

Electrocardiogram-Derived Tidal Volume During Treadmill Stress Test

Javier Milagro, David Hernando, Jesús Lázaro, José A. Casajús, Nuria Garatachea, Eduardo Gil and Raquel Bailón

Abstract—Objective: Electrocardiogram (ECG) has been regarded as a source of respiratory information with the main focus in the estimation of the respiratory rate. Although little research concerning the estimation of tidal volume (TV) has been conducted, there are several ECG-derived features that have been related with TV in the literature, such as ECG-derived respiration, heart rate variability or respiratory rate. In this work, we exploited these features for estimating TV using a linear model. **Methods:** 25 young (33.4 ± 5.2 years) healthy male volunteers were recruited for performing a maximal (MaxT) and a submaximal (SubT) treadmill stress test, which were conducted in different days. Both tests were automatically segmented in stages attending to the heart rate. Afterwards, a subject-specific TV model was calibrated for each stage, employing features from MaxT, and the model was later used for estimating the TV in SubT. **Results:** During exercise, the different proposed approaches led to relative fitting errors lower than 14% in most of the cases and than 6% in some of them. **Conclusion:** Low achieved fitting errors suggest that TV can be estimated from ECG during a treadmill stress test. **Significance:** The results suggest that it is possible to estimate TV during exercise using only ECG-derived features.

Index Terms—tidal volume, electrocardiogram, respiration, exercise, heart rate variability

I. INTRODUCTION

MONITORING respiratory activity is very important in several applications, e.g., respiratory rate is a sensitive clinical parameter in a multitude of pulmonary diseases [1]. Another important respiratory parameter is the tidal volume

This work was supported by grant BES-2015-073694 from Ministerio de Economía y Competitividad (MINECO) and by Fondos FEDER through project RTI2018-097723-B-I00. Also by Government of Aragón and European Social Fund (EU) through Grupo de Referencia BSICoS (T39_17R) and by CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN) through Instituto de Salud Carlos III. This project has received funding from the European Unions Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 745755. The computation was performed by the ICTS NANBIOSIS, specifically by the High Performance Computing Unit of CIBER-BBN at University of Zaragoza. *Asterisk indicates corresponding author.*

J. Milagro*, D. Hernando, J. Lázaro, E. Gil and R. Bailón are with the Biomedical Signal Interpretation and Computational Simulation (BSI-CoS) group at the Aragón Institute of Engineering Research (I3A), IIS Aragón, University of Zaragoza, Zaragoza, Spain, and CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN), Madrid, Spain, (e-mail: milagro@unizar.es).

J. Lázaro is also with the Department of Biomedical Engineering at the University of Connecticut, Storrs CT, USA.

J. A. Casajús and N. Garatachea are with the Growth, Exercise, Nutrition and Development (GENUD) group at the Faculty of Health and Sport Sciences, IIS Aragón, University of Zaragoza, Zaragoza, Spain, and CIBER in Physiopathology of Obesity and Nutrition (CIBERObsn), Madrid, Spain. Copyright (c) 2017 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

(TV), which is defined as the volume of air inhaled or exhaled during a respiratory cycle. TV is useful for monitoring some respiratory diseases, such as Cheyne-Stokes respiration and sleep apnea. Both respiratory rate and TV have been studied as indicators of the anaerobic threshold [2], [3], which is related to the fitness level. The assessment of the fitness level results interesting for heart failure patients [4] and for athletes. Respiration monitoring techniques are usually based on plethysmography, pneumography or spirometry. These techniques require cumbersome devices which remain inconvenient for some applications, and which may interfere with natural breathing. Thus, some alternatives have been presented in the literature.

Although electrocardiogram (ECG) mainly represents the electrical cardiac activity measured on the surface of the skin, it is known to carry some respiratory information. Essentially, ECG is modulated by respiration through three different mechanisms: respiratory sinus arrhythmia (RSA), changes in the relative position of the recording electrodes, and changes in thoracic impedance. RSA is an extra-cardiac modulation of heart rate (HR) which reflects as a tachycardia during inspiration followed by a bradycardia during expiration. The origin of RSA is not completely understood, and three main non-exclusive hypotheses have been proposed, suggesting a central [5], baroreflex-mediated [6] or mechanical [7] origin, as well as a combination of them [8]. On the other hand, variations in the relative position of the recording electrodes due to chest movement during respiration result in alterations of the electrical pathway between them. Moreover, changes in the amount and composition of air inside the lungs or in the flow or composition of blood in the chest lead to different thoracic impedance value and distribution. These impedance changes directly affect electrical propagation, which is reflected as a modulation of the ECG morphology.

ECG morphology changes, as synchronized with respiration, have been exploited by several authors to extract respiratory information from ECG features such as R or R-to-S waves amplitude [9], the QRS-complex slopes [10] and QRS-complex area [11] variations, or vectocardiogram rotations [12]. This family of methods is usually referred to as ECG-derived respiration (EDR), since they allow to extract respiratory information only from ECG, without need of additional sensors. Despite the interest in EDR, research efforts have focused in estimating respiratory rate, with very few publications concerning TV estimation. Moreover and to the best of our knowledge, most of the studies aiming to estimate TV are based on ECG-unrelated techniques such as image acquisition [13], traqueal sounds [14] and inductive [15] or opto-electronic plethysmography [16].

Moody *et al.* already reported proportionality between TV and EDR [11]. Almost 30 years later, Sayadi *et al.* conducted a conceptual study aiming to derive TV using only ECG or different intra-cardiac signals in a controlled environment [17]. For this purpose, they employed mechanically ventilated swines, varying both TV and respiratory rate ranging from 0 to 750 ml and from 7 to 14 breaths/min (0.12 to 0.23 Hz) respectively. Each configuration was maintained for a minimum of 90 seconds, hence allowing stable measurement periods. In this work, we addressed TV estimation from ECG in rest and during a treadmill test, which constitutes a highly non-stationary scenario. Subject-specific models for TV based on ECG derived features were proposed and calibrated during a maximal effort test. These models were then validated for TV estimation in a submaximal treadmill test, conducted in a different day.

II. MATERIALS AND METHODS

A. Dataset

25 male volunteers aged 33.4 ± 5.2 years were recruited. All of them were apparently healthy and active, practicing aerobic training at least 3 times per week. None of the subjects were active smokers or reported any respiratory disorder by the time of the study, and only one of them was under medication intake (fluoxetine) during the recordings. They performed a maximal and a submaximal treadmill (Quasar MED LT h/p Cosmos, Nussdorf-Traunstein, Germany) test in different days, denoted as MaxT and SubT respectively. Both tests were divided in 3 stages: a 5-minute resting stage, during which the subjects remained sat and without talking, an exercise stage and a recovery stage. The resting stage was common to both tests, whereas different protocols were followed in the exercise stage. During MaxT, the volunteers started to run at a initial speed of $8 \text{ km}\cdot\text{h}^{-1}$ which was increased $1 \text{ km}\cdot\text{h}^{-1}$ per minute until they stopped due to volitional exhaustion. At this point, maximum HR was annotated for each subject. On the contrary, during SubT the procedure was similar to MaxT, but speed was kept constant once the subjects reached the 90% of their maximum HR, and they were asked to keep running for 2 more minutes at the reached speed. Finally, the recovery stage was shared by both tests and lasted between 3 and 5 minutes during which the subjects were required to remain running at a comfortable speed of $8 \text{ km}\cdot\text{h}^{-1}$.

Multi-lead ECG was acquired with a high-resolution holter (Mortara 48-hour H12+, Mortara Instrument, Milwaukee, Wisconsin). Leads I, II, III, aVL, aVR, aVF, V4, V5 and V6 were recorded at a sampling rate of 1000 Hz, whilst an Oxycon Pro device (Jaeger/Viasys, Germany) was used for recording breath-by-breath minute ventilation (V_E) and respiratory rate (F_r). A breath-by-breath HR signal was also acquired with the latter device. Recordings were performed at University of Zaragoza (Spain), and the protocol was approved by the institutional ethics committee following the ethical principles of the Declaration of Helsinki. Written informed consent was received from all the volunteers, and demographics are summarized in Table I.

TABLE I

DEMOGRAPHICS OF THE SUBJECTS IN THE PRESENTED DATASET. ALL THE VALUES ARE GIVEN AS MEAN \pm STANDARD DEVIATION, EXCEPT FROM THE NUMBER OF SUBJECTS (N) AND THE MAXIMUM HEART RATE (THE LATTER IS PROVIDED AS MEDIAN (25th PERCENTILE, 75th PERCENTILE) SINCE IT WAS NOT NORMALLY DISTRIBUTED). (BMI: BODY MASS INDEX).

N	Age (years)	Height (cm)	Weight (kg)	BMI (kg/m^2)	Max. HR (bpm)
25	33.4 ± 5.2	178 ± 5.5	74.8 ± 7.0	23.6 ± 2.1	180 (172, 186)

B. Preprocessing

Beat detection and delineation was accomplished in each ECG lead using the wavelet-based approach proposed in [18]. Instantaneous HR, $x_{\text{HR}}(n)$, was derived from beat time occurrence series as $60/d_{\text{IF}}^u(n)$, where $d_{\text{IF}}^u(n)$ represents the unevenly sampled interval function [19]. The reference for TV was obtained as $V_T(k) = V_E(k)/F_r(k)$, where index k represents the k -th breath, and $x_{\text{HR}}(n)$ was synchronized with $V_T(k)$. Synchronization was performed by employing the derivatives of $x_{\text{HR}}(n)$ and the HR signal provided by the Oxycon Pro device, synchronized with $V_T(k)$. First, both HR signals were resampled at 4 Hz by linear interpolation. Afterwards, correlation between the derivatives of the interpolated signals was calculated, and the time difference was obtained as the maximum of this correlation. Finally, the interpolated and synchronized version of $x_{\text{HR}}(n)$ was resampled at the time instants k by linear interpolation, and $x_{\text{HR}}(k)$ and $V_T(k)$ were smoothed using a 10-sample median filter.

The described preprocessing was applied in MaxT and SubT. In order to distinguish the notation between signals corresponding to each test, superindexes were employed so that $x_{\text{HR}}^m(k)$ and $V_T^m(k)$ refer to MaxT whereas $x_{\text{HR}}^s(k)$ and $V_T^s(k)$ allude to SubT.

Both stress tests were segmented into 5 different stages: a rest stage, three exercise stages, and a recovery stage. I_{rest} corresponds to the initial resting stage during which volunteers remained sat, and it lasts from the beginning of the recording until 30 seconds before exercise onset, so that transition from sat to the treadmill was discarded. I_{recov} refers to the recovery stage, and it expands from 30 seconds after maximum HR was reached until the end of the acquisition. The initial time offset of 30 seconds was included to avoid transition from exercise to recovery stage, since subjects did not behave in the same way after reaching maximum HR: whereas some of them remained running, some others jumped from the treadmill and then started running again. The segmentation of the three exercise stages was performed automatically from $x_{\text{HR}}^m(k)$ as percentages of the range of variations of the HR: 0-60% (I_{0-60}), 60-80% (I_{60-80}) and 80-100% (I_{80-100}), considering the mean HR at I_{rest} as 0% and the maximum HR as 100%. Whereas the definition of I_{rest} and I_{recov} was similar for MaxT and SubT, the percentages of maximum HR used for defining the other 3 stages were only calculated from $x_{\text{HR}}^m(k)$, and these values were employed in both tests. An example of this segmentation process is displayed in Fig. 1.

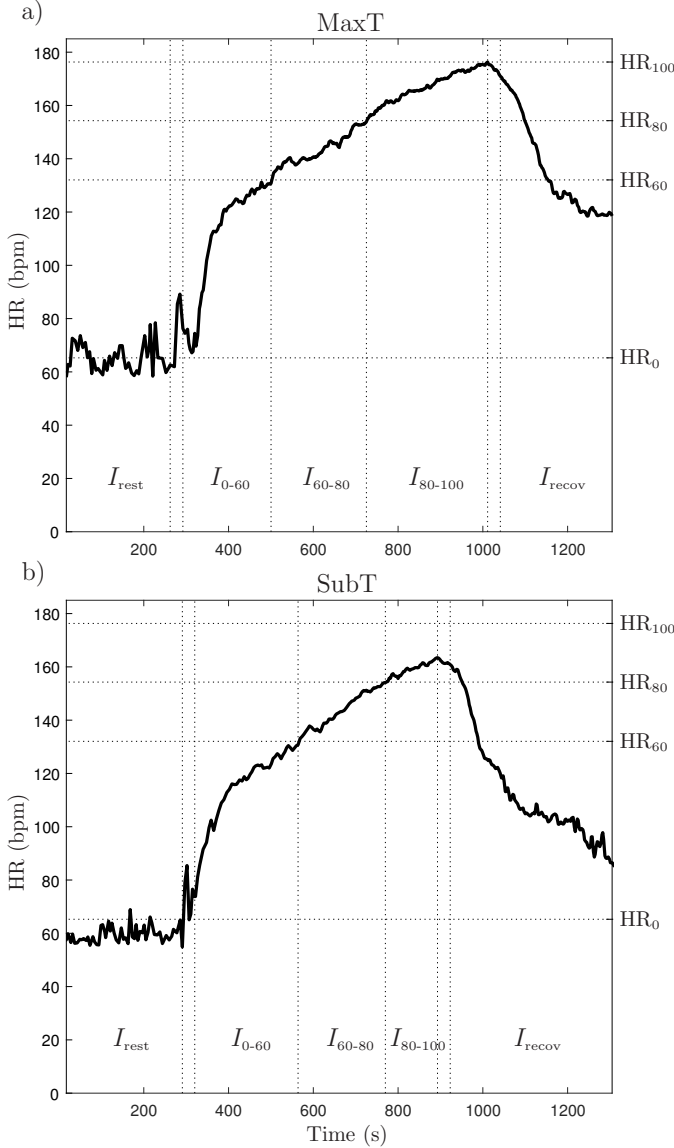


Fig. 1. Stage segmentation for MaxT, a), and SubT, b), of one subject is shown. Vertical lines indicate onset and offset of the different stages for each test, whereas horizontal lines mark the different percentages of maximum HR used for segmentation. In both cases, there is a 30-second interval between I_{rest} and I_{0-60} , and between I_{80-100} and I_{recov} , in order to exclude the transition from rest to exercise and from exercise to recovery respectively.

C. Tidal Volume Estimation

The proposed TV estimation approach consists of a linear model which was calibrated using the data in MaxT (training set) and evaluated using the data in SubT (test set). Both calibration and estimation were performed for each subject and stage. For the calibration process of each stage all the data samples in the current stage were used. Therefore, for stage I_i , the offset and slope (α_{I_i} and β_{I_i} respectively) of a linear model were estimated in a least squares sense by fitting $V_{T,I_i}^m(k)$ with a determined feature of MaxT, $\theta_{I_i}^m(k)$. The selection of appropriate features remains essential for a proper TV estimation, and several options are described below. Afterwards, TV was estimated in SubT as:

$$\hat{V}_{T,I_i}^s(k) = \alpha_{I_i} + \theta_{I_i}^s(k)\beta_{I_i}, \quad (1)$$

where $\theta_{I_i}^s$ represents the employed feature in stage I_i and in SubT.

Several features were tested as possible TV predictors: the amplitude of different EDR series (in a single-lead and a multi-lead approach), the instantaneous HR, the high-frequency (HF) power of the heart rate variability (HRV) signal and the respiratory rate. Moreover, a multi-parametric model including two or more of these features was also considered. The different methodologies followed for feature extraction are described below, and when not indicated the same procedure was applied for feature extraction in MaxT and SubT. It is important to note that all the features were normalized with respect to MaxT in order to minimize inter-day changes in measurements (different electrode position, different basal state, etc), so that:

$$\begin{aligned} \tilde{\theta}_{I_i}^m(k) &= \frac{\theta_{I_i}^m(k) - \mu_{\theta_{I_i}^m}}{\sigma_{\theta_{I_i}^m}}, \\ \tilde{\theta}_{I_i}^s(k) &= \frac{\theta_{I_i}^s(k) - \mu_{\theta_{I_i}^m}}{\sigma_{\theta_{I_i}^m}}, \end{aligned} \quad (2)$$

being $\mu_{\theta_{I_i}^m}$ and $\sigma_{\theta_{I_i}^m}$ the mean and standard deviation of $\theta_{I_i}^m(k)$ during interval I_i respectively. $\tilde{\theta}_{I_i}^m$ and $\tilde{\theta}_{I_i}^s$ represent the normalized versions of $\theta_{I_i}^m(k)$ and $\theta_{I_i}^s(k)$ during interval I_i respectively, although they will be referred to as $\theta^m(k)$ and $\theta^s(k)$ for simplicity.

D. Single-lead EDR

Amplitude difference between peaks and nadirs of the EDR series obtained from each lead were used as features for TV estimation. For this purpose, several EDR signals were considered: R-S amplitudes [9], QRS upslopes and downslopes, and R wave angles [10]. The resulting series were evenly sampled at 4 Hz and low-pass filtered at 1.5 Hz in order to discard HF components that are unrelated with respiration, and they were referred to as $x_{EDR}(t)$. As local maxima/minima in the EDR signal amplitude should be related with the end of expiration/inspiration, thus when electrodes are closer to/farther from the heart, peaks and nadirs in $x_{EDR}(t)$ were detected, and the difference between the amplitude of each peak and its corresponding nadir was calculated. The amplitude difference series were linearly interpolated at the time instants when expirations occur and smoothed with a 10-sample median filter, and the outputs of this process, $\theta^m(k)$, were used as features for estimating the TV. An example of this procedure using the R-S amplitude as EDR signal is displayed in Fig. 2.

E. Multi-lead EDR

During inspiration and expiration processes thorax expands or contract differently in the three spatial dimensions, and hence the use of spatial information could result in a better TV estimation. In this way, a multi-lead approach consisting in the combination of EDR signals extracted from different leads is

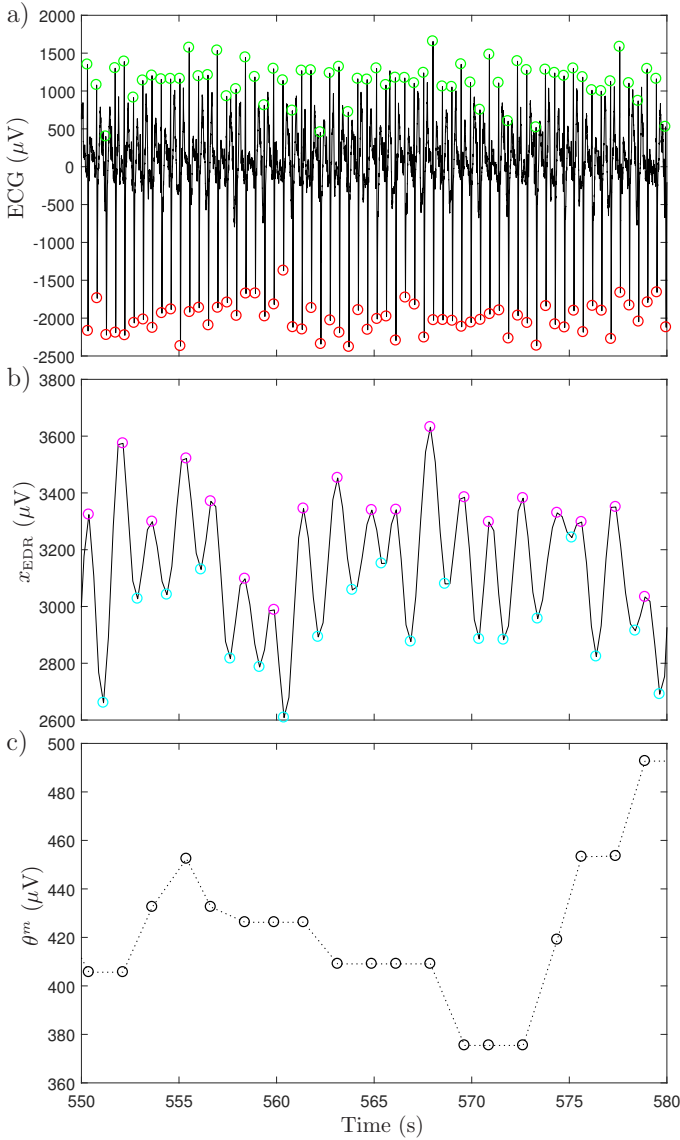


Fig. 2. Derivation of the amplitude difference series using R-S amplitude as EDR. In a), R (green circles) and S (red circles) waves were detected in the ECG signal, and the difference between them was calculated for each beat to obtain a R-S amplitude series. $x_{\text{EDR}}(t)$ is a low-pass filtered version of this series (b). Finally, the peaks and nadirs in $x_{\text{EDR}}(t)$ were detected (magenta and cyan circles respectively), and the series generated from the difference between them was resampled at the times when breaths occur and smoothed with a 10-sample median filter. The result, $\theta^m(k)$ (c), was used as a feature for our linear model.

proposed. The procedure is similar to the single-lead approach, and only differs in the definition of $x_{\text{EDR}}(t)$. In order to account for this three-dimensional behavior, in the multi-lead approach the EDR signals of three different leads were combined using a principal component analysis (PCA), and $x_{\text{EDR}}(t)$ was obtained as the first component of this PCA. The employed leads were selected as those forming the most orthogonal combination possible, so that spatial information was maximized. In this case, leads V4, V6 and aVF were selected. Also the possibility of combining all the available leads was contemplated in the analysis.

As in the single-lead approach, R-S amplitudes, QRS up-

slopes and downslopes, and R wave angles were used as EDR signals.

F. Instantaneous Heart Rate

When body metabolic demand increases, TV and HR increase in order to enlarge gas exchange. In this way, it is expectable that TV and HR present some correlation, so HR was considered as a possible TV estimator. $x_{\text{HR}}^m(k)$ and $x_{\text{HR}}^s(k)$ were used as features for the linear model in MaxT and SubT respectively:

$$\begin{aligned}\theta^m(k) &= x_{\text{HR}}^m(k), \\ \theta^s(k) &= x_{\text{HR}}^s(k).\end{aligned}\quad (3)$$

G. Heart Rate Variability

Relationship between heart rate variability (HRV) and TV has been a recurrent topic in the literature [20]–[22], where association of an increased TV with a higher HF power (P_{HF}) has been reported. For this reason, we considered P_{HF} as a potential feature for estimating TV. Although HRV is drastically reduced during moderate exercise [23], [24], the mechanical effect that breathing exerts over the sinus node appears to be responsible of increased P_{HF} during high intensity exercise [25]. However, P_{HF} calculation should be addressed carefully during physical activity, since the increased F_r during exercise could yield to a shift of power towards higher frequency components [26]. In this way, we adopted the methodology proposed in [27] for the calculation of P_{HF} , where it was determined in a time-frequency basis ($P_{\text{HF}}(t)$), using an adaptive band centered in F_r . Moreover, the presence of non-respiratory-related frequency components that might lay within the HF band [24] was taken into account (see [27] for details). The obtained $P_{\text{HF}}(t)$ was resampled at the time instants when breaths occur, and the resulting discrete series ($P_{\text{HF}}^m(k)$ and $P_{\text{HF}}^s(k)$ for MaxT and SubT respectively) were used as features for TV estimation:

$$\begin{aligned}\theta^m(k) &= P_{\text{HF}}^m(k), \\ \theta^s(k) &= P_{\text{HF}}^s(k).\end{aligned}\quad (4)$$

Prior to HRV analysis, ectopic beats and misdetections were identified and corrected using the method described in [28] (less than a 0.1% of the beats were labeled as ectopics or misdetections).

H. Respiratory Rate

Minute ventilation is defined as the volume of air inhaled or exhaled per minute. In this way, it is proportional to both TV and F_r . When an increase of gas exchange is required, the request can be satisfied by increasing either TV or/and F_r . However, in a very demanding situation such as a maximal effort test, the role of both magnitudes is closely related [29], so F_r was also considered as a possible TV estimator. Respiratory rate was estimated from the ECG as proposed by Lázaro *et al.* [10]. The EDR series calculated from the

upslopes, downslopes and angles of the R waves of all the available leads were employed. Same parameters than in [10] were used for F_r estimation in I_{rest} , whereas they were modified in the exercise stages. Concretely, the Welch's periodogram parameters were set to $T_s = 12$ s, $T_m = 4$ s and $t_s = 1$ s. Also the peakness threshold for averaging was reduced to 50% ($\xi = 0.5$), and faster changes in F_r were allowed by increasing δ from 0.1 to 0.2 (see [10] for details). Finally, the upper limit of the EDR signals filtering was set to 1.5 instead of 1 Hz in order to adapt it to the studied scenario, where F_r can exceed 60 breaths/minute.

Afterwards, the estimated F_r for MaxT and SubT, \hat{F}_r^m and \hat{F}_r^s , were used as features for TV estimation:

$$\begin{aligned}\theta^m(k) &= \hat{F}_r^m(k), \\ \theta^s(k) &= \hat{F}_r^s(k).\end{aligned}\quad (5)$$

I. Multi-parametric Model

Since in the literature TV has been related with all the previously described parameters independently, it might occur that combining the TV information obtained from different sources yields to a better estimation, so we considered the possibility of merging information from all the presented parameters using a multi-linear model, so that the final TV estimation was obtained as:

$$\hat{V}_{T_i}^s(k) = \alpha_{I_i} + \theta_{I_i}^{s,1}(k)\beta_{I_i}^1 + \theta_{I_i}^{s,2}(k)\beta_{I_i}^2 + \dots + \theta_{I_i}^{s,L}(k)\beta_{I_i}^L, \quad (6)$$

where $\theta_{I_i}^{s,l}(k)$ represents the different proposed features, i.e., parameters extracted from the single-lead and multi-lead EDR approaches, instantaneous HR, HRV and F_r . α_{I_i} and $\beta_{I_i}^l$ represent the parameters of the multi-linear model estimated from MaxT, and L is the number of parameters considered in the model.

J. Subject-independent Model

The possibility of having a single model which can be applied in a subject-independent basis was also addressed. For this purpose, the median of each of the coefficients of all the subject-specific multi-parametric models described above was calculated:

$$\begin{aligned}\bar{\alpha}_{I_i} &= \text{median}([\alpha_{I_i}(1), \alpha_{I_i}(2), \dots, \alpha_{I_i}(N)]), \\ \bar{\beta}_{I_i}^l &= \text{median}([\beta_{I_i}^l(1), \beta_{I_i}^l(2), \dots, \beta_{I_i}^l(N)]), \\ &\quad \forall l \in [1, \dots, L],\end{aligned}\quad (7)$$

being $\bar{\alpha}_{I_i}$ and $\bar{\beta}_{I_i}^l$ the coefficients of the subject-independent multi-parametric model, $\alpha_{I_i}(n)$ and $\beta_{I_i}^l(n)$ the coefficients of the multi-parametric model for subject n , and N and L the total number of subjects and parameters considered in the model respectively.

K. Performance Measurement

Median and interquartile range (IQR) of the absolute (ϵ_a) and relative error (ϵ_r) were calculated for each methodology, stage and subject in order to quantify the accuracy of the estimation. Medians of medians and IQRs were computed among subjects for each methodology and stage. Moreover, accuracy in the estimation of F_r was quantified as the inter-subject median of the median absolute error in each stage.

III. RESULTS

Median TV estimation errors obtained for each stage and approach are displayed in Tables II and III. Also graphical examples of the different methodologies are depicted in Figure 3. From the 25 volunteers recruited for this study, a total of 11 subjects had to be discarded due to missing TV signal or ECG, or bad quality signals either in MaxT or SubT.

The estimation errors obtained with the single-lead approach are summarized in Table II, where the results obtained for the different EDR methods can be compared. Although similar results were obtained for all the leads, lead II was the best performing independently of the considered EDR. The use of the downslopes of the R waves led to the lowest estimation errors in I_{rest} , I_{60-80} and I_{80-100} , with median relative errors of 11.68, 7.40 and 5.81%, respectively. On the other hand, the upslopes led to the best results in I_{0-60} , whereas the R-S amplitude was the best performing EDR in I_{recov} , with median relative errors of 17.01 and 14.07% respectively. Median IQRs of the estimation error were similar for all the approaches and stages.

Estimation errors obtained with the multi-lead, HR, HRV, F_r and multi-parametric approaches are reflected in Table III. Lowest relative fitting errors were obtained when combining the downslopes series in lead II and the instantaneous HR with the multi-parametric approach, except in I_{rest} , where the multi-lead approach remained the best option. Scatter plots of the performance of the multi-parametric approach for all the subjects and in the different stages are displayed in Fig. 4. Nevertheless, similar estimation errors were achieved in the multi-lead, HR, F_r and the multi-parametric approaches for all the stages, except for I_{rest} in F_r and the multi-parametric option, and I_{0-60} in F_r and multi-lead, where larger errors were observed. On the other hand, slightly increased errors in most of the stages were obtained for the HRV approach. In most of the cases, median relative error was lower than 14%, being it lower than 7.5% in I_{80-100} for all the approaches, and as low as 5.06% in the multi-parametric approach. Highest estimation errors were generally obtained for HRV and F_r , although their performance is comparable to the other approaches in the majority of the stages. Median IQRs of the estimation error were similar for all the approaches and stages, except for increased values in I_{0-60} for all the approaches excluding the HR and the multi-parametric options.

In Table IV, the median value of each coefficient obtained for the subject-specific multi-parametric models, as well as the ratio of the contributions of the downslopes and the HR are summarized. Whereas the contribution of the HR is in median always higher than the contribution of the EDR calculated

TABLE II

INTER-SUBJECT MEDIANS OF MEDIAN AND IQR OF THE FITTING ERRORS OBTAINED WITH THE SINGLE-LEAD EDR APPROACH. BEST RESULTS WERE ACHIEVED FOR LEAD II, SO RESULTS OBTAINED IN LEAD II WITH EACH OF THE CONSIDERED EDRS ARE DISPLAYED. THE MEDIAN AND IQR OF THE ABSOLUTE AND RELATIVE ERROR CORRESPONDING TO THE LOWEST RELATIVE ERROR IN EACH STAGE ARE HIGHLIGHTED IN BOTH TYPE.

		Π_{R-S}		Π_{upslope}		$\Pi_{\text{downslope}}$		Π_{angle}	
		Median	IQR	Median	IQR	Median	IQR	Median	IQR
I_{rest}	ϵ_a (liters)	0.11	0.11	0.10	0.11	0.11	0.12	0.09	0.09
	ϵ_r (%)	13.06	11.36	13.38	12.42	11.68	12.26	12.40	12.11
I_{0-60}	ϵ_a (liters)	0.37	0.33	0.27	0.31	0.39	0.30	0.31	0.36
	ϵ_r (%)	20.31	22.80	17.01	16.80	22.34	22.49	19.77	26.57
I_{60-80}	ϵ_a (liters)	0.27	0.23	0.24	0.23	0.18	0.23	0.29	0.21
	ϵ_r (%)	11.35	6.95	9.17	6.56	7.40	7.28	10.86	7.00
I_{80-100}	ϵ_a (liters)	0.19	0.11	0.19	0.13	0.16	0.14	0.17	0.12
	ϵ_r (%)	6.03	4.57	5.96	4.48	5.81	4.62	6.27	4.24
I_{recov}	ϵ_a (liters)	0.42	0.32	0.49	0.34	0.39	0.29	0.48	0.32
	ϵ_r (%)	14.07	12.76	15.26	11.43	15.77	11.32	16.68	12.30

TABLE III

INTER-SUBJECT MEDIANS OF MEDIAN AND IQR OF THE FITTING ERRORS OBTAINED WITH THE MULTI-LEAD, HR, HRV, F_r AND MULTI-PARAMETRIC APPROACHES. RESULTS CONCERNING THE MULTI-LEAD APPROACH WERE ACHIEVED CONSIDERING THE LEADS V4, V6 AND aVF, WHEREAS THOSE OF THE MULTI-PARAMETRIC APPROACH WERE OBTAINED FROM A COMBINATION OF THE SINGLE-LEAD AND HR APPROACHES (IN THE SINGLE-LEAD APPROACH, LEAD II AND QRS DOWNSLOPES WERE EMPLOYED). THE MEDIAN AND IQR OF THE ABSOLUTE AND RELATIVE ERROR CORRESPONDING TO THE LOWEST RELATIVE ERROR IN EACH STAGE ARE HIGHLIGHTED IN BOTH TYPE.

		Multi-lead		HR		HRV		F_r		Multi-parametric	
		Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
I_{rest}	ϵ_a (liters)	0.09	0.12	0.10	0.10	0.12	0.10	0.17	0.13	0.16	0.12
	ϵ_r (%)	11.87	14.23	12.74	12.57	16.19	14.10	17.37	16.85	17.61	15.87
I_{0-60}	ϵ_a (liters)	0.39	0.35	0.25	0.22	0.38	0.39	0.56	0.54	0.23	0.25
	ϵ_r (%)	21.86	22.31	15.12	14.36	20.41	25.38	28.85	40.43	12.96	15.87
I_{60-80}	ϵ_a (liters)	0.29	0.22	0.15	0.13	0.27	0.18	0.21	0.23	0.14	0.11
	ϵ_r (%)	11.24	6.80	7.64	5.66	10.98	7.02	9.02	7.49	7.41	4.68
I_{80-100}	ϵ_a (liters)	0.15	0.14	0.16	0.13	0.16	0.14	0.18	0.11	0.14	0.12
	ϵ_r (%)	6.97	5.02	6.14	4.13	5.32	4.03	7.14	4.81	5.06	4.01
I_{recov}	ϵ_a (liters)	0.37	0.31	0.27	0.28	0.54	0.34	0.36	0.35	0.28	0.28
	ϵ_r (%)	15.33	8.66	11.55	12.04	18.58	15.74	13.75	14.61	11.41	12.03

from the downslopes, the former turns highest during the I_{0-60} and the I_{recov} stages. On the other hand, and since best performance was generally achieved with the multi-parametric model, the subject-independent model was composed by the median of the coefficients of all the subject-specific multi-parametric models, which are those displayed in Table IV. The performance of this subject-independent model is also summarized in Table IV, where larger absolute and relative errors when compared with the subject-specific model were obtained in all the stages, except from I_{rest} .

Although not displayed in the tables, median F_r estimation error was computed for all the stages in MaxT and SubT independently. Estimation errors lower than 0.035 Hz were obtained for most of the stages in both MaxT and SubT, whilst a maximum error of 0.077 Hz was obtained for I_{0-60} in SubT.

IV. DISCUSSION

In this work, we addressed the possibility of estimating TV only from ECG recordings. The use of a dataset composed by signals acquired during maximal and submaximal treadmill stress tests results in a challenging but suitable environment for testing the behavior of the proposed methodologies in a highly non-stationary scenario. Also rest periods were considered, thus representing a more stationary situation. Although several different approaches were proposed, all of them are based on a first order lineal model, so that the complexity relies in the selection of adequate ECG-derived features.

Both MaxT and SubT were automatically divided in five stages, according to the exercise onset and offset and different percentages of the maximum reached HR. Afterwards, TV estimation was performed for each stage. For this purpose, we calibrated the parameters of the linear model in (1) using

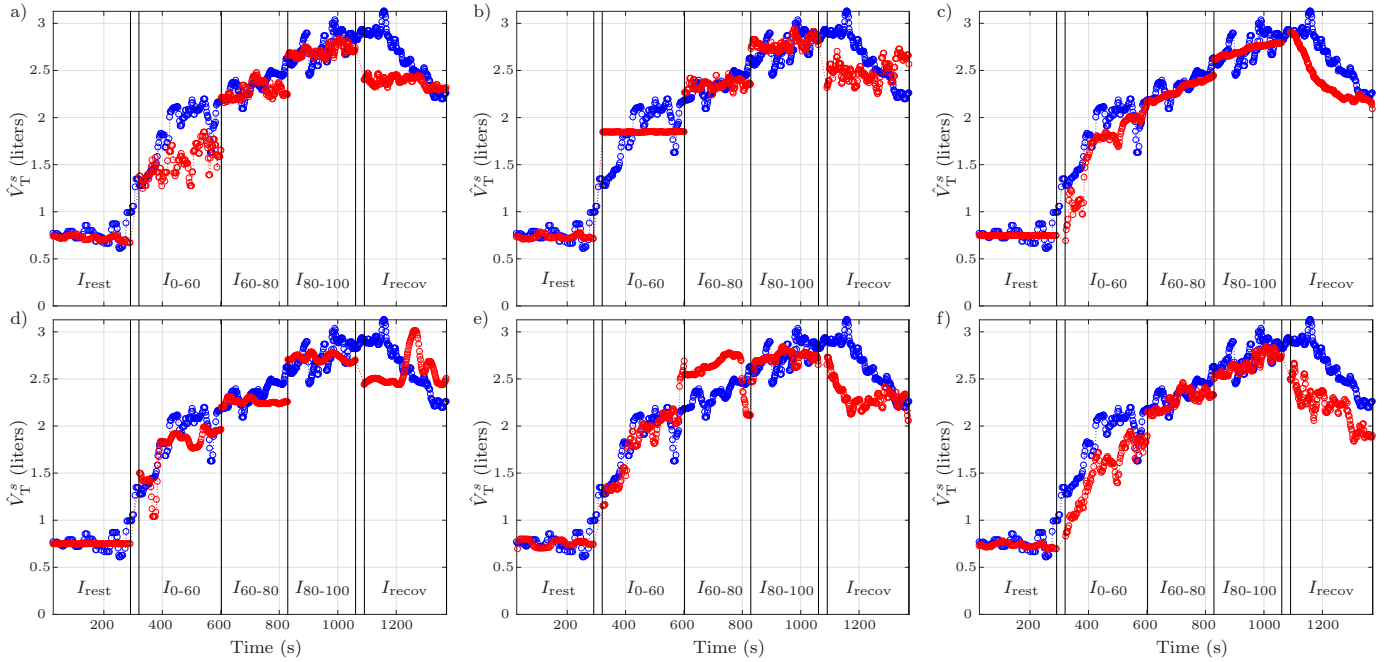


Fig. 3. Real (blue) and estimated (red) TV (\hat{V}_T^s) for a given subject is displayed. The different estimations were obtained from the best performing lead (II in this case) of the single-lead EDR approach (a), the multi-lead EDR approach (b), the instantaneous HR (c)) approach, the HRV approach (d), the respiratory rate approach (e) and the best performing feature combination (lead II in the EDR approach plus instantaneous HR in this case) of the multi-parametric approach (f).

all the samples in the stage of interest of MaxT, and used them for estimating the TV in the same stage of SubT. *Moody et al.* suggested that the amplitude of an EDR signal (calculated using the QRS-complex area) obtained from any set of electrodes should be roughly proportional to the TV [11]. In order to test this hypothesis, we applied four different EDR techniques to all the available leads: the R-S amplitudes [9], and the upslopes, downslopes and angles of the R waves [10]. Although results obtained for all the leads were similar, best estimation accuracy was obtained when using lead II, independently of the employed EDR. When comparing the different EDRs, lowest error was achieved with the downslopes in most of the stages, as displayed in Table II, which might suggest a more linear relationship with TV than the other EDRs. For all the leads and EDRs, a lower performance was observed in the I_{0-60} stage. The most likely explanation for the performance reduction is the fast changes that occur during this stage, with a sudden increase in TV, HR and F, that are not completely followed by the EDRs or the first-order linear model. Also changes in the position of the electrodes used for ECG acquisition (since both tests were performed in different days, it is probable that they were attached in slightly different places) or in the basal state of the subjects could constitute additional sources of estimation error.

In a second approach, we addressed the possibility of including spatial information by a combination of EDRs extracted from three "quasi-orthogonal" leads (as lead V2 was missing, there were not three leads that were completely orthogonal) through PCA, so that the main variations in those EDRs, which are expected to be produced by respiration, were maximized. Also the option of combining all the avail-

able leads was considered, although it did not outperformed the three-lead option. Despite the similar results obtained in the single-lead and multi-lead EDR approaches, median relative error was slightly higher when accounting for spatial information, except for I_{rest} and I_{recov} , probably due to fact that sources of noise such as movement during running may contribute to the first component of PCA. Nonetheless, since the different spatial dimensions of thoracic expansion do not contribute equally to the total TV, most of the information might be obtained from a single lead capturing variations along a preferable dimension. Additionally, the non-standard leads employed by *Lázaro et al.* in [10] were considered, although their performance was not higher than that of the multi-lead approach.

HRV has also been related with TV in the literature, as P_{HF} has been reported to be affected both by TV and respiratory rate [20]–[22]. In a previous study using the same dataset, *Hernando et al.* proposed a method for calculating P_{HF} in a time-frequency basis and considering the presence of non-respiratory-related components [27], so we adopted this methodology and used the obtained P_{HF} as a feature for estimating the TV. Results displayed in Table III revealed that the performance using HRV was similar than for the single-lead EDR approach in I_{80-100} and in I_{0-60} , but fitting error was higher in I_{rest} , I_{60-80} and I_{recov} (although it increased less than a 5% in all the cases). The lowered performance during I_{60-80} and I_{recov} could be related with the fact that HRV is drastically reduced during moderate exercise [23], [24] and so its variations might uncouple from those in TV, thus resulting in an increased estimation error in these stages, whereas during high intensity exercise (as in I_{80-100}) P_{HF} might recover the

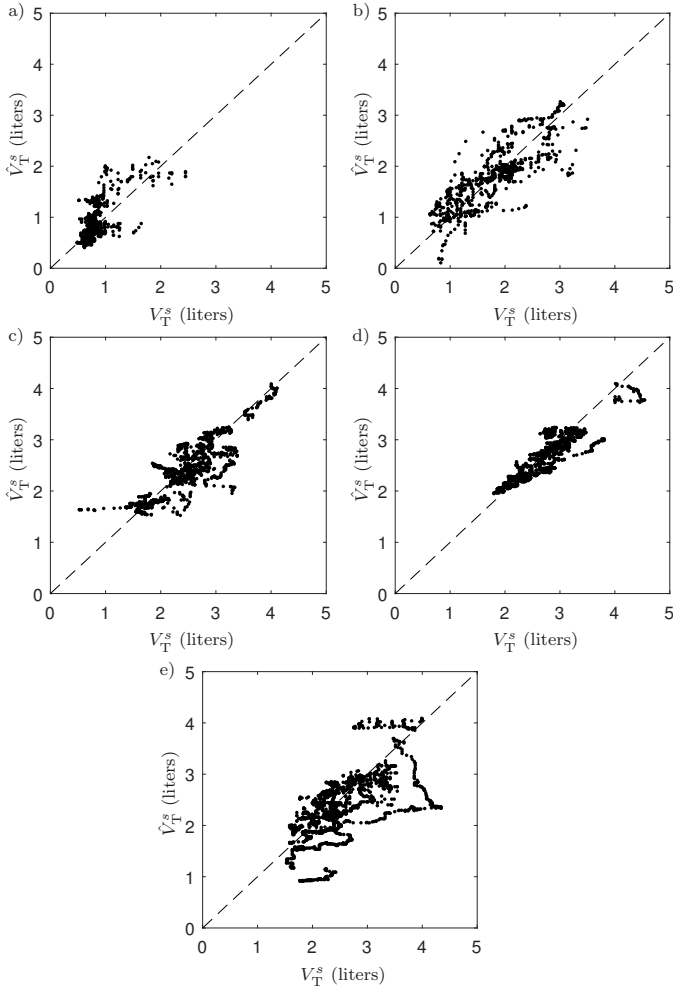


Fig. 4. Scatter plots of the TV estimated from the multi-parametric approach when combining the downslopes of lead II and the HR (\hat{V}_T^s) against the real one (V_T^s) for all the subjects and each of the stages (a): I_{rest} , b): I_{0-60} , c): I_{60-80} , d): I_{80-100} and e): I_{recov} . Dashed lines indicate $\hat{V}_T^s = V_T^s$.

coupling with TV (possibly due to mechanical stretching of the sinus node [25]), thus resulting in a reduced estimation error. In the case of I_{rest} , decreased performance could have its origin in the fact that low respiratory rates might result in wrong estimations of P_{HF} , due to spectral shift towards low frequency band. Differences in the metabolic demands and in the autonomic nervous system state (reflected in HRV) from day to day could also contribute to estimation error.

Since metabolic demands of the body increase during exercise, the demanded gas exchange increases as well, so both minute ventilation and HR increase consequently. Due to the very close relation between TV and F_r in the control of minute ventilation during exercise [29], F_r was considered for estimating the TV. First, F_r was estimated as proposed by *Lázaro et al.* [10]. Essentially, F_r is calculated from the EDRs derived from the upslopes, downslopes and angles of the R waves of all the available leads, which are combined in a short-term basis when they spectra are peaky enough. Although the same parameters employed in [10] were employed in I_{rest} , they were modified during the exercise stages attending to the fast and wide changes observed in F_r in the studied scenario.

Median absolute F_r estimation error was generally lower than 0.035 Hz. Results in Table III reveal a relative error lower than 10% in I_{60-80} and I_{80-100} , lower than 18% in I_{rest} and I_{recov} and higher than 28% in I_{0-60} respectively. The decreased accuracy in the latter stage might be explained by the different response of the subjects to the increasing ventilation demands, as observed by *Carey et al.* during incremental exercise [30]. Whereas in some subjects F_r increased in parallel with the exercise load, some others satisfied the ventilation demands during moderate exercise by mainly regulating the TV, with a lower contribution of F_r , which increased slowly as exercise became more intense. In order to study the effect of F_r estimation in the results, we repeated the calculations using the original F_r provided by the Oxycon Pro device, concluding that the error in F_r estimation did not contribute noticeably to the error in TV estimation.

On the other hand, HR increases together with exercise intensity. In this way, we observed that the use of HR as TV estimator resulted in low fitting errors, with a median relative error lower than 13% in all the stages except in I_{0-60} , where it increased up to 15.12%. Despite the high performance of this approach (especially in I_{0-60} , when compared to the other approaches), these results should be regarded carefully, since the relation between HR and TV could be not that direct in other scenarios.

Finally, we also considered the combination of several features using a data fusion algorithm such as a multi-linear model. From all the possible feature combinations, lowest fitting errors were obtained when combining the single-lead EDR and the HR approaches. This feature combination outperformed all the other approaches in all the stages, except in I_{rest} , thus highlighting the multi-source origin of the physiological mechanisms underlying the respiratory-related modulation of cardiac activity. In this way, although the EDR and HR signals may share some information regarding respiratory activity, they also contain non-redundant information that is exploited by this multi-parametric approach, thus resulting in a better TV model than any of the considered features separately. However, and as displayed in Table IV, the contribution of the HR was always dominant, independently on the considered stage. This dominance was weak during I_{rest} , but increased in the other stages, especially during I_{0-60} and I_{recov} . This behavior is most likely related with the similar profile of the instantaneous HR and the TV during the treadmill tests, so that when abrupt transitions occurs (such as those in I_{0-60} and I_{recov}), HR turns in the best estimator of the TV, whose changes are poorly followed by the other considered features. As displayed in Fig. 4, the performance of this approach was similar for all the subjects, with larger variations occurring in I_{rest} , I_{0-60} and I_{recov} . Whereas the presence of outliers in I_{0-60} was to be expected due to abrupt changes in TV, larger performance variations during I_{rest} are most likely due to a lower linear coupling between the target features and the TV during spontaneous breathing in some of the subjects. On the other hand, the subjects behaved distinctly after reaching the maximum HR in MaxT (some of them jumped from the treadmill and started running again), and therefore it is difficult to establish whether large deviations in I_{recov} are explained by this fact or by an uncoupling between TV and the estimation features.

TABLE IV

MEDIAN (IQR) OF THE PARAMETERS OF THE MULTI-PARAMETRIC MODEL FOR EACH STAGE (PARAMETERS OF THE MULTI-LINEAR MODEL WHEN THE QRS DOWNSLOPES IN LEAD II AND THE INSTANTANEOUS HR ARE CONSIDERED, SO THAT $\bar{\alpha}_{I_i}$ IS THE OFFSET, AND $\bar{\beta}_{I_i}^1$ AND $\bar{\beta}_{I_i}^2$ ARE THE CONTRIBUTIONS OF THE DOWNSLOPES AND THE HR RESPECTIVELY). ALSO THE MEDIAN (IQR) OF THE ABSOLUTE AND RELATIVE ERRORS OBTAINED WHEN ESTIMATING THE TV USING THE MEDIAN MODEL ARE DISPLAYED.

	$\bar{\alpha}_{I_i}$	$\bar{\beta}_{I_i}^1$	$\bar{\beta}_{I_i}^2$	$ \bar{\beta}_{I_i}^2/\bar{\beta}_{I_i}^1 $	ϵ_a (liters)	ϵ_r (%)
I_{rest}	0.800 (0.232)	0.021 (0.088)	0.036 (0.101)	1.128 (1.594)	0.10 (0.08)	14.04 (11.77)
I_{0-60}	1.734 (0.488)	0.016 (0.070)	0.358 (0.352)	6.186 (21.097)	0.48 (0.28)	22.72 (22.19)
I_{60-80}	2.410 (0.561)	-0.020 (0.042)	0.110 (0.084)	3.076 (4.743)	0.45 (0.14)	18.74 (4.88)
I_{80-100}	2.828 (0.728)	0.001 (0.054)	0.089 (0.1849)	5.791 (14.302)	0.32 (0.14)	10.23 (5.47)
I_{recov}	2.561 (0.942)	0.001 (0.094)	0.230 (0.257)	7.955 (52.689)	0.49 (0.39)	19.75 (12.33)

Nevertheless, since the multi-parametric approach was generally the best performing, it was used for estimating a subject-independent model, built as the median of all the previously trained subject-specific models. As summarized in Table IV, relative errors lower than 20% were obtained for most of the stages (relative error raised to 23% in I_{0-60}). As expected, the median absolute and relative errors were larger than in the case of the subject-specific model for all the stages, except in the case of I_{rest} . This is possibly related with the fact that, for those subjects presenting large estimation errors in this particular stage, a median model results in a better approach, given that the average TV during rest is similar for people with similar characteristics.

There are also some limitations that must be highlighted. First, the dataset was composed only by healthy men in a relatively small age range, which were used to aerobic training. In this way, the algorithm performance in subjects with a different age range or physical condition, or with cardiorespiratory disorders remains unknown and should be evaluated in further studies. Regarding the high linear relation between TV and HR, the scope of this work was limited to a treadmill test, and hence this coupling might be reduced in other scenarios.

In summary, we proposed a simple method for estimating TV from only the ECG. Several different features that are related with respiration were considered as TV estimators, and all the methodologies were tested in rest and also in a highly non-stationary scenario such as an effort treadmill test. The promising results with low fitting errors suggest that it might be possible to develop a subject-specific model that could be applied to estimate the TV in a day-independent basis. Nevertheless, further research should be conducted in order to improve TV estimation from ECG. In this study we only considered estimation during an exercise test, but this method could be useful in many other applications, e.g., in the monitoring of respiratory disorders such as Cheyne-Stokes respiration, chronic obstructive pulmonary disease or asthma. For this purpose, validation in these scenarios remains crucial. Moreover, the proposed model could be regarded as an interesting tool in several activities that are centered in the control of respiration, like meditation, yoga or mindfulness, and in different fields of sports science.

V. CONCLUSION

A methodology for estimating TV from several features derived from ECG during a treadmill stress test has been presented, considering the possibility to develop a subject-oriented model independent on the measurement day. Recordings from two different days were employed, being the first used for calibrating the model and the second for testing. During exercise, the different proposed approaches led to fitting errors lower than 14% in most of the cases and than 6% in some of them, suggesting that TV can be estimated from ECG in non-stationary conditions. Best results were obtained when combining the information provided by a single-lead EDR signal based on the downslopes of the R waves and the instantaneous HR.

REFERENCES

- [1] B. Krieger *et al.*, "Continuous noninvasive monitoring of respiratory rate in critically ill patients," *Chest*, vol. 90, no. 5, pp. 632–634, 1986.
- [2] D. G. Carey *et al.*, "Respiratory rate is a valid and reliable marker for the anaerobic threshold: implications for measuring change in fitness," *J Sports Sci Med*, vol. 4, no. 4, p. 482, 2005.
- [3] L. Mirmohamadsadeghi *et al.*, "The respiration pattern as an indicator of the anaerobic threshold," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 546–549, IEEE, 2015.
- [4] K. Yoshimura *et al.*, "Dynamics of cardiorespiratory response during and after the six-minute walk test in patients with heart failure," *Physiother Theory Pract*, pp. 1–12, 2018.
- [5] D. L. Eckberg, "Point: counterpoint: respiratory sinus arrhythmia is due to a central mechanism vs. respiratory sinus arrhythmia is due to the baroreflex mechanism," *J Appl Physiol*, vol. 106, no. 5, pp. 1740–1742, 2009.
- [6] J. M. Karemaker, "Counterpoint: respiratory sinus arrhythmia is due to the baroreflex mechanism," *J Appl Physiol*, vol. 106, no. 5, pp. 1742–1743, 2009.
- [7] G. Blain, O. Meste, and S. Bermon, "Influences of breathing patterns on respiratory sinus arrhythmia in humans during exercise," *Am J Physiol Heart Circ Physiol*, vol. 288, no. 2, pp. H887–H895, 2005.
- [8] C. Julien *et al.*, "Comments on point: counterpoint: respiratory sinus arrhythmia is due to a central mechanism vs. respiratory sinus arrhythmia is due to the baroreflex mechanism," *J Appl Physiol*, vol. 106, no. 5, pp. 1745–1749, 2009.
- [9] C. Mason and L. Tarassenko, "Quantitative assessment of respiratory derivation algorithms," in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 2, pp. 1998–2001, IEEE, 2001.
- [10] J. Lázaro *et al.*, "Electrocardiogram derived respiratory rate from QRS slopes and R-wave angle," *Ann Biomed Eng*, vol. 42, no. 10, pp. 2072–2083, 2014.
- [11] G. B. Moody *et al.*, "Derivation of respiratory signals from multi-lead ECGs," *Computers in cardiology*, vol. 12, no. 1985, pp. 113–116, 1985.

- [12] R. Bailón, L. Sornmo, and P. Laguna, "A robust method for ECG-based estimation of the respiratory frequency during stress testing," *IEEE Trans Biomed Eng*, vol. 53, no. 7, pp. 1273–1285, 2006.
- [13] B. A. Reyes *et al.*, "Tidal volume and instantaneous respiration rate estimation using a volumetric surrogate signal acquired via a smartphone camera," *IEEE J Biomed Health Inform*, vol. 21, no. 3, pp. 764–777, 2017.
- [14] N. Reljin, B. A. Reyes, and K. H. Chon, "Tidal volume estimation using the blanket fractal dimension of the tracheal sounds acquired by smartphone," *Sensors*, vol. 15, no. 5, pp. 9773–9790, 2015.
- [15] V.-P. Seppa, J. Viik, and J. Hyttinen, "Assessment of pulmonary flow using impedance pneumography," *IEEE Trans Biomed Eng*, vol. 57, no. 9, pp. 2277–2285, 2010.
- [16] C. M. A. Reinaux *et al.*, "Tidal volume measurements in infants: Opto-electronic plethysmography versus pneumotachograph," *Pediatr Pulmonol*, vol. 51, no. 8, pp. 850–857, 2016.
- [17] O. Sayadi *et al.*, "An optimized method for estimating the tidal volume from intracardiac or body surface electrocardiographic signals: implications for estimating minute ventilation," *Am J Physiol Heart Circ Physiol*, vol. 307, no. 3, pp. H426–H436, 2014.
- [18] J. P. Martínez *et al.*, "A wavelet-based ecg delineator: evaluation on standard databases," *IEEE Trans Biomed Eng*, vol. 51, no. 4, pp. 570–581, 2004.
- [19] O. Rompelman, A. Coenen, and R. Kitney, "Measurement of heart-rate variability: part 1 - comparative study of heart-rate variability analysis methods," *Med Biol Eng Comput*, vol. 15, no. 3, p. 233, 1977.
- [20] J. A. Hirsch and B. Bishop, "Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate," *Am J Physiol Heart Circ Physiol*, vol. 241, no. 4, pp. H620–H629, 1981.
- [21] D. L. Eckberg, "Human sinus arrhythmia as an index of vagal cardiac outflow," *J Appl Physiol*, vol. 54, no. 4, pp. 961–966, 1983.
- [22] T. E. Brown *et al.*, "Important influence of respiration on human RR interval power spectra is largely ignored," *J Appl Physiol*, vol. 75, no. 5, pp. 2310–2317, 1993.
- [23] F. Cottin and Y. Papelier, "Regulation of the cardiovascular system during dynamic exercise: integrative approach," *Crit Rev Phys Rehabil Med*, vol. 14, no. 1, 2002.
- [24] R. Bailón *et al.*, "Influence of running stride frequency in heart rate variability analysis during treadmill exercise testing," *IEEE Trans Biomed Eng*, vol. 60, no. 7, pp. 1796–1805, 2013.
- [25] F. Cottin, C. Médigue, and Y. Papelier, "Effect of heavy exercise on spectral baroreflex sensitivity, heart rate, and blood pressure variability in well-trained humans," *Am J Physiol Heart Circ Physiol*, vol. 295, no. 3, pp. H1150–H1155, 2008.
- [26] R. Bailón *et al.*, "Analysis of heart rate variability using time-varying frequency bands based on respiratory frequency," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 6674–6677, IEEE, 2007.
- [27] D. Hernando *et al.*, "Methodological framework for heart rate variability analysis during exercise: application to running and cycling stress testing," *Med Biol Eng Comput*, vol. 56, no. 5, pp. 781–794, 2018.
- [28] J. Mateo and P. Laguna, "Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal," *IEEE Trans Biomed Eng*, vol. 50, no. 3, pp. 334–343, 2003.
- [29] M. J. Tipton *et al.*, "The human ventilatory response to stress: rate or depth?," *J Physiol*, vol. 595, no. 17, pp. 5729–5752, 2017.
- [30] D. Carey, G. Pliego, and R. Raymond, "How endurance athletes breathe during incremental exercise to fatigue: interaction of tidal volume and frequency," *J Exerc Physiol Online*, vol. 11, no. 4, 2008.