons in each layer are interconnected with associated weights. The weight of each neuron modifies with the output of first layer neurons before the next processing takes place. The sigmoidal function was used in this network. The architecture of network was optimized with 9:4:5, where, the inputs selected were the values of CD, SD_1, SD_2, SD_1/SD_2, α_1, α_2, L_{max}, SampEn and ApEn indices.

The network was tested with varied learning rate from 0.01 to 0.9 and hidden layer nodes from 1 to 500 and optimized with learning rate parameter of 0.4 and hidden layer nodes of 4. The inbuilt Matlab function (Lavenberg-Marquardt algorithm) has been used to implement the network. The network was trained for one thousand iterations. The supervised network has suggested to be advantageous with effective training and better system behaviour (Aggarwal et al., 2007; Sinha et al., 2007a, 2007b; Shukla and Aggarwal, 2017).

Support vector machine (SVM)

SVM was implemented with radial basis function (RBF) using windows. SVM has been utilised for data classification, which is simpler to use than ANN. A hyperplane divides a pattern into two classes. A nonlinear SVM maps input vectors into a high dimensional space $z = \psi(x)$, where equation of the hyperplane is given as $F(x) = w.\ \psi(x) + bias$. The margin that separates the hyperplane and the data is maximized denoted by $\frac{2}{||w||}$ where, w is weight. RBF kernel is given by

$$K(x, x') = f\left(\frac{1}{2}||x - x'||^2\right) \quad (1)$$

This includes Gaussian RBF kernel vectors near boundary are called support vectors. x' is the centre and σ is variance (Amari and Wu, 1999).

$$K(x, x') = e^{-\left(\frac{1}{2\sigma^2}||x-x'||^2\right)} \quad (2)$$

Statistical analysis

Student's *t*-test and analysis of variance (ANOVA) was executed to compare different ECOG PS among different nonlinear HRV parameters and on the basis of gender using R 3.0.1 (R Foundation for Statistical Computing, Vienna, Austria) at significance of $P \leq 0.05$ and $P \leq 0.0001$. All the data for nonlinear HRV parameters of lung cancer were shown as mean ± standard error (SE).

Results

In present work, indices of nonlinear HRV analysis were observed for different ECOG scale for lung cancer.

Nonlinear HRV features

The obtained results suggested the lower HRV and decrease in RR interval with increase in cancer severity from ECOG1 to ECOG4. CD (Fig. 1) and PP (SD_1, SD_2 and SD_1/SD_2) (Figs. 2–3) values were

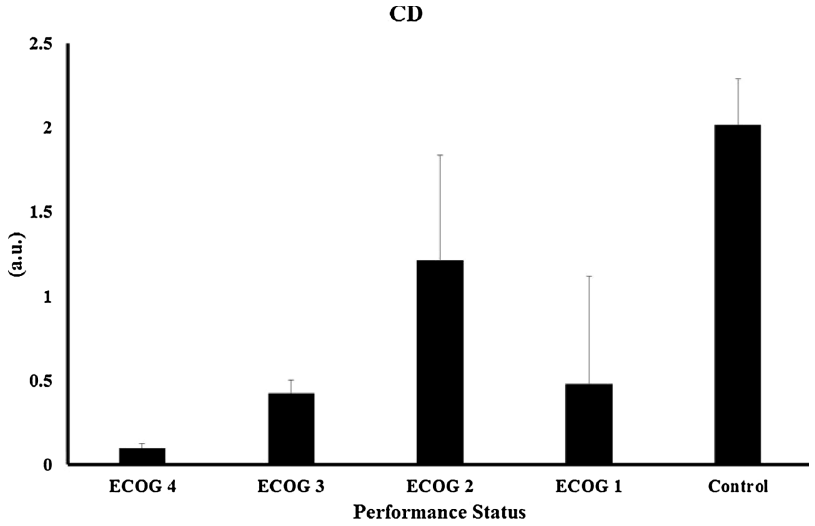

Fig. 1. Correlation dimension (CD) analysis with progression in lung cancer.

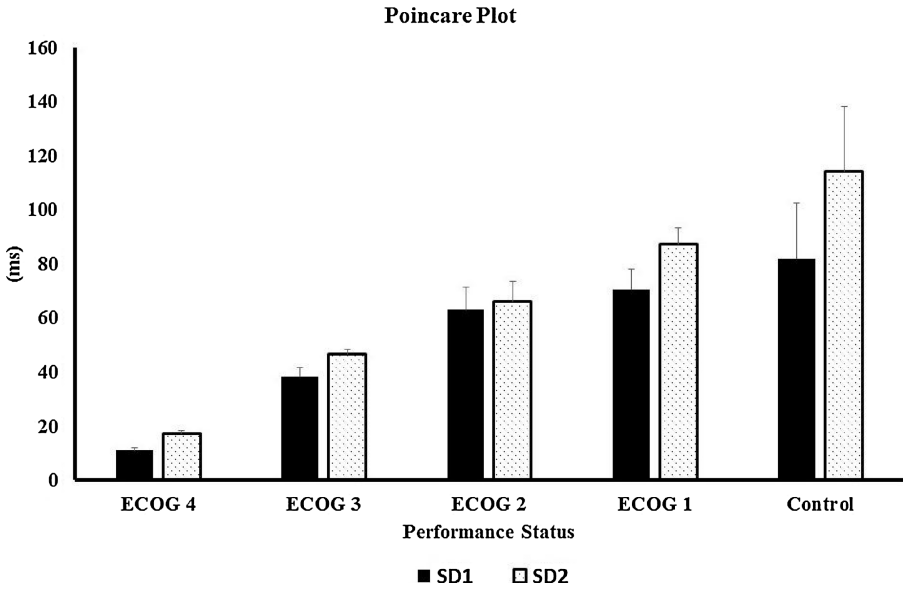

Fig. 2. Short term variability (SD_1) and long term variability (SD_2) analysis of Poincare plot.

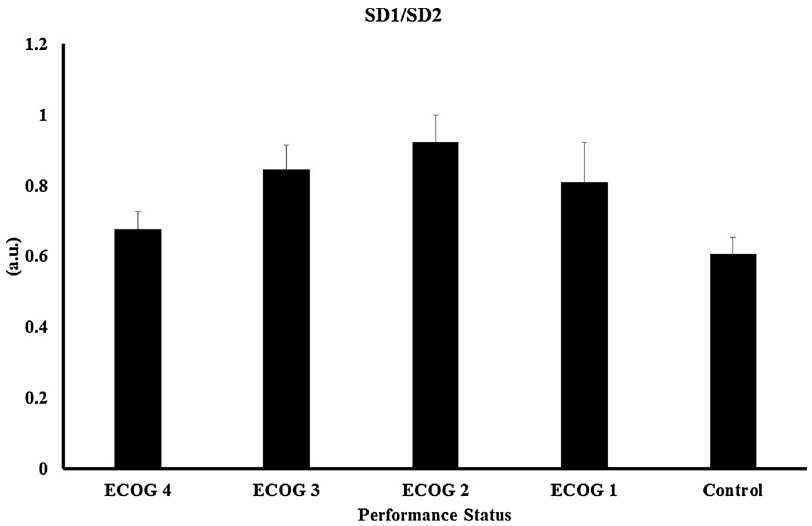

Fig. 3. Represents the ratio of parasympathetic to sympathetic activity ratio using Poincare plot (PP) analysis.

observed to reduce with ECOG 1 to 4. However, an exception was observed in ECOG1 of CD and SD_1/SD_2 ratio analysis with reduced value than ECOG2 and ECOG3, respectively. Further, RP L_{max} (Fig. 4) and entropy (ApEn and SampEn) (Fig. 5) value rises with disease progression suggested low HRV or slow variation in peak to peak interval with more rhythmic heart beats under diseased condition. Exception was found in ECOG1 PS of entropy that showed higher value than ECOG3 values. The α_1 and α_2 of DFA analysis increases with disease progression (Fig. 6). Where, α_1 and α_2 represents the parasympathetic and sympathetic tone, respectively. However, in α_1, exception was observed at ECOG3 with lesser than ECOG2 value. The α_2 value revealed to be higher than α_1 in ECOG3 and ECOG4 case. However, lower values were also observed in ECOG1 and ECOG2 cases where α_1 was higher than α_2. This higher value of α_2 indicates slower RR interval variation with higher degree of smoothness. Overall, results suggest higher SNS tone that increases with disease severity from ECOG1 to ECOG4 due to withdrawal of vagal tone. The graphical results obtained from Kubios HRV for PP and DFA analysis for different ECOG scale along with control group has been represented in Fig. 7–11 respectively. Control males have lower values of PP descriptors (SD_1, SD_2), SD_1/SD_2, ShanEn, L_{mean}, SampEn, REC and DET than control females whereas it is vice-versa in ApEn, L_{max} and CD. Subject males have higher values of SD_1, SD_2, SampEn, REC and CD than their female counterparts. All the values of non-linear measures based on gender are given in Table 1. The relationship between age and all HRV measures could not be established because there was no uniformity in pattern. But, in all age groups of more than or equal to 65 years, 41 to 64 years and 0 to 40 years, the values of SD_1, SD_2, L_{mean}, REC, DET, ShanEn and CD decreased in ECOG4 from ECOG3. All the values on the basis of age is given in Table 2.

Statistical analysis

Students' two-tailed *t*-test was conducted on non-linear measures of HRV at statistical significance of 0.05 and termed significant. Few non-linear domain measures were also tested and found significant at 0.0001 and was termed as highly significant. Statistically, CD was observed to be significant between ECOG3 and ECOG4 (p = 1.43e-04). Further, control group & ECOG1, Control & ECOG3 and Control & ECOG4 were highly significant. SD_1 was

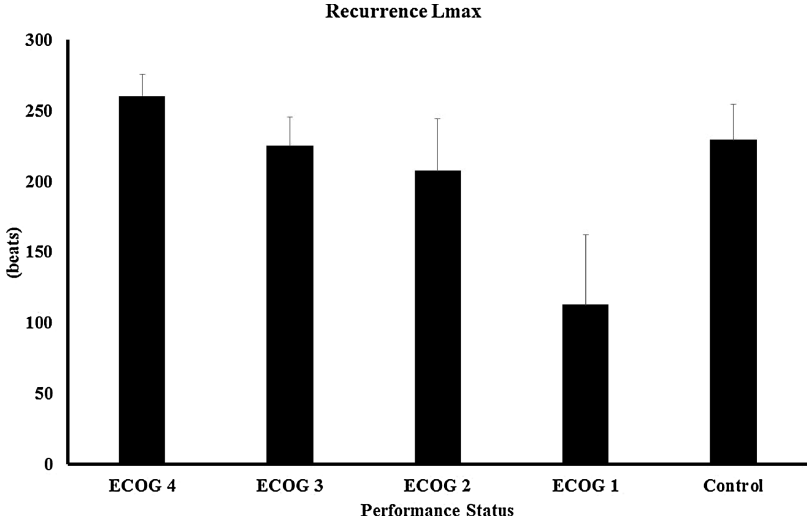

Fig. 4. Recurrence plot analysis for L_{max} parameter with severity of disease.

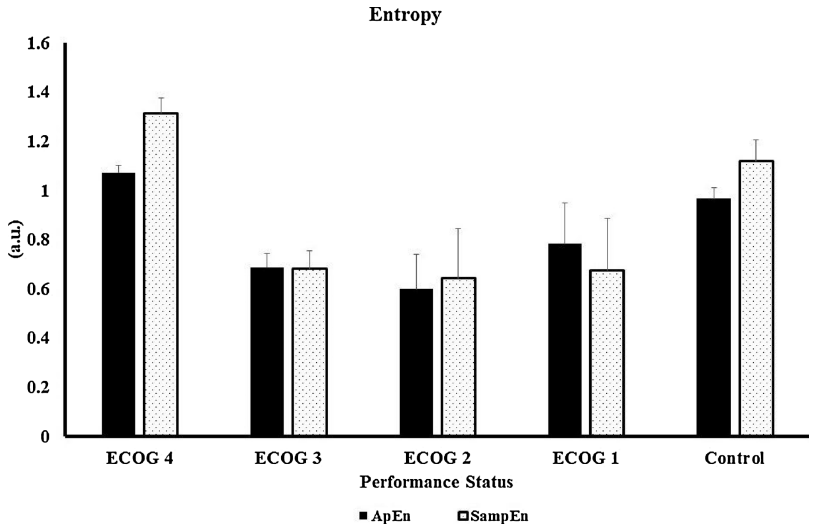

Fig. 5. Entropy [Approximate entropy (ApEn) and Sample entropy (SampEn)] analysis shows the increase in entropy with illness.

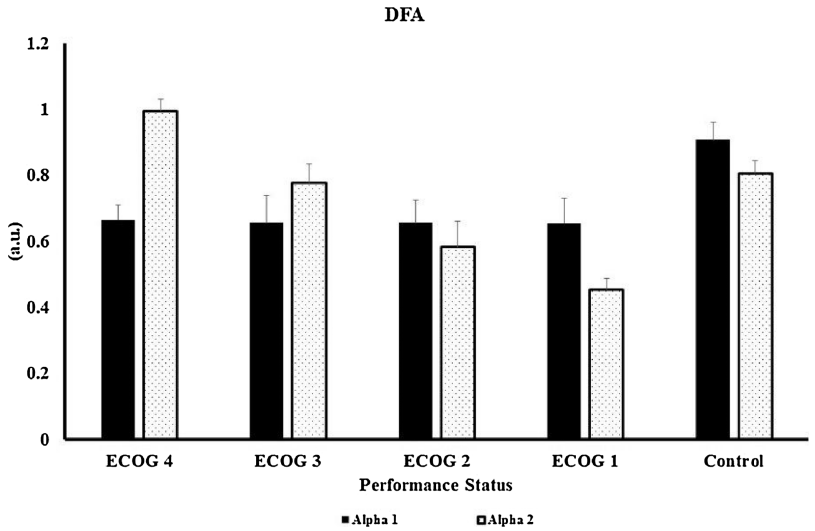

Fig. 6. Represents the DFA exponent analysis with performance status of cancer stage.

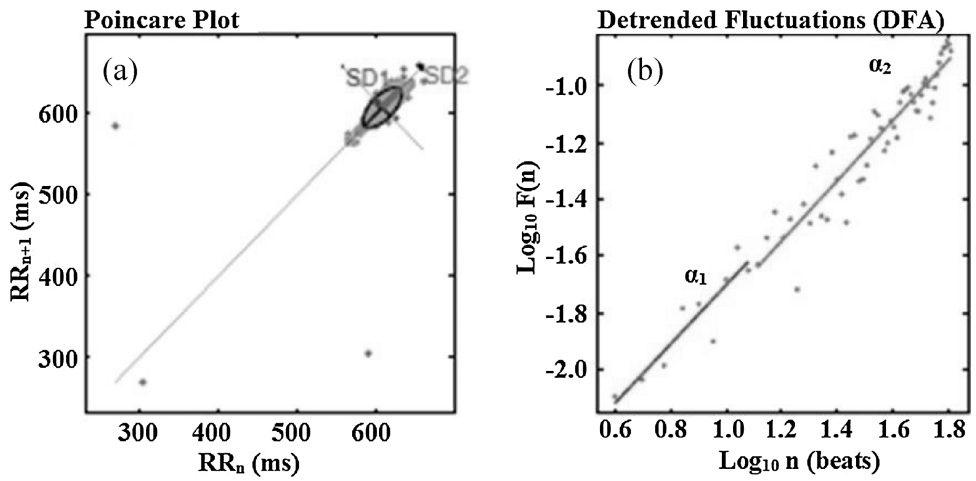

Fig. 7. (a) Poincare plot and (b) DFA for ECOG4 lung cancer.

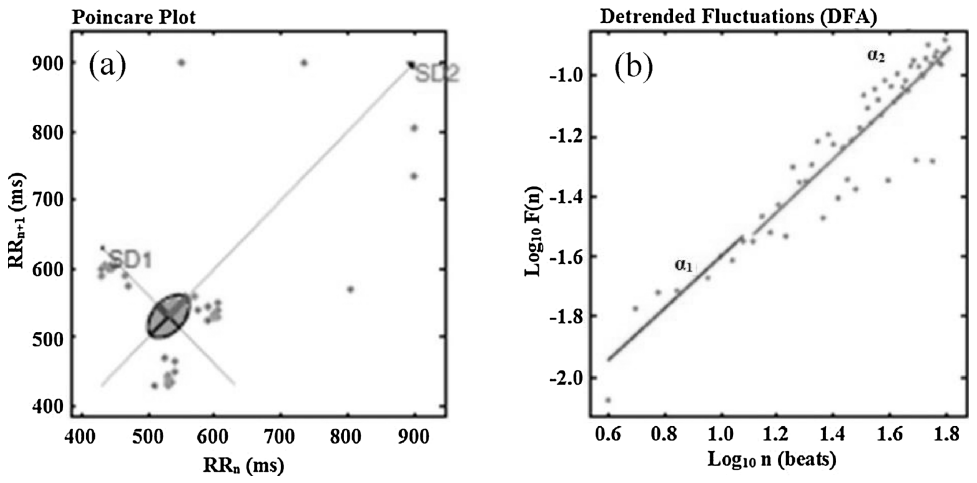

Fig. 8. (a) Poincare plot and (b) DFA for ECOG3 lung cancer.

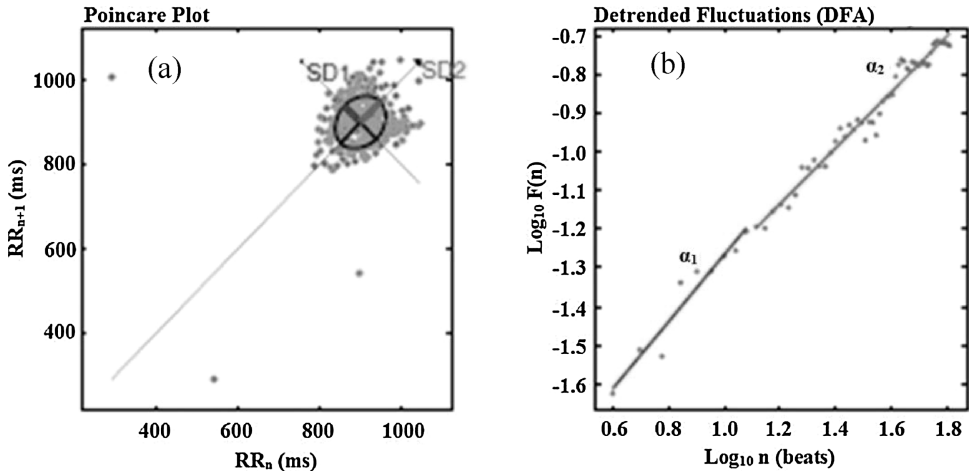

Fig. 9. (a) Poincare plot and (b) DFA for ECOG2 lung cancer.

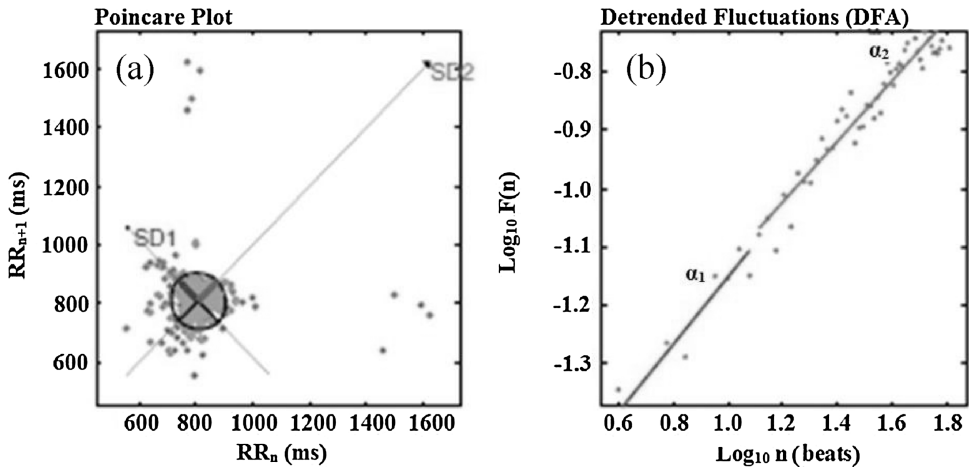

Fig. 10. (a) Poincare plot and (b) DFA for ECOG1 lung cancer.

significant between Control & ECOG3 ($p = 0.04365$) and ECOG2 & ECOG3 ($p = 8.99e\text{-}03$). SD_1 shows significance between ECOG2 & ECOG3 ($p = 0.02432$), ECOG2 & ECOG4 ($p = 4.64E\text{-}04$), ECOG1 & ECOG3 ($p = 0.001071$) and ECOG1 & ECOG4 ($p = 1.83e\text{-}04$). In SD_2, ECOG2 & ECOG3 ($p = 0.03307$), ECOG1 & ECOG2 ($p = 0.003214$), ECOG1 & ECOG3 ($p = 1.63e\text{-}04$), ECOG2 & ECOG4 ($p = 2.42e\text{-}04$) and Control & ECOG4 ($p = 3.66e\text{-}04$) was significant whereas between Control & ECOG3, ECOG1 & ECOG4 and ECOG3 & ECOG4 was found

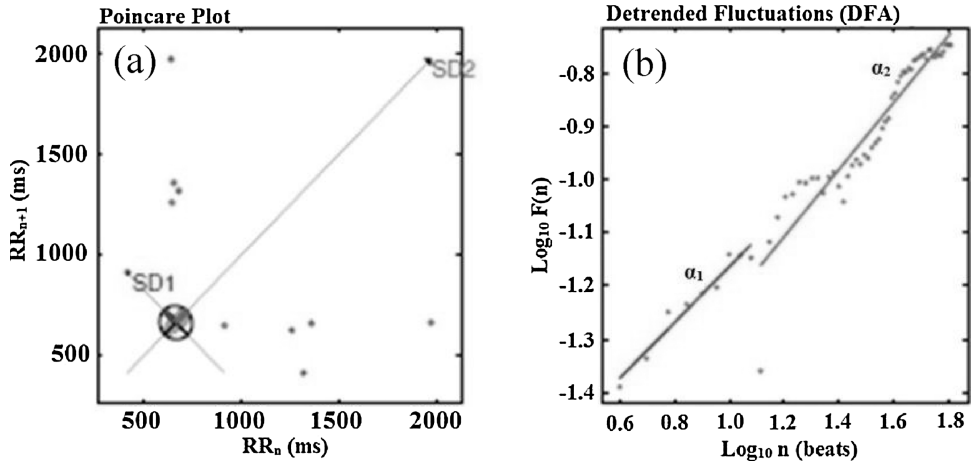

Fig. 11. (a) Poincare plot and (b) DFA for controls.

highly significant. SD_1/SD_2 between Control & ECOG1 ($p = 0.01708$), Control & ECOG2 ($p = 3.33e-04$) and ECOG1 & ECOG4 ($p = 0.01085$) was observed significant while, ECOG2 & ECOG4 was highly significant. L_{max} was found not significant. SampEn between Control & ECOG2 ($p = 0.05673$), ECOG4 & ECOG3 ($p = 1.85e-3$), ECOG4 & ECOG1 ($p = 1.03e-02$), ECOG4 & ECOG2 ($p = 1.17e-02$), Control & ECOG1 ($p = 3.23e-02$) and Control & ECOG3 ($p = 3.78e-04$) was observed to be significant. α_1 between Control & ECOG1 ($p = 0.03064$), Control & ECOG2 ($p = 0.01058$), Control & ECOG4 ($p = 5.23e-04$) and Control & ECOG3 ($p = 0.0149$) were significant. α_2 was significant between ECOG3 & ECOG4 ($p = 0.005535$), ECOG1 & ECOG3 ($p = 8.53e-04$), Control & ECOG2 ($p = 2.76e-02$), Control & ECOG4 ($p = 2.27e-03$) and ECOG2 & ECOG4 ($p = 1.15e-03$). However, high significance was observed between ECOG1 & ECOG4 and Control & ECOG1.

ANOVA test revealed significance of SD_1 ($p = 4.82E-07$, $F = 9.9406$), SD_2 ($p = 7.3E-09$, $F = 12.9$), α_1 ($p = 0.02089$, $F = 3.0047$), α_2 ($p = 4.41e-06$, $F = 8.4733$), ApEn ($p = 7.78e-11$, $F = 16.309$), SampEn ($p = 2.48e-08$, $F = 12.079$) and CD ($p = 1.87e-15$, $F = 25.56$).

Based on gender, in males SD_1 ($p = 0.00195$, $F = 4.797$), SD_2 ($p = 0.000357$, $F = 6.0455$), SD_1/SD_2 ($p = 0.01968$, $F = 3.1661$), DET ($p = 0.04581$, $F = 2.5807$) and α_2 ($p = 0.0002108$, $F = 6.444$) were significant. ApEn, SampEn and CD were highly significant. In females, ShanEn ($p = 0.02168$, $F = 3.0212$), α_2 ($p = 0.001428$, $F = 5.183$), ApEn ($p = 0.001246$, $F = 5.2882$) and SampEn ($p = 0.003713$, $F = 4.4507$) was significant. CD was highly significant between control male and other male subjects. CD also found to be highly significant between female subjects and ECOG3 and 4 PS of male group. SampEn observed to be significant between male and female subjects. SD_1 and SD_1/SD_2 was found significant between male and females of ECOG3 PS. The correlation between within ECOG PS states based on gender is given in Table 3.

Classification accuracy

The ANN was optimized for different learning rate and hidden layer nodes. The observed results suggested 0.4 learning rate to be best with accuracy of 88.8% and overall accuracy of 93.09% were obtained at 04 hidden layer nodes (Tables 4 and 5). ANN exhibited ECOG1 83.3%, ECOG2 50%, ECOG3 90%, ECOG4 95% and Controls 86.7%. The confusion matrix is given in Fig. 12 which states that 1 subject in ECOG1, 4 in ECOG2, 3 in ECOG3, 3 in ECOG4 and 4 in Controls have been misclassified. Moreover, with SVM, 100% accuracy has been observed. The accuracy observed with training data set was 97.4% and 97.9% for Regression and Confusion matrix, respectively. Further, 84.6% and 60% was obtained with test data set for Regression and Confusion matrix, respectively. The overall accuracy was 93.09% and 88.8% was found for Regression and Confusion matrix, respectively. Thus, with Regression analysis higher accuracy was achieved.

Discussion

The obtained results for different nonlinear HRV indices indicates reduced HRV with severity of lung cancer. Due to decrease in complexity pattern, HRV is reduced in lung cancer patients. The increase in sympathetic activity and decrease in parasympathetic activity as lung cancer progresses from ECOG1 to ECOG4 scale has been suggested from current findings. The reduced HRV signifies stress level that is common with cancer patients (Walsh and Nelson, 2002). However, fatigue may occur for those who fail to recover from stress-related metabolic demands. There exists an active role of ANS in cancer related fatigue (Thayer and Sternberg, 2006).

CD and PP (SD_1, SD_2 and SD_1/SD_2) analysis revealed reduced values with disease severity. However, ECOG1 of CD and SD_1/SD_2 deviate with lower value than ECOG2 and ECOG3. Further, SD_2 value was more as compared to SD_1. The lower CD was also observed in cardiomyopathy and stressed subjects than normal subjects (Carvajal et al., 2005; Schubert et al., 2009; Skinner et al., 1993) and less varied RR interval (Acharya et al., 2006). Past research work has also suggested lower reflexive control of heart beat interval with advance heart diseases and CD to be better predictor of risk in ventricular fibrillation (Skinner et al., 1993). Further, in contradiction, Fojt and Holcik (1998) reported higher CD value under pathology as compared to normal condition. The obtained decrease in SD_1, SD_2 and SD_1/SD_2 with the progression of lung cancer have been supported with published findings that suggested lower HRV with reduction in SD_1 and SD_2 values in polycystic ovarian syndrome (Saranya et al., 2015). Further, decrease in SD_1/SD_2 with low HRV was also reported (Fojt and Holcik, 1998). Thus, present results are in accordance with published literature. Further, the decreased SD_1 value indicates withdrawal of PNS activity, whereas SD_2 is a global activity indicator (Liu et al., 2010).

The rise in values of RP L_{max} and entropy (ApEn and SampEn) suggested reduced vagal tone with enhanced SNS activity with severity of lung cancer. Observed findings agreed with published results that suggested increase in L_{max} value under diabetic and vasovagal syncope condition with low HRV or impaired ANS

Table 1
Non-linear Heart Rate Variability measures based on gender (values in mean ± SE) where, 'M', 'F' and 'G' stands for male, female and gender respectively.

Non-linear measures of HRV	SD_1		SD_2		SD_2/SD_1		L_{mean}		L_{max}		REC	
G	M	F	M	F	M	F	M	F	M	F	M	F
Control	80.66 ± 32.69	104.06 ± 33.21	120.16 ± 42.12	131.76 ± 35.68	2.27 ± 0.35	1.46 ± 0.12	22.99 ± 7.01	24.46 ± 7.37	248.15 ± 39.31	182.15 ± 27.73	48.78 ± 5.95	55.72 ± 4.96
ECOG1	83.73 ± 8.67	70.72 ± 0	103.99 ± 6.65	87.40 ± 0	1.34 ± 0.23	1.24 ± 0	40.47 ± 18.97	16.14 ± 0	207.6 ± 56.75	113 ± 0	69.46 ± 6.89	58.86 ± 0
ECOG2	75.48 ± 19.95	58.89 ± 9.79	76.76 ± 7.47	62.75 ± 9.32	1.07 ± 0.18	1.17 ± 0.13	37.49 ± 27.35	31.78 ± 7.19	199.5 ± 38.5	210.17 ± 49.06	63.28 ± 21.74	61.9 ± 8.72
ECOG3	42.25 ± 3.39	26.32 ± 6.28	45.17 ± 1.65	49.52 ± 4.4	1.45 ± 0.3	2.89 ± 0.85	34.18 ± 5.05	52.89 ± 18.54	198.81 ± 20.45	311.5 ± 50.3	65.81 ± 3.53	64.79 ± 7.06
ECOG4	11.62 ± 1.67	11.45 ± 1.49	16.03 ± 1.74	17 ± 1.62	1.81 ± 0.2	1.88 ± 0.2	29.63 ± 4.89	24.77 ± 4.32	248.58 ± 23.11	256.03 ± 21.18	53.68 ± 4.8	47.36 ± 4.04

Non-linear measures of HRV	DET		ShanEn		α_1		α_2		ApEn		SampEn		CD	
G	M	F	M	F	M	F	M	F	M	F	M	F	M	F
Control	98.75 ± 0.37	99.1 ± 0.32	3.4 5 ± 0.15	3.49 ± 0.11	0.95 ± 0.09	0.78 ± 0.05	0.83 ± 0.05	0.76 ± 0.06	0.99 ± 0.07	0.89 ± 0.07	1.18 ± 0	1.02 ± 0	2.07 ± 0.44	2.21 ± 0.37
ECOG1	99.64 ± 0.19	99.73 ± 0	3.66 ± 0.15	3.34 ± 0	0.68 ± 0.09	0.65 ± 0	0.55 ± 0.04	0.45 ± 0	0.52 ± 0.19	0.78 ± 0	0.47 ± 0.14	0.67 ± 0.14	1.11 ± 0.78	0.48 ± 0
ECOG2	98.57 ± 1.42	99.29 ± 0.59	3.56 ± 0.45	3.56 ± 0.26	0.62 ± 0.25	0.67 ± 0.06	0.66 ± 0.06	0.56 ± 0.1	0.68 ± 0.46	0.57 ± 0.74	0.91 ± 0.18	0.56 ± 0.18	2.34 ± 2.23	0.84 ± 0.54
ECOG3	99.72 ± 0.08	99.71 ± 0.1	3.78 ± 0.9	4.15 ± 0.26	0.64 ± 0.09	0.73 ± 0.21	0.69 ± 0.06	1.02 ± 0.11	0.67 ± 0.06	0.79 ± 0.07	0.64 ± 0.18	0.86 ± 0.18	0.37 ± 0.09	0.46 ± 0.11
ECOG4	98.57 ± 0.36	98.43 ± 0.26	3.75 ± 0.13	3.63 ± 0.11	0.63 ± 0.07	0.64 ± 0.06	0.96 ± 0.04	0.96 ± 0.05	1.09 ± 0.04	1.06 ± 0.04	1.23 ± 0.11	1.41 ± 0.18	0.05 ± 0.02	0.12 ± 0.04

Table 2
Non-linear Heart Rate Variability measures based on age (value in mean ± SE).

Age (years)	PS	SD_1 (ms)	SD_2 (ms)	SD_1/SD_2	L_{mean} (beats)	L_{max} (beats)	REC (%)	DET (%)	ShanEn	α_1	α_2	ApEn	SampEn	CD
> = 65	ECOG 3	47.41 ± 3.33	47.53 ± 2.43	1.03 ± 0.07	31.43 ± 5.39	168.88 ± 18.96	65.18 ± 3.64	99.74 ± 0.08	3.77 ± 0.1	0.6 ± 0.09	0.67 ± 0.06	0.65 ± 0.05	0.6 ± 0.07	0.36 ± 0.09
	ECOG 4	11.59 ± 1.95	14.81 ± 1.93	0.81 ± 0.09	32.23 ± 6.62	244.89 ± 29.46	54.48 ± 6.04	98.47 ± 0.46	3.79 ± 0.17	0.57 ± 0.09	0.93 ± 0.05	1.08 ± 0.05	1.24 ± 0.09	0.04 ± 0.01
40–64	ECOG 3	23.64 ± 4.74	44.56 ± 2.08	0.55 ± 0.11	47.19 ± 12.52	298.22 ± 37.61	65.26 ± 5.29	99.76 ± 0.08	4.24 ± 0.2	0.75 ± 0.2	0.99 ± 0.11	0.71 ± 0.13	0.73 ± 0.16	0.43 ± 0.1
	ECOG 4	12.5 ± 1.48	18.36 ± 1.55	0.69 ± 0.06	25.78 ± 1.04	251.38 ± 19.55	49.89 ± 4.02	98.6 ± 0.24	3.66 ± 0.1	0.67 ± 0.06	0.93 ± 0.04	1.07 ± 0.04	1.38 ± 0.1	0.12 ± 0.04
0–40	ECOG 3	36.41 ± 7.35	43.18 ± 6.69	0.89 ± 0.21	59.41 ± 23.94	275 ± 63.57	73.07 ± 15.4	99.46 ± 0.36	3.86 ± 0.37	0.56 ± 0.29	0.62 ± 0.11	0.75 ± 0.26	0.76 ± 0.31	0.34 ± 0.33
	ECOG 4	5.73 ± 1.63	12.17 ± 4.66	0.51 ± 0.07	12.92 ± 2.38	256.25 ± 58.69	34.84 ± 5.68	97.64 ± 1.08	3.29 ± 0.25	0.68 ± 0.11	1.13 ± 0.05	1.1 ± 0.09	1.38 ± 0.31	0.05 ± 0.05

Table 3
Statistical analysis of nonlinear HRV parameters based on gender [Student's *t*-test and ANOVA ($p \leq 0.05$, $^*p \leq 0.0001$)].

T-TEST		Male					Female			
		Control	ECOG1	ECOG2	ECOG3	Subject	Control	ECOG1	ECOG2	ECOG3
Male	Subject	*CD				SampEn				
	ECOG1	*SD_2, REC, DET, SampEn, *α_1								
	ECOG3	SD_1/SD_2, SampEn, α_1, *CD	*SD_1, *SD_2			*CD				
	ECOG4	* SD_2, *SD_1/SD_2, α_1, *CD	*SD_1, *SD_2, *α_1	SD_1/SD_2, SampEn	*SD_1, *SD_2, REC, *DET, * SampEn, * α_1	*CD				
Female	Subject						L_{max}, SD_1, SD_2,	*CD		
	ECOG3				SD_1, SD_1/SD_2		SD_1, SD_2, ShanEn		SD_1, SD_1/SD_2, α_1	*CD
	ECOG4						L_{max}, SD_1, SampEn, α_1, *CD		SD_1, *SD_2, *SampEn, α_1	*SD_2, *DET, SampEn, CD

Table 4
Effects of learning rate on effectiveness of 9:100:5 network (no. of iterations = 1000). The Maximum accuracy is highlighted in bold.

Learning rate	% Accuracy
0.01	79.74
0.02	85.36
0.04	84.17
0.06	85.24
0.08	84.05
0.1	70.05
0.2	69.83
0.4	**88.80**
0.6	85.15
0.8	79.52
0.9	80.67

Table 5
Effects of number of hidden layer nodes on performance of ANN (testing results). The Maximum accuracy is highlighted in bold.

Hidden layer	% Accuracy
1	80.12
4	**93.09**
8	83.47
10	86.87
50	79.92
100	85.79
150	76.10
200	84.06
250	78.40
300	89.35
350	82.46
400	70.40
450	70.96
500	83.60

(Schlenker et al., 2014). Further, work of Mohebbi et al. (2011) revealed rise in L_{max} and Shannon entropy values in paroxysmal atrial fibrillation as compared to control. L_{max} has already been suggested to be inversely proportional to Largest Lyapunov exponent (LLE) in chaotic system (Schumacher, 2004). Further, Acharya et al. (2006) reported decreased LLE value for slow varying signals. Thus, it can be interpreted that L_{max} value increases for slowly varied signals. Literature review also suggested increase in ApEn under major depressive and panic disorder (Baumert et al., 2009). Further, the increase in entropy has also been discussed in syncopic subjects (Schlenker et al., 2014). The pattern of complexity gets lowered in depressed and schizophrenic patients in measuring ApEn and LLE in comparison with healthy controls indicating decreased cardiac vagal function and increased sympathetic activity (Bar et al., 2008; Yeragani et al., 2002). However, exception was observed in ECOG1 PS of entropy that showed larger value than ECOG3 values.

The value of α_1 and α_2 rises with the progression of disease from ECOG1 to ECOG4. However, an exception ECOG3 of α_1 was observed with lesser than ECOG2 value. Where, α_1 and α_2 represents the parasympathetic and sympathetic tone, respectively. The results suggested significant increase in SNS activity with higher ECOG PS state. The α_2 value also revealed to be higher than α_1 in ECOG3 and ECOG4 case. However, lower values were also observed in ECOG1 and ECOG2 cases where α_1 was higher than α_2. The slower RR variation can be suggested with larger α_2 value in higher ECOG state (highly critical) as compared to lower ECOG state (less critical). In support, literature demonstrated higher α_2 values than α_1 in individuals with cardiac anomalies (Roy and Ghatak, 2013) and suggested to be better indices of the physiological condition than α_1 (Yeh et al., 2010). α_2 indicates the collapse of nonlinear processes in regulation of cardiovascular system resulting in highly correlated fractal dynamics (De Souza et al., 2014). Literature suggested that age does not have any effect on HRV time-domain measures (Shukla and Aggarwal, 2017). In present study, it has been found that higher number of subjects were in the category of more than or equal to 65 years of age and no correlation was established between HRV and age. Also, varied age groups existed in each category of ECOG PS Scale. However, gender has been suggested to affect HRV indices with higher values of SDNN, RMSSD, Mean RR and RR TI in controls than subjects. Few studies mentioned that age and gender does not show any significant variation in HRV indices. Also, healthy subjects have higher HRV than the cancer patients. HRV reduces with increasing age (Gribbin et al., 1971).

The obtained results for ANN and SVM suggested the accuracy of 93.09% and 100% in predicting and classifying the lung cancer stage based on nonlinear HRV indices. The application of soft computing tools (ANN and SVM) for prediction analysis during complex decision-making can provide estimates of severity of disease status in patients. Literature review revealed application of ANN and SVM in prognosis of prostate and breast cancer (Cosma et al., 2016; Utomo et al., 2014). However, to the best author knowledge, none of the study has utilised HRV indices as input to machine learning. The accuracy may vary with other techniques of AI for offline and real-time applications. The real-time testing is yet to be done and part of our future work.

The decrease in complexity pattern with reduced HRV was observed in lung cancer. Further, the enhanced sympathetic tone with withdrawal of PNS activity has been suggested as the PS progresses from ECOG1 to ECOG4. In accordance to the obtained results, literature revealed lower HRV with suppressed PNS function and decreased survival rate among patients of various types and stages of cancer. However, the work was performed on time domain cancer HRV analysis (De Couck and Gidron, 2013; Guo et al., 2015; Kim et al., 2010; Shukla and Aggarwal, 2017). The observed findings having prognostic role as the modulation in vagal activity are of importance to researchers and clinicians. Implementation of AI on obtained nonlinear HRV indices can be more predictive in diagnosis of PS of lung cancer. However, the role of expert clinicians cannot be neglected for higher level diagnosis. Further, vagus nerve modulates the PNS activity in response to internal or external stressors including cancer (Walsh and Nelson, 2002) with enhanced SNS and suppressed PNS activity mediates the metastasis of cancer (Aggarwal et al., 2014). The ANS dysfunction activates or inhibits the vascular endothelial growth factor (VEGF) via release of norepinephrine, dopamine and bradykinin. With the advancement of cancer, norepinephrine level also increased, which induces VEGF expression that leads to angiogenesis (Walsh and Nelson, 2002). Thus, appropriate diagnosis of PS for appropriate treatment is highly required to reduce the mortality rate among cancer patients. With present findings, HRV can be one of such tools to be effective in prognosis.

Conclusion

HRV analysis has proven to be significant in several cardiac anomalies cases and of psychological origin. The HRV measures were influenced by gender but not by age in lung cancer subjects. The sympathetic activity was higher in males as compared to females in lung cancer which was vice-versa in healthy controls. The obtained nonlinear HRV results suggested lower HRV due to enhanced sympathetic and reduced parasympathetic activity with disease progression. Current work also suggested the role of HRV in prognosis of performance status of lung cancer with overall accuracy of 93.09% and 100% required to reduce the mortality rate and enhanced recovery. ANN could diagnose ECOG 4 very

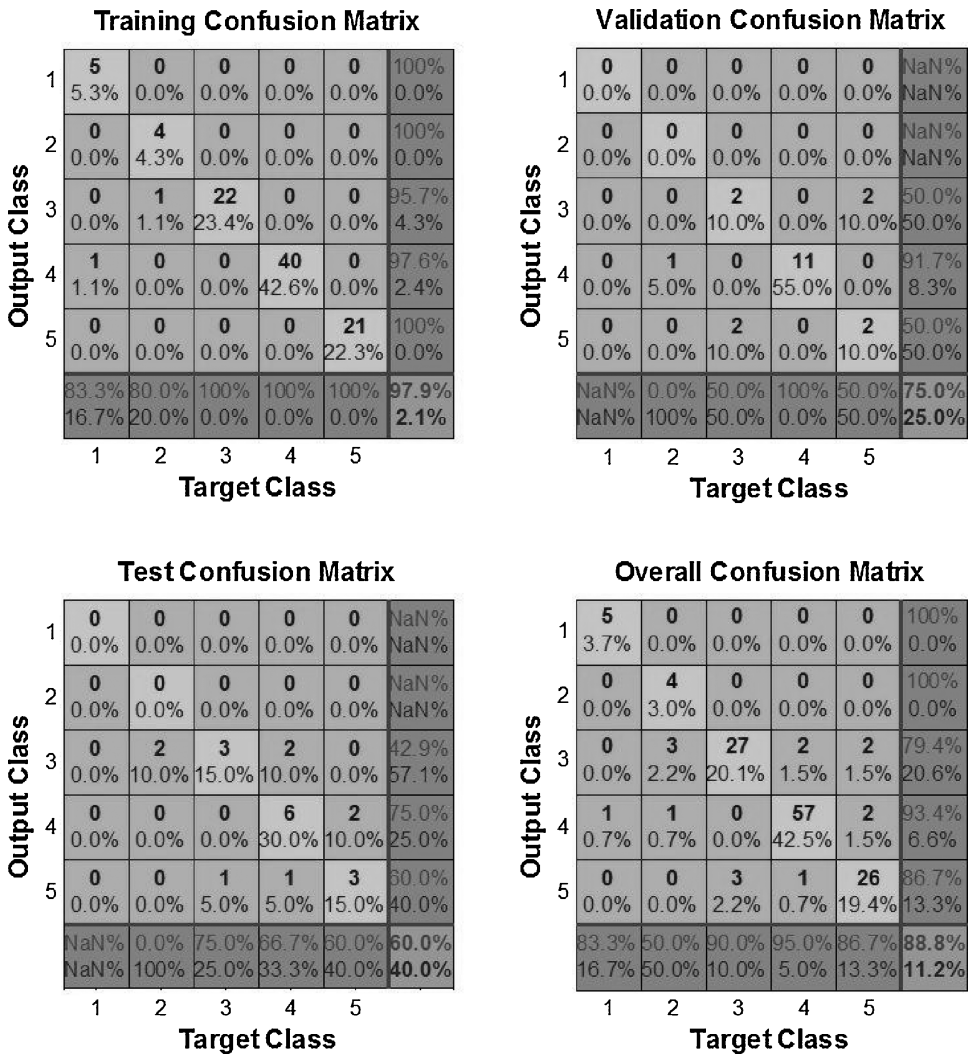

Fig. 12. The confusion matrix for 9:4:5 ANN architecture.

efficiently. In parallel, the proposed automated system may provide aid to expert clinicians for better and diagnosis of performance status of disease based on HRV features. Different patient's condition may exhibit variation in data. Thus, more varied number of cancer cases are required in future to optimize the system. However, expert clinician advice cannot be ignored.

Conflict of interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgement

Authors are grateful to Dr. Rajesh Singh, Dr. Seema, Dr. Richa Madhavi and Dr. Dinesh Sinha of Indira Gandhi Institute of Medical Sciences, Cancer Centre, Patna, India. Also, authors are thankful to medical oncologist Dr. Shreeniwas Raut, HMRI Paras Hospital, Patna for their clinical inputs and permitting for data collection in the hospital. Authors also extend gratitude to Prof. Rakesh Kumar Sinha (Department of Bio-Engineering, Birla Institute of Technology, Mesra, Ranchi) for their technical inputs in upgrading the manuscript.

References

Acharya, U.R., Lim, C.M., Joseph, P., 2002. Heart rate variability analysis using correlation dimension and detrended fluctuation analysis. ITBM-RBM 23, 333–339.

Acharya, U.R., Joseph, K.P., Kannathal, N., Lim, C.M., Suri, J.S., 2006. Heart rate variability: a review. Med. Biol. Eng. Comput. 44, 1031–1051.

Aggarwal, Y., Karan, B.M., Das, B.N., Aggarwal, T., Sinha, R.K., 2007. Backpropagation ANN-based prediction of exertional heat illness. J. Med. Syst. 31, 547–550.

Aggarwal, Y., Singh, N., Sinha, R.K., 2012. Electrooculogram based study to assess the effects of prolonged eye fixation on autonomic responses and its possible implication in man-machine interface. Health Technol. 2, 89–94.

Aggarwal, Y., Singh, N., Ghosh, S., Sinha, R.K., 2014. Eye gaze-induced mental stress alters the heart rate variability analysis. J. Clin. Eng. 39, 79–89.

Amari, S., Wu, S., 1999. Improving support vector machine classifiers by modifying kernel functions. Neural Netw. 12, 783–789.

Bar, K.J., Wernich, K., Boettger, S., Cordes, J., Boettger, M.K., Loffler, S., Agelink, M.W., 2008. Relationship between cardiovagal modulation and psychotic state in patients with paranoid schizophrenia. Psych. Res. 157, 255–257.

Baumert, M., Lambert, G.W., Dawood, T., Lambert, E.A., Esler, M.D., McGrane, M., et al., 2009. Short-term heart rate variability and cardiac norepinephrine spillover in patients with depression and panic disorder. Am. J. Physiolo. Heart Circ. Physiol. 297, H674–H679.

Carvajal, R., Wessel, N., Vallverdu, M., Caminal, P., Voss, A., 2005. Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy. Comput. Meth. Prog. Biomed. 78, 133–140.

Chiang, J.K., Kuo, T.B., Fu, C.H., Koo, M., 2013. Predicting 7-day survival using heart rate variability in hospice patients with non-lung cancers. J. Nat. Cancer Inst. 11, 812–813.

Cosma, G., Acampora, G., Brown, D., Rees, R.C., Khan, M., Pockley, A.G., 2016. Prediction of pathological stage in patients with prostate cancer: a neuro-fuzzy

model. PLoS One. 11, e0155856. doi:http://dx.doi.org/10.1371/journal.pone.0155856.
De Couck, M., Gidron, Y., 2013. Norms of vagal nerve activity, indexed by heart rate variability, in cancer patients. Cancer Epidemiology 37, 737–741.
De Couck, M., Brummelen, D.V., Schallier, D., Greve, J.D., Gidron, Y., 2013. The relationship between vagal nerve activity and clinical outcomes in prostate and non-small cell lung cancer patients. Oncol. Reports 30, 2435–2441.
De Couck, M., Marechal, R., Moorthamers, S., Van Laethem, J.-L., Gidron, Y., 2016. Vagal nerve activity predicts overall survival in metastatic pancreatic cancer, mediated by inflammation. Cancer Epidemiology. 40 (47), 51.
De Souza, A.C.A., Cisternas, J.R., De Abreu, L.C., Roque, A.L., Monteiro, C.B.M., Adami, F., et al., 2014. Fractal correlation property of heart rate variability in response to the postural change maneuver in healthy women. Int. Arch. Med. 7, 25.
Fojt, O., Holcik, J., 1998. Applying nonlinear dynamics to ECG signal processing. IEEE Eng. Med. Biol. Mag. 17, 96–101.
Garland, L.H., Beier, R.L., Coulson, W., Heald, J.H., Stein, R.L., 1968. The apparent sites of origin of carcinomas of the lung. Radiology 78, 1–11.
Gribbin, B., Pickering, T.G., Sleight, P., Peto, R., 1971. Effect of age and high blood pressure on barorefiex sensitivity in man. Circ. Res. 29, 424–431.
Guo, Y., Koshy, S., Hui, D., Palmer, J.L., Shin, K., Bozkurt, M., Yusuf, S.W., 2015. Prognostic value of heart rate variability in patients with cancer. J. Clin. Neurophysiol. 32, 516–520.
Kim, D.H., Kim, J.A., Choi, Y.S., Kim, S.H., Lee, J.Y., Kim, Y.E., 2010. Heart rate variability and length of in hospice cancer patients. J. Korean Med. Sci. 25, 1140–1145.
Li, S., Sun, Y., Gao, D., 2013. Role of the nervous system in cancer metastasis. Oncol. Letters. 5, 1101–1111.
Lilenbaum, R.C., Cashy, J., Hensing, T.A., Young, S., Cella, D., 2008. Prevalence of poor performance status in lung cancer patients: implications for research. J. Thorac. Oncol. 3, 125–129.
Liu, C., Liu, C., Shao, P., Li, L., Sun, X., Wang, X., Liu, F., 2010. Comparison of different threshold values *r* for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups. Physiol. Meas. 32, 167–180.
Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., Kurths, J., 2002. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys. Rev. E 66 (2 Pt 2), 026702.
Mohebbi, M., Ghassemian, H., Asl, B.M., 2011. Structure of the recurrence plot of heart rate variability signal as a tool for predicting the onset of paroxysmal atrial fibrillation. J. Med. Sig. Sens. 1, 113–121.
Roy, B., Ghatak, S., 2013. Nonlinear methods to assess changes in heart rate variability in type 2 diabetic patients. Arq. Bras. Cardiol. 101, 317–327.
Saranya, K., Pal, G.K., Habeebullah, S., Pal, P., 2015. Analysis of Poincare plot of heart rate variability in the assessment of autonomic dysfunction in patients with polycystic ovary syndrome. Int. J. Clin. Exp. Physiol. 2, 34.
Schlenker, J., Nedelka, T., Riedlbauchoven, L., Socha, V., Hana, K., Kutilek, P., 2014. Recurrence quantification analysis: a promising method for data evaluation in medicine. Eur. J. Biomed. Inform. 10, en35–en40.
Schubert, C., Lambertz, M., Nelesen, R.A., Bardwell, W., Choi, J.B., Dimsdale, J.E., 2009. Effects of stress on heart rate complexity −a comparison between short-term and chronic stress. Biol. Psychol. 80, 325–332.
Schumacher, A., 2004. Linear and nonlinear approaches to the analysis of R–R interval variability. Biol. Res. Nursing. 5, 211–221.
Shukla, R.S., Aggarwal, Y., 2017. Time-domain heart rate variability-based computer-aided prognosis of lung cancer. Ind. J. Cancer doi:http://dx.doi.org/10.4103/ijc.IJC_395_17 in press.
Signorini, M.G., Marchetti, F., Cerutti, S., 2001. Applying nonlinear noise reduction in the analysis of heart rate variability. IEEE Eng. Med. Biol. Mag. 20, 59–68.
Sinha, R.K., Aggarwal, Y., Das, B.N., 2007a. Backpropagation artificial neural network detects changes in electro-encephalogram power spectra of syncopic patients. J. Med. Syst. 31, 63–68.
Sinha, R.K., Aggarwal, Y., Das, B.N., 2007b. Backpropagation artificial neural network classifier to detect changes in heart sound due to mitral valve regurgitation. J. Med. Syst. 31, 205–209.
Skinner, J.E., Pratt, C.M., Vybiral, T., 1993. A reduction in the correlation dimension of heartbeat intervals precedes imminent ventricular fibrillation in human subjects. Am. Heart J. 125, 731–743.
Sorensen, J.B., Klee, M., Palshof, T., Hansen, H.H., 1993. Performance status assessment in cancer patients. An inter-observer variability study. Br. J. Cancer 67, 773–775.
Tarvainen, M.P., Niskanen, J.P., Lipponen, J.A., Ranta-Aho, P.O., Karjalainen, P.A., 2014. Kubios HRV −heart rate variability analysis software. Comput. Meth. Prog. Biomed. 113, 210–220.
Thayer, J., Sternberg, E., 2006. Beyond heart rate variability: vagal regulation of allostatic systems. Ann. New York Acad. Sci. 1088, 361–372.
Utomo, C.P., Kardiana, A., Yuliwulandari, R., 2014. Breast cancer diagnosis using artificial neural networks with extreme learning techniques. Int. J. Adv. Res. Artif. Intel. 3, 10–14.
Voss, A., Schroeder, R., Heitmann, A., Peters, A., Perz, S., 2015. Short-term heart rate variability-influences of gender and age in healthy subjects. PLoS One 10, e0118308. doi:http://dx.doi.org/10.1371/journal.pone.0118308.
Walsh, D., Nelson, K.A., 2002. Autonomic nervous system dysfunction in advanced cancer. Support. Care Cancer 10, 523–528.
Walsh, D., Donnelly, S., Rybicki, L., 2000. The symptoms of advanced cancer: relationship to age, gender, and performance status in 1,000 patients. Support. Care Cancer 8, 175–179.
Weinberg, R.A., 1996. E2F and cell proliferation: a world turned upside down. Cell 85, 457–459.
Yeh, R.-G., Shieh, J.-S., Han, Y.-Y., Wang, Y.-J., Tseng, S.-C., 2006. Detrended fluctuation analysis of short-term heart rate variability in surgical intensive care units. Biomed. Eng. Appl. Basis. Commun. 18, 21–26.
Yeh, R.-G., Chen, G.-Y., Shieh, J.-S., Kuo, C.-D., 2010. Parameters investigation of detrended fluctuation analysis for short-term human heart rate variability. J. Med. Biol. Eng. 30, 277–282.
Yeragani, V.K., Radhakrishna, R.K., Tancer, M., Uhde, T., 2002. Nonlinear measures of respiration: respiratory irregularity and increased chaos of respiration in patients with panic disorder. Neuropsychobiology 46, 111–120.