Coronary artery disease diagnosis based on exercise electrocardiogram indexes from repolarisation, depolarisation and heart rate variability

R. Bailón[1] J. Mateo[1] S. Olmos[1] P. Serrano[2] J. García[1]
A. del Río[2] I. J. Ferreira[2] P. Laguna[1]

[1]Communications Technology Group, Aragón Institute of Engineering Research (I3A), University of Zaragoza, Spain
[2]Department of Medicine, 'Lozano Blesa', University Hospital, Zaragoza, Spain

Abstract—*Several indexes have been reported to improve the accuracy of exercise test electrocardiogram (ECG) analysis in the diagnosis of coronary artery disease (CAD), compared with the classical ST depression criterion. Some of them combine repolarisation measurements with heart rate (HR) information (such as the so-called ST/HR hysteresis); others are obtained from the depolarisation period (such as the Athens QRS score); finally, there are heart rate variability (HRV) indexes that account for the nervous system activity. The aim of this study was to identify the best exercise ECG indexes for CAD diagnosis. First, a method to automatically estimate repolarisation and depolarisation indexes in the presence of noise during a stress test was developed. The method is divided into three stages: first, a preprocessing step, where QRS detection, filtering and baseline beat rejection are applied to the raw ECG, prior to a weighted averaging; secondly, a post-processing step in which potentially noisy averaged beats are identified and discarded based on their noise variance; finally, the measurement step, in which ECG indexes are computed from the averaged beats. Then, a multivariate discriminant analysis was applied to classify patients referred for the exercise test into two groups: ischaemic (positive coronary angiography) and low-risk (Framingham risk index <5%). HR-corrected repolarisation indexes improved the sensitivity (SE) and specificity (SP) of the classical exercise test (SE = 90%, SP = 79% against SE = 65%, SP = 66%). Depolarisation indexes also achieved an improvement over ST depression measurements (SE = 78%, SP = 81%). HRV indexes obtained the best classification results in our study population (SE = 94%, SP = 92%) by means of the very high-frequency power (VHF) (0.4–1 Hz) at stress peak.*

Keywords—*Coronary artery disease diagnosis, Depolarisation, Exercise electrocardiogram, Heart rate variability, Repolarisation, Robust estimation*

Med. Biol. Eng. Comput., 2003, **41**, 561–571

1 Introduction

THE TRADITIONAL interpretation of the exercise electrocardiogram (ECG) for diagnosis of coronary artery disease (CAD) is based on the ST-segment depression. However, the diagnostic accuracy obtained in clinical studies is limited. A meta-analysis in GIANROSSI *et al.* (1989), including 147 published works comparing exercise-induced ST depression with coronary angiography, reported sensitivities ranging from 23% to 100% (68% ± 16%) and specificities from 17% to 100% (77% ± 17%).

The proposal of new indexes to improve the diagnostic performance of ECG analysis from stress testing has been the subject of numerous investigations.

ECG indexes that combine information on the ST level and heart rate (HR) have been shown to improve the CAD diagnostic accuracy of exercise testing compared with the classical ST-segment depression. Several reported indexes can be extracted from the diagram of the ST depression against the HR (often referred to as an ST/HR diagram) during an exercise test: HR-adjusted ST depression (DETRANO *et al.*, 1986; MORISE and DUVAL, 1995; HERPIN *et al.*, 1995; OKIN and KLIGFIELD, 1995*b*; MORISE, 1997), HR-adjusted ST integral (OKIN *et al.*, 1991*b*; HERPIN *et al.*, 1995) and ST/HR slope (KLIGFIELD and OKIN, 1999) are among them. Conventional HR-adjusted ST indexes

Correspondence should be addressed to Raquel Bailón; email: rbailon@posta.unizar.es

Paper received 12 March 2003 and in final form 4 June 2003
MBEC online number: 20033811

usually consider only the exercise phase of the test. However, the recovery phase of the exercise test has been proven to provide valuable information in cardiovascular disease diagnosis (OKIN *et al.*, 1989; WATANABE *et al.*, 2001; MORSHEDI-MEIBODI *et al.*, 2002). The so-called ST/HR hysteresis, which measures the average ST depression difference between the exercise and recovery phases relative to HR, has been proposed recently (LEHTINEN *et al.*, 1996; LEHTINEN, 1999).

On the other hand, indexes measured on the depolarisation period, such as the Athens QRS score (MICHAELIDES *et al.*, 1990; VANCAMPEN *et al.*, 1996; TOTH *et al.*, 2001) and the QRS duration (MICHAELIDES *et al.*, 1993), have also been proposed to improve the exercise test CAD diagnostic value. In other works, alternative combinations of QRS amplitude indexes have shown a better performance than the Athens QRS score (GARCÍA *et al.*, 2000).

Finally, some investigators have examined the relationships of sympathetic and parasympathetic nervous system activity (quantified by heart rate variability (HRV)) to the incidence of CAD. HRV measurements have been proposed, in combination with other ECG indexes, to improve markedly the identification of patients at risk for sudden cardiac death (TASK FORCE OF ESC AND NASPE, 1996). Investigators from the Framingham Heart Study reported that the estimation of HRV by ambulatory monitoring offered prognostic information beyond that provided by the evaluation of traditional cardiovascular risk factors (TSUJI *et al.*, 1994; 1996). Other studies described the relationship between HRV and ischaemic cardiomyopathy during ambulatory recordings (JAGER *et al.*, 1997; BILCHICK *et al.*, 2002). HRV during exercise testing has not been as widely studied; however, in DILAVERIS *et al.* (1998) and MATEO *et al.* (2001*a*) it was shown that HRV behaviour during exercise is not the same in patients with CAD as in non-ischaemic patients.

ECG signal recordings are often contaminated with undesirable signals (noise), mainly owing to environmental and physiological conditions and to the acquisition system. During an exercise test, the noise level is increased as the patient is constantly moving.

Repolarisation and depolarisation indexes have been observed to be highly sensitive to exercise test noise, especially when automatically estimated (MATEO *et al.*, 2001*b*).

The aim of this work was to identify the best exercise ECG indexes to discriminate between patients with CAD and low CAD-risk subjects. First, it was necessary to develop a signal processing method to obtain a robust automatic estimation of repolarisation and depolarisation indexes in noisy exercise test recordings.

2 Materials and methods

2.1 *Study population*

In the University Hospital 'Lozano Blesa' of Zaragoza, Spain, the ECGs of 844 patients referred for a treadmill exercise test (following the Bruce protocol (BRUCE *et al.*, 1973)) were recorded, including 66 asymptomatic Spanish Army volunteers. Standard leads (V1, V3–V6, I, II, III, aVR, aVL and aVF) and RV4 were digitally recorded at 1 kHz sampling rate with an amplitude resolution of 0.6 μV. The classical standard lead V2 was substituted by lead RV4 so that more information could be extracted from the right part of the heart, as suggested in MICHAELIDES *et al.* (1996), to improve the diagnostic accuracy of the test. The investigation conformed to the principles outlined in the Declaration of Helsinki. The procedures and protocols used in this study were approved by the Ethics Committee of the University 'Lozano Blesa' of Zaragoza. Informed consent was obtained from all subjects prior to data collection.

Patients were classified into different groups, according to the purposes of the several stages of the study:

Ischaemic: This group comprised 79 patients with significant stenoses in at least one major coronary artery, as revealed by coronary angiography (used as gold standard).

Low-risk: The gold standard to include patients in this group was the Framingham risk index. The Framingham risk algorithm computes the 10 year predicted risk of developing manifest CAD using data relating to several risk factors (age, total and HDL cholesterol, blood pressure, diabetes and smoking) (D'AGOSTINO *et al.*, 2000). The Framingham risk index was calculated for all subjects. Information on all risk factors was not available for all patients. The score given to risk factors with missing information was zero. Only patients with information on at least four (out of six) risk factors presenting an index lower than 5% were included in the low-risk group. Three of the patients classified as low-risk by the Framingham index were indeed ischaemic, as shown by coronary angiography, and were therefore excluded from this group. Finally, the low-risk group consisted of 44 patients (22 with information on six risk factors, six on five and 16 on four), all presenting a negative clinical and electrical exercise test.

Asymptomatic volunteer: This group comprised the 66 asymptomatic volunteers from the Spanish Army, who underwent an exercise test with negative results for CAD.

Non-ischaemic: This group was defined in a previous work (MATEO *et al.*, 2001*b*) and it is considered here for purposes of comparison. It included 286 patients: 220 patients with a negative clinical and electrical exercise test and reaching at least 90% of the maximum age-related heart rate plus the 66 volunteers. (The maximum age-related heart rate is computed as $HR_{max} = 220 - age$. HR_{max} and *age* are measured in beats min^{-1} and years, respectively.)

The remaining 480 non-classified patients were not analysed in this study.

A total of nine ECGs (seven from the ischaemic group, one from the low-risk group and one from the asymptomatic volunteer group) were excluded from the analysis owing to power-line interference, excessive baseline variations, signal loss or an excessive ectopic beat rate.

A summary of the study population characteristics is given in Table 1.

2.2 *ECG indexes measurement*

2.2.1 *Repolarisation and depolarisation indexes:* For repolarisation and depolarisation index estimation, two different time periods were considered: the beginning of the recording (S_1) and the exercise peak (S_2), defined as the instant of maximum heart rate (see Fig. 1). The duration of S_1 and S_2 was 11 beats.

Repolarisation indexes:

- ST indexes (*Repo*): The ST level was estimated by averaging 10 ms of ECG signal at an HR-dependent distance from the QRS fiducial point (BADILINI *et al.*, 1996)

$$STpoint = QRSpoint + (40\,\text{ms} + 1.2 \cdot RR^{1/2})$$

  The QRS fiducial point was defined as the centre of gravity of the whole QRS complex (see further detail in Section 2.2.2). The iso-electric level was obtained by averaging

Table 1 Study population characteristics

Characteristic	Ischaemic	Low-risk	Asymptomatic volunteer	Non-ischaemic
Number	72	43	65	285
Age year*	59 ± 10	41 ± 13	35 ± 1	49 ± 15
Sex (male/female)	68/4	28/15	65/0	218/66
MaxHR, beats min^{-1}*	132 ± 19	173 ± 14	185 ± 11	167 ± 17
10 year CAD risk, %*	18 ± 11	3 ± 1	†	9 ± 8

*Mean ± standard deviation (SD). †Information not available. *MaxHR* = maximum heart rate achieved

10 ms of the PR interval starting 70 ms before the QRS fiducial point. The ST level at exercise peak (S_2) was denoted ST_p. The ST difference (ΔST) was computed between S_2 and S_1. The absolute value of the ST difference ($|\Delta ST|$) was also considered, to take into account either a depression or an elevation of ST level, caused by sub-endocardial or sub-epicardial ischaemia, respectively (BAYÉS, 1992).

- HR-corrected repolarisation indexes (*Repo/HR*): Several indexes were measured from the ST/HR diagram, which is the plot of ST depression against HR during an exercise test: the ST difference corrected by the HR increment between S_2 and S_1 ($\Delta ST_c = \Delta ST/\Delta HR$) (DETRANO *et al.*, 1986), the corresponding absolute value ($|\Delta ST_c|$) and the ST/HR hysteresis (*STHL*), defined as the integrated difference between ST depression during exercise and recovery over HR from 3 min after stress peak to the maximum HR, normalised by the HR increment (LEHTINEN *et al.*, 1996) (see Fig. 2).

Depolarisation indexes (Depo): The amplitude of the Q-, R- and S-waves and QRS duration were automatically measured on the corresponding beat (see Section 2.2.2) using the system described in GARCÍA *et al.* (2000) and LAGUNA *et al.* (1994). ECG indexes ΔQ, ΔR, ΔS and ΔQRS_d were computed as the differences between the S_2 and S_1 values. Owing to the difficulty in correctly estimating the QRS duration on each lead, especially owing to the high noise level at the exercise peak, the ΔQRS_d considered in this study was the averaged QRS duration among all leads, after visual inspection and rejection of outlier measurements.

2.2.2 *Robust estimation of repolarisation and depolarisation indexes:* Exercise ECG recordings are usually very noisy, especially around the stress peak when the patient's muscular activity is high, particularly affecting automatic estimation of the repolarisation and depolarisation indexes. Therefore it was necessary to develop a method to obtain a robust automatic estimation of the repolarisation and depolarisation indexes in exercise test recordings. The method is divided into three stages: first, a preprocessing step, where QRS detection, filtering and baseline beat rejection are applied to the raw ECG, prior to a weighted averaging; then, in the post-processing step, potentially noisy averaged beats are identified and discarded based on their noise variance; finally, in the measurement step, ECG indexes are computed from the averaged beats. This method was previously evaluated by means of simulated exercise test recordings (BAILÓN *et al.*, 2002).

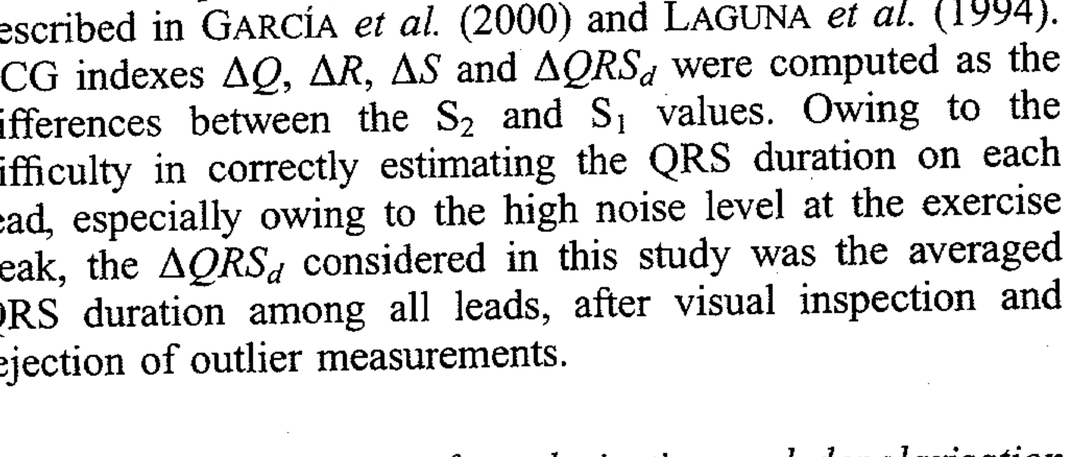

Fig. 1 *Time periods for ECG indexes' estimation plotted over HR series*

Preprocessing: The preprocessing step comprised: QRS detection and selection of normal beats (MOODY and MARK, 1982); baseline wander attenuation using cubic spline interpolation (iso-electric level knots estimated averaging 20 ms of signal starting 80 ms before QRS mark); rejection of beats presenting a difference in their mean iso-electric level with respect to adjacent beats of more than 600 μV; and, finally, owing to the beat-to-beat highly non-stationary characteristics of exercise test noise, a running weighted averaging was applied to the signal, assuming short-term stationarity for the ECG signal. Each beat was multiplied by a coefficient inversely proportional to its noise variance (HOKE *et al.*, 1984). This noise variance was estimated as the signal power after high-pass filtering (cutoff frequency = 15 Hz) in the interval ($QRS - 150$ ms–$QRS + 0.7RR$ ms), where *QRS* represents the QRS mark, and *RR* is the interval between two consecutive beats (in ms). A non-causal running weighted averaging of 11 beats was performed, sliding five beats each time. The underlying assumption was the local stationarity of the ECG signal within those 11 beats. An averaged beat and its corresponding ensemble variance (estimation of the ensemble noise) were obtained for each five-beat period.

Post-processing: Despite the first stage, unreliable measurements can still occur owing to the high noise content in exercise test recordings. In the post-processing step, those

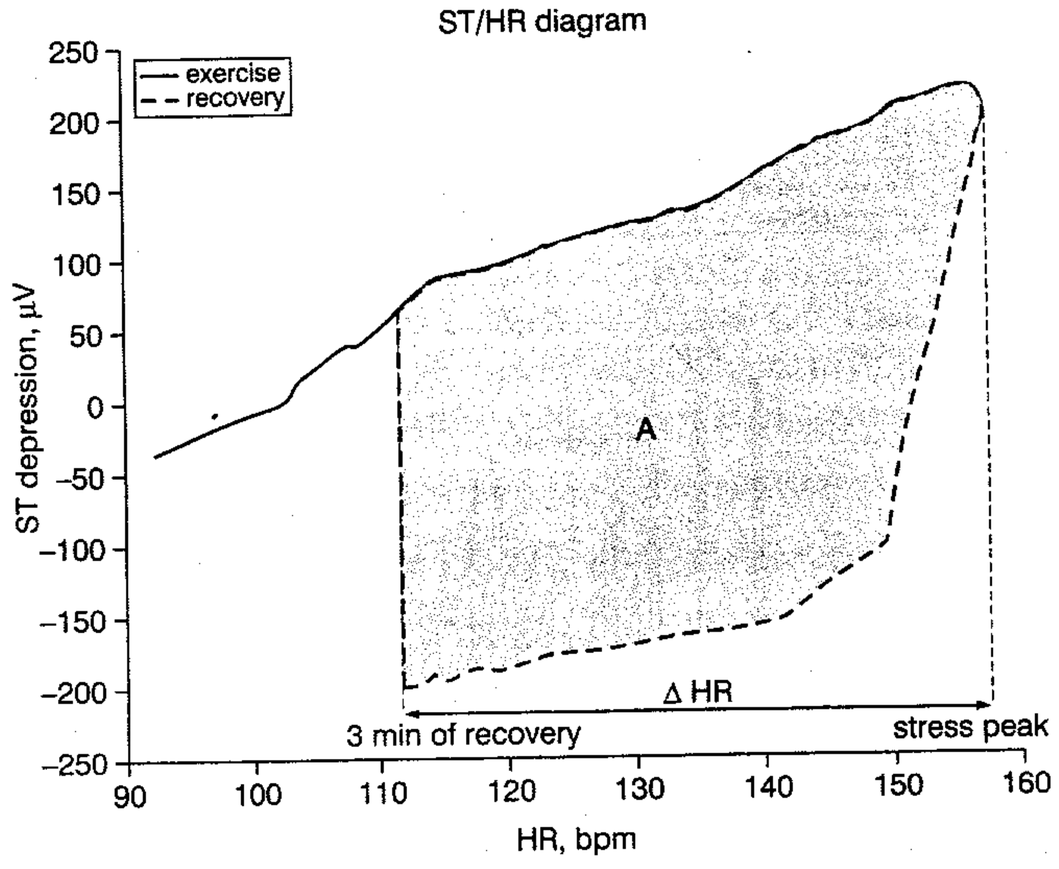

Fig. 2 *ST/HR diagram: ST/HR hysteresis = A/ΔHR; A = area between the (– – –) recovery and (———) exercise ST depression values; ΔHR = HR difference between stress peak and first 3 min of recovery*

averaged beats presenting a high noise variance, called outliers, were rejected to avoid unreliable measurements. As the noise level is not constant during the exercise test, the rejection threshold should vary according to the noise level at the different stages of the test. We used an adaptive-threshold rejection method based on the median absolute deviation (MAD) of the averaged beat noise variance (HAMPEL *et al.*, 1986). Beats whose noise variance differed from their median more than their threshold were discarded (see Fig. 3). The threshold for each averaged beat was defined as the median value (estimated in an interval of 2 min) plus its respective MAD, estimated in an interval of 5 min.

Owing to the high noise level at stress peak and its non-stationary characteristics, many or even all averaged beats surrounding the exercise peak can be considered as outliers and therefore discarded. If all the beats in an interval of 15 s before and after stress peak were outliers, the averaged beat presenting the minimum noise variance in that interval would be kept to avoid loosing all data at stress peak. The stress peak was identified as the maximum of the HR trend after a causal moving-average filtering of 5 beats, as the signal is especially noisy at that point.

Measurement: Finally, in the measurement step, the repolarisation and depolarisation indexes described in Section 2.2.1 were measured on the non-rejected weighted averaged beats. Each averaged beat was assigned an RR value, the median of the RR values of the 11 original beats included in its average. The ST/HR diagram was constructed from the averaged beat series: a single ST depression value was obtained for each HR value (mean of all ST values corresponding to the same HR, differentiating between exercise and recovery phases). In addition to the previous processing, a median filtering of nine samples was needed to smooth the ST trend (applying the median filtering without the previous processing was insufficient to obtain a reliable ST/HR diagram). ECG indexes ST_p, ΔST, $|\Delta ST|$, ΔST_c, $|\Delta ST_c|$, $STHL$, ΔQ, ΔR, ΔS and ΔQRS_d were estimated on each lead for each exercise test record.

Simulation study set-up: Evaluation of the robust estimation method is difficult, as no reference can be obtained in actual ECG records. Therefore a noise-free exercise test ECG was simulated from a set of 15 weighted averaged beats, called templates, extracted from the resting, exercise and recovery phases of a real stress test record. The HR and ST depression of each template was modified to follow a known ST/HR pattern. A continuous signal was obtained through concatenation of template repetitions, modified so that HR and ST depression varied linearly. Four different ST/HR patterns, based on actual cases reported in previous works (LEHTINEN *et al.*, 1996), were used. The final simulated stress ECG had a duration of approximately 11 min.

To take into account the presence of noise in exercise ECG recordings, an additive noise model was used. Noise records were estimated as the residual of raw stress ECG recordings (different from the one used to construct the simulated ECG) and the corresponding averaged beat series. Owing to beat-to-beat morphological variations, estimated noise records can include large QRS residuals. To cancel these spike-like residuals, a similar rejection procedure was applied, based on the MAD method previously described. Rejected samples were substituted by a random value within the threshold limits. A total of 54 different noise records (RMS levels ranging from 114 to 979 μV), extracted from different leads of exercise test recordings, were used.

Fig. 4 shows fragments of 10 s duration of the simulated records at different stages of the exercise test (resting, exercise and recovery, respectively).

2.2.3 *HRV indexes (HRV):* HRV was measured from the raw ECG (after QRS detection and selection of normal beats (MOODY and MARK (1982)) on three different 2 min duration intervals: the beginning of the exercise P_1, just before the peak of maximum exercise P_2 and during the recovery period P_3, starting 30 s after the maximum exercise peak (see Fig. 1).

- Time-domain HRV indexes (*HRV time*): Standard deviation of the normal-to-normal (NN) QRS intervals *SDNN* and root mean squared of successive NN differences *RMSSD* were calculated after linear detrending of the HR series in the P_1, P_2 and P_3 periods. The slopes *SLP* from linear detrending in each period were also included in the analysis.
- Frequency-domain HRV indexes (*HRV frequency*): The power spectral density (PSD) of HRV was estimated from the linearly detrended and interpolated heart timing signal (MATEO and LAGUNA, 2000) resampled at 2 Hz, reducing the effect of ectopic beats by the method proposed in MATEO and LAGUNA (2003). The fast Fourier transform (FFT) was applied over 2 min episodes in P_1, P_2 and P_3 (see Fig. 1). ECG indexes were defined as the spectral power in the following frequency bands: very low frequency *VLF* (0–0.04 Hz), low frequency *LF* (0.04–0.15 Hz), high frequency *HF* (0.15–0.4 Hz), and very high-frequency *VHF* (0.4–1 Hz). (The duration of the episodes considered in the frequency analysis (2 min) limits to 0.0083 Hz the lowest measurable frequency. Therefore, *VLF* is rigorously defined in this work between 0.0083 and 0.04 Hz). Also the total power *AF* was considered. The *VHF* band is proposed in this study for exercise test recordings. The power in the *VHF* band cannot be evaluated in resting conditions because of the low HR, which results in a low HR series sampling rate. However, the power in the *VHF* band becomes significant during the P_2 and P_3 periods, when the mean HR is above 120 beats min^{-1} (HR series sampling rate above 2 Hz). Fig. 5 shows the PSD of the HR trend given in Fig. 1 during the three periods analysed. Note the presence of spectral power in the *VHF* band during the P_2 and P_3 periods.

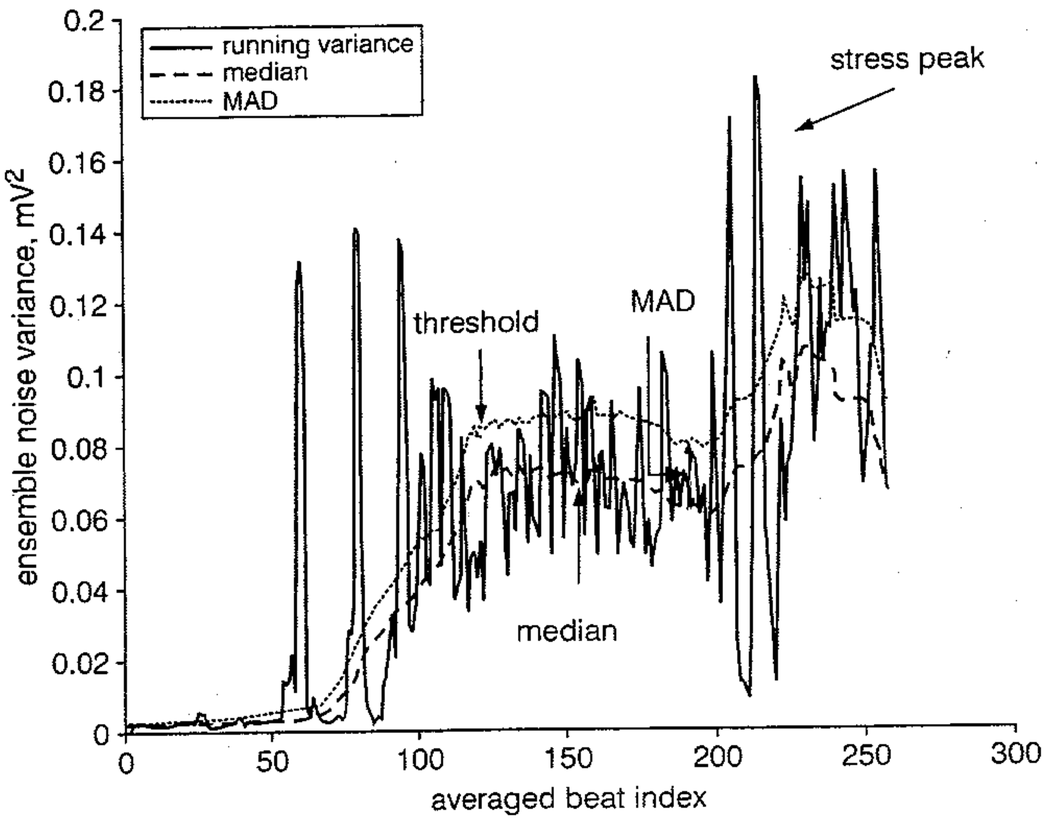

Fig. 3 *Example of beat rejection based on MAD. (——) Running variance; (– – –) median; (·····) MAD*

2.3 *Statistical analysis*

The significance test and discriminant analysis assume that the implied variables are Gaussian. However, the statistical distribution of the main HRV indexes is highly asymmetric

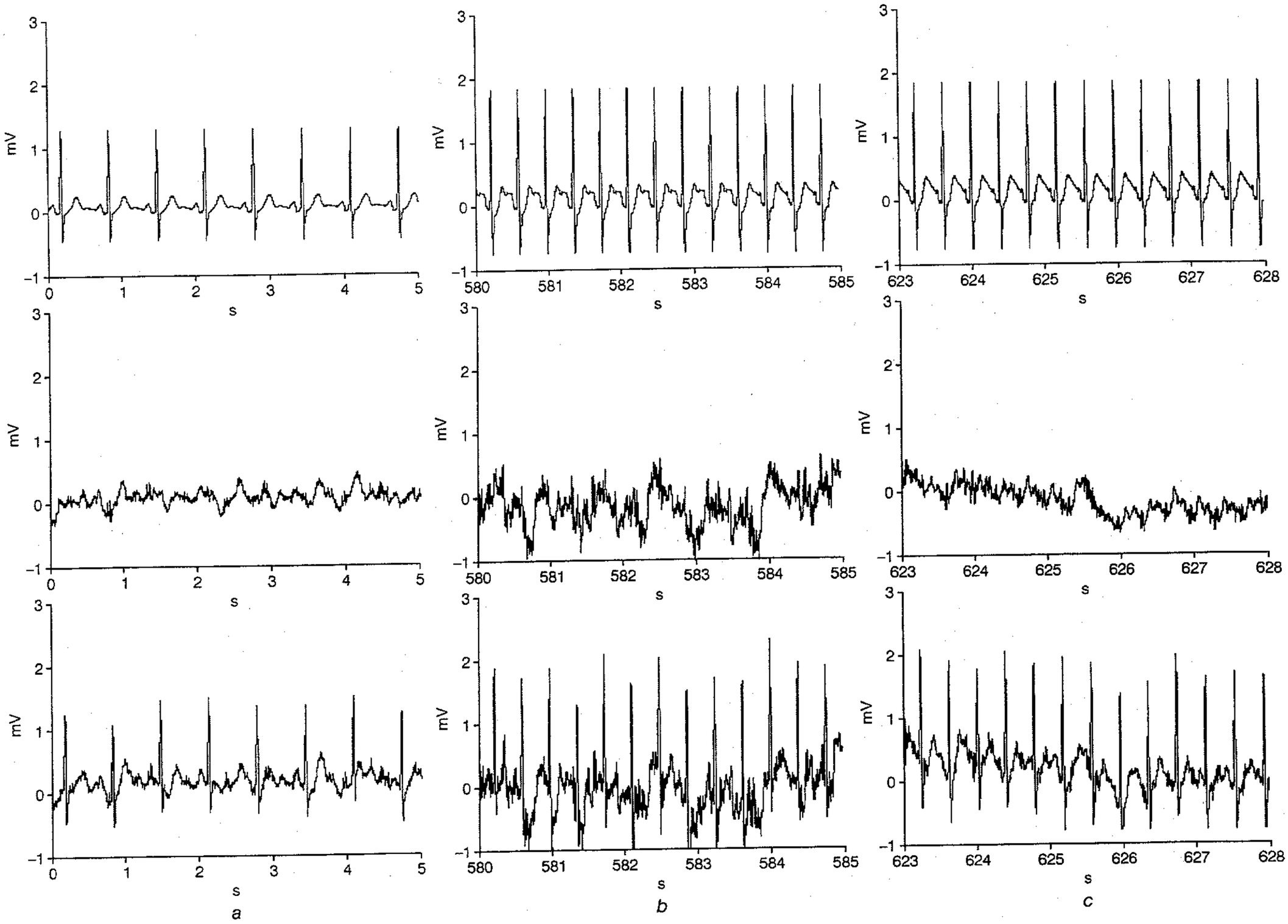

Fig. 4 *10 s fragments of simulated stress test records (noise-free ECG, noise and noisy ECG) during (a) resting, (b) exercise and (c) recovery phases*

and not Gaussian. Thus all HRV indexes were logarithmically transformed to reduce their skewness (the logarithmic transformation reduced the mean skewness of all HRV indexes from 2.80 to 0.003). Multivariate discriminant analysis (spss 11.0) was used to identify the exercise ECG indexes that best classify the patients in two groups. We used the stepwise approach in the analysis. The criterion followed in the variable inclusion/rejection was the Wilk's lambda minimisation ($F = 3.84$ for inclusion and $F = 2.71$ for rejection). The classification results were calculated with cross-validated estimation (leave-one-out). Multivariate discriminant analysis was independently applied to different sets of indexes. Four classification problems were analysed, considering different pairs of groups (refer to Section 2.1: ischaemic against non-ischaemic, ischaemic against low-risk, ischaemic against low-risk + asymptomatic volunteer and low-risk against asymptomatic volunteer). As multivariate discriminant procedures require that all observations have complete data, and not all the variables could be estimated for all the patients (such as wave amplitudes not measurable in some leads, outlier QRS widths and so on), the number of cases included in each discriminant analysis was not the same. The number of stepwise selected variables used in the discriminant analysis was truncated, when necessary, to follow the criterion of *number of variables* $< (\textit{smallest group size})^{1/2}$.

Five variables were calculated to assess the classification performance: sensitivity *SE*, specificity *SP*, positive predictive value *P*+, negative predictive value *P*− and exactness *EX*.

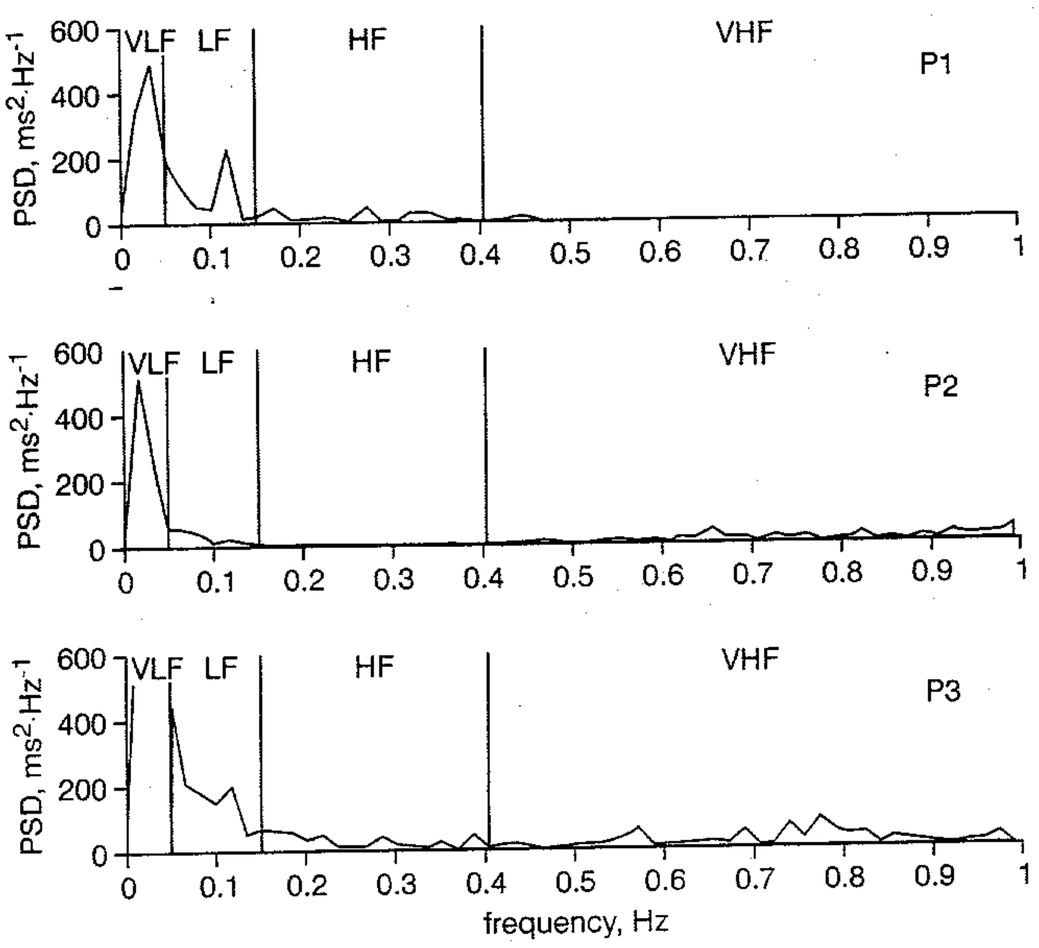

Fig. 5 *Distribution of PSD in VLF, LF, HF and VHF bands during P_1, P_2 and P_3 periods*

3 Results

3.1 *Robust index estimation method*

3.1.1 *Simulation results:* The robust index estimation method described in Section 2.2.2 was evaluated on a total of 216 records, simulated from the four different ST/HR patterns and the 54 different noise records (refer to Section 2.2.2).

Table 2 shows the reduction in the estimation error of the indexes *STHL*, ΔQ, ΔR and ΔS achieved by the method. The estimation error of *STHL* was reduced by approximately

Table 2 Reduction in estimation error of different ECG indexes by robust method described in Section 2.2.2

ECG index	*STHL**	ΔQ	ΔR	ΔS
Mean, %	14.71	22.41	9.63	5.86
meanNORP/meanRP, μ*V*	34/29	36/28	101/92	103/97
SD, %	15.85	2.27	2.67	2.54
SDNORP/SDRP, μV	63/53	127/124	158/154	173/169

Mean (*meanNORP*) and standard deviation (*SDNORP*) of estimation error without application of robust signal processing described in Section 2.2.2; mean (*meanRP*) and standard deviation (*SDRP*) of estimation error with application of robust signal processing in Section 2.2.2.
*As explained in BAILÓN *et al.* (2002)

15% both in mean (5 μV) and in standard deviation (10 μV). The reduction in the estimation error of the indexes ΔQ, ΔR and ΔS was of 22% (8 μV), 10% (9 μV) and 6% (6 μV) in mean, respectively, and of nearly 3% in standard deviation in all cases (3 μV, 4 μV and 4 μV respectively).

3.1.2 *Classification results:* The classification problem presented in MATEO *et al.* (2001*b*) was reproduced to test the clinical performance of the robust index estimation method. Those variables reported in MATEO *et al.* (2001*b*) as the best discriminatory indexes were used to classify patients in two groups: ischaemic and non-ischaemic. The ischaemic group was weighted by a factor of four to compensate for the different group sizes, for purposes of comparison with the results in MATEO *et al.* (2001*b*).

Table 3 shows the diagnostic performance (in terms of *EX*) obtained by the indexes estimated using the method presented in Section 2.2.2, in comparison with the one published before (MATEO *et al.* 2001*b*) for each variable set.

In the (*Repo*, *Repo/HR*), (*Depo*, *Repo*, *Repo/HR*) and (*Depo*, *Repo*, *Repo/HR*, *HRV*) sets, the numbers of classification variables were too large relative to the smallest group size (9 against 58, 10 against 48 and 11 against 45, respectively). Diagnostic performance was then recalculated for them using a lower number of variables, following the criterion given in Section 2.3.

Note that the diagnostic performance achieved by depolarisation and repolarisation indexes estimated by the method presented in Section 2.2.2 is higher than that reported in MATEO *et al.* (2001*b*) for all variable subsets, even when a lower number of variables is used. Classification results for HRV indexes were exactly the same as in MATEO *et al.* (2001*b*), as the beat locations used for HRV analysis were the same in both studies.

The most significant variables of each subset are printed in Table 3 in the same order as selected by the stepwise method.

Table 3 Effect of robust ECG index estimation method described in Section 2.2.2 in CAD diagnosis performance: ischaemic against non-ischaemic

Variable set	*N*	Cases	*EX**, %	*EX*†, %
(*ST peak*) STp_{aVL}	1	71/102	58	56
(Repo) ΔST_{aVL}, $\lvert\Delta ST_{V3}\rvert$, ΔST_{RV4}, $\lvert\Delta ST_{II}\rvert$, STp_{V5}, ΔSTc_{V3}	6	70/100	69	58
(*Repo*, *Repo/HR*)	9‡	58/90	92	85
$\lvert\Delta STc_{V3}\rvert$, $STHLv_3$, $STHL_{V1}$, $STHL_{II}$, $\lvert\Delta ST_{V6}\rvert$, $\lvert\Delta STc_{V6}\rvert$, ΔSTc_{V3}	7**	59/90	90	
(ΔQ, ΔR, ΔS) ΔR_{aVF}, ΔS_{V6}, ΔQ_{V6}	3	63/80	83	69
(*QRS width*) $\Delta QRSd_{III}$, $\Delta QRSd_{V6}$, ΔQRS_{II}	3	42/64	70	57
(*Depo*) ΔR_{aVF}, ΔS_{V6}, ΔQ_{V6}, ΔR_I, $\Delta QRSd_{III}$	5	47/55	81	70
(*Depo*, *Repo*, *Repo/HR*)	10†	48/55	92	86
$\lvert\Delta STc_{V3}\rvert$, ΔS_{V6}, $STHL_{aVR}$, $STHL_{V1}$, ΔS_{aVL}, $\lvert\Delta ST_{V6}\rvert$, ΔSTc_{V3}	7**	54/63	91	
(*HRV time*) $RMSSD_{P2}$, $SDNN_{P1}$, $SDNN_{P2}$, $RMSSD_{P3}$	4	71/275	73	73
(*HRV frequency*) VHF_{P2}, HF_{P3}, SLP_{P3}, VLF_{P2}, HF_{P2}, AF_{P1}	6	63/240	79	79
(*HRV*) VHF_{P2}, $RMSSD_{P2}$, VLF_{P2}, LF_{P2}, $SDNN_{P1}$, LF_{P1}, HF_{P3}	7	63/240	83	83
(*Depo*, *Repo*, *Repo/HR*, *HRV*)	11†	45/55	97	91
$STHL_{V6}$, VHF_{P2}, $RMSSD_{P2}$, $\lvert\Delta STc_{V3}\rvert$, LF_{P2}, VLF_{P2}	6**	56/83	92	

N = number of variables; cases = ischaemic/non-ischaemic
*Applying robust ECG index estimation method described in Section 2.2.2
†Without robust index estimation as in MATEO *et al.* (2001*b*)
Depo = depolarisation indexes. *Repo* = repolarisation indexes. *Repo/HR* = HR-adjusted repolarisation indexes. *HRV* = HRV indexes
‡Number of variables too large relative to sample size
**Using conservative number of variables (≈ (smallest group size)$^{1/2}$)

3.2 *ECG index diagnostic performance*

3.2.1 *Ischaemic against low-risk:* The next goal in our study was to identify those variables that best discriminate between the ischaemic and low-risk groups (refer to Section 2.1).

First, we applied the analysis of variance (ANOVA) method to test the hypothesis that the group means were different. The most significant variables for discriminating the groups ($p<0.001$) are shown in Table 4.

Then, multivariate discriminant analysis was independently applied to different sets of variables. A summary of the results is given in Table 5 in terms of *SE*, *SP*, *P+*, *P−* and *EX*. The classification variables used in each case are listed in the same order as selected by the stepwise method.

The sensitivity and specificity achieved by repolarisation indexes ($SE=65\%$, $SP=66\%$) were improved when HR information was added to ST level indexes ($SE=90\%$, $SP=79\%$). Note that ST/HR hysteresis indexes alone obtained a diagnostic performance ($SE=89\%$, $SP=75\%$) similar to all HR-corrected repolarisation indexes. A combination of depolarisation indexes outperformed classical repolarisation indexes ($SE=78\%$, $SP=81\%$). Observe that addition of depolarisation indexes to HR-corrected repolarisation indexes did not add further improvement ($SE=90\%$, $SP=76\%$ against $SE=90\%$, $SP=79\%$). HRV indexes obtained the best classification results ($SE=94\%$, $SP=92\%$).

The next attempt was to use the ischaemic against non-ischaemic discriminatory variables obtained in MATEO *et al.* (2001*b*) to classify the patients of the present study in their corresponding group: ischaemic or low-risk. The classification

Table 4 Most significant variables ($p<0.001$) for discriminating ischaemic against ow-risk, according to ANOVA

Repolarisation	*STHL* in all leads
Depolarisation	ΔQ_{V5}, ΔQ_{V6}, ΔQ_{aVR}, ΔQ_{II}, ΔQ_{aVF} ΔR_{V6}, ΔR_{aVR}, ΔR_{II}, ΔR_{aVF}
HRV	VLF_{P1}, LF_{P1}, HF_{P1}, VHF_{P1}, AF_{P1} VHF_{P2}, AF_{P2} $RMSSD_{P3}$, SLP_{P3}

capabilities of the different variable sets were compared with those achieved previously in MATEO *et al.* (2001*b*) (see Table 5).

It can be appreciated that classification results did not change significantly. Note that only in the (*Depo*, *Repo*, *Repo/HR*) and the (*HRV*) sets was diagnostic performance slightly lower. In the (*Depo*, *Repo*, *Repo/HR*, *HRV*) analysis, two different sets of four and six variables were used, achieving an *EX* of 82% and 86%, respectively. A new set including the most significant variables of depolarisation, repolarisation and HRV was defined, obtaining an *EX* of 85% (see the last row in Table 5). The most relevant variable was the *VHF* power at exercise peak, accounting for 9% of the total *EX*.

3.2.2 *Ischaemic against low-risk + asymptomatic volunteer:* The variable sets previously cited were used to classify patients into two groups: ischaemic and a new group including the low-risk and asymptomatic volunteer groups.

Classification results are summarised in Table 6. A slight increase in diagnostic performance can be observed when

Table 5 Classification results: ischaemic against low-risk

Variable set	*N*	Cases	*SE*, %	*SP*, %	*P+*, %	*P−*, %	*EX*, %
(*ST peak*) STp_{aVL}	1	72/42	54	71	76	48	61
(*Repo*) *STp_{aVL}, STp_{V1}, $\lvert\Delta ST_{RV4}\rvert$, $\lvert\Delta ST_{V6}\rvert$, STp_{III}, STp_{aVF} †ΔST_{aVL}, $\lvert\Delta ST_{V3}\rvert$, ΔST_{RV4}, $\lvert\Delta ST_{II}\rvert$, STp_{V5}, $\lvert\Delta ST_{V6}\rvert$	6 6	71/41 71/41	65 58	66 63	77 73	52 46	65 60
(*ST/HR hysteresis*) $STHL_{II}$, $STHL_{V3}$, $STHL_{aVR}$, $STHL_{I}$, $STHL_{V1}$	5	61/36	89	75	86	79	84
(*Repo*, *Repo/HR*) *$STHL_{V5}$, ΔST_{V6}, $\lvert\Delta ST_{V1}\rvert$, STp_{I}, $STHL_{II}$ †$\lvert\Delta STc_{V3}\rvert$, $STHL_{V3}$, $STHL_{V1}$, $STHL_{II}$, $\lvert\Delta ST_{V6}\rvert$, $\lvert\Delta STc_{V6}\rvert$, ΔSTc_{V3}	5 7	59/34 60/34	90 87	79 88	88 93	82 79	86 87
(*Depo*) *ΔQ_{II}, ΔQRS_d, ΔQ_{aVL}, ΔS_{aVR}, ΔS_{V1} †ΔR_{aVF}, ΔS_{V6}, ΔQ_{V6}, ΔR_{I}, $\Delta QRSd_{III}$	5 5	58/27 48/26	78 83	81 69	90 83	63 69	79 78
(*Depo*, *Repo*, *Repo/HR*) *ΔQ_{V6}, ΔS_{V4}, $\lvert\Delta ST_{RV4}\rvert$, $STHL_{V3}$, $STHL_{II}$ †$\lvert\Delta STc_{V3}\rvert$, ΔS_{V6}, $STHL_{aVR}$, $STHL_{V1}$, ΔS_{aVL}, $\lvert\Delta ST_{V6}\rvert$, ΔSTc_{V3}	5 7	58/29 55/24	90 80	76 62	88 83	79 58	85 75
(*HRV time*) *$RMSSD_{P2}$, $SDNN_{P1}$, $SDNN_{P2}$, $RMSSD_{P3}$, $SDNN_{P3}$ †$RMSSD_{P2}$, $SDNN_{P1}$, $SDNN_{P2}$, $RMSSD_{P3}$	5 4	72/43 72/43	74 74	77 77	84 84	63 63	75 75
(*HRV frequency*) *LF_{P1}, VHF_{P2}, HF_{P3}, SLP_{P3}, VLF_{P2}, HF_{P2} †VHF_{P2}, HF_{P3}, SLP_{P3}, VLF_{P2}, HF_{P2}, AF_{P1}	6 6	64/39 64/39	89 89	87 90	92 93	83 83	88 89
(*HRV*) *VHF_{P2}, VLF_{P2}, $SDNN_{P1}$, HF_{P3}, SLP_{P3}, $SDNN_{P2}$ †VHF_{P2}, $RMSSD_{P2}$, VLF_{P2}, LF_{P2}, $SDNN_{P1}$, LF_{P1}, HF_{P3}	6 7	65/40 64/39	94 84	92 90	95 93	90 78	93 86
(*Depo*, *Repo*, *Repo/HR*, *HRV*) *ΔQ_{V6}, $STHL_{II}$, $STHL_{V3}$, $\lvert\Delta ST_{V1}\rvert$ †$STHL_{V6}$, VHF_{P2}, $RMSSD_{P2}$, $\lvert\Delta STc_{V3}\rvert$, LF_{P2}, VLF_{P2} ‡VHF_{P2}, $STHL_{V5}$, VLF_{P2}, ΔQ_{II}, ΔQRS_d, ΔST_{V6}	4 6 6	59/30 56/31 54/28	85 84 85	77 90 86	88 94 92	72 76 75	82 86 85

N = number of variables; cases = ischaemic/low-risk
Depo = depolarisation indexes; *Repo* = repolarisation indexes; *Repo/HR* = HR-adjusted repolarisation indexes; *HRV* = HRV indexes
*Classification variables selected by stepwise method to discriminate between ischaemic and low-risk
†Classification variables obtained in MATEO *et al.* (2001*b*) to discriminate between ischaemic and non-ischaemic
‡Most significant variables of depolarisation, repolarisation and HRV to discriminate between ischaemic and low risk

Table 6 Classification results: (1) ischaemic against low-risk; (2) ischaemic against low-risk + asymptomatic volunteer; (3) low-risk against asymptomatic volunteer

Variable set		(1)		(2)		(3)	
	N	Cases	*EX*, %	Cases	*EX*, %	Cases	*EX*, %
(*ST peak*)							
STp_{aVL}	1	72/42	61	72/101	62	42/59	50
(*Repo*)							
*STp_{aVL}, ST_{pV1}, $\|\Delta ST_{RV4}\|$, $\|\Delta ST_{V6}\|$, STp_{III}, STp_{aVF}	6	71/41	65	71/100	65	41/59	57
†ΔST_{aVL}, $\|\Delta ST_{V3}\|$, ΔST_{RV4}, $\|\Delta ST_{II}\|$, STp_{V5}, $\|\Delta ST_{V6}\|$	6	71/41	60	71/99	65	41/58	54
(*ST/HR hysteresis*)							
$STHL_{II}$, $STHL_{V3}$, $STHL_{aVR}$, $STHL_{I}$, $STHL_{V1}$	5	61/36	84	61/92	88	36/56	74
(*Repo, Repo/HR*)							
*$STHL_{V5}$, ΔST_{V6}, $\|\Delta ST_{V1}\|$, STp_{I} , $STHL_{II}$	5	59/34	86	59/89	87	34/55	73
†$\|\Delta STc_{V3}\|$, $STHL_{V3}$, $STHL_{V1}$, $STHL_{II}$, $\|\Delta ST_{V6}\|$, $\|\Delta STc_{V6}\|$, ΔSTc_{V3}	7	60/34	87	60/89	88	34/55	71
(*Depo*)							
*ΔQ_{II}, ΔQRS_{d}, ΔQ_{aVL}, ΔS_{aVR}, ΔS_{V1}	5	58/27	79	58/58	74	27/31	57
†ΔR_{aVF}, ΔS_{V6}, ΔQ_{V6}, ΔR_{I}, $\Delta QRSd_{III}$	5	48/26	78	48/55	83	26/29	64
(*Depo, Repo, Repo/HR*)							
*ΔQ_{V6}, ΔS_{V4}, $\|\Delta ST_{RV4}\|$, $STHL_{V3}$, $STHL_{II}$	5	58/29	85	58/72	91	29/43	74
†$\|\Delta STc_{V3}\|$, ΔS_{V6}, $STHL_{aVR}$, $STHL_{V1}$, ΔS_{aVL}, $\|\Delta ST_{V6}\|$, ΔSTc_{V3}	7	55/24	75	55/63	85	24/39	75
(*HRV time*)							
*$RMSSD_{P2}$, $SDNN_{P1}$, $SDNN_{P2}$, $RMSSD_{P3}$, $SDNN_{P3}$	5	72/43	75	72/105	82	43/62	72
†$RMSSD_{P2}$, $SDNN_{P1}$, $SDNN_{P2}$, $RMSSD_{P3}$	4	72/43	75	72/105	81	43/62	72
(*HRV frequency*)							
*LF_{P1}, VHF_{P2}, HF_{P3}, SLP_{P3}, VLF_{P2}, HF_{P2}	6	64/39	88	64/92	92	39/53	71
†VHF_{P2}, HF_{P3}, SLP_{P3}, VLF_{P2}, HF_{P2}, AF_{P1}	6	64/39	89	64/92	91	39/53	64
(*HRV*)							
*VHF_{P2}, VLF_{P2}, $SDNN_{P1}$, HF_{P3}, SLP_{P3}, $SDNN_{P2}$	6	65/40	93	65/94	94	40/54	68
†VHF_{P2}, $RMSSD_{P2}$, VLF_{P2}, LF_{P2}, $SDNN_{P1}$, LF_{P1}, HF_{P3}	7	64/39	86	64/92	93	39/53	72
(*Depo, Repo, Repo/HR, HRV*)							
*ΔQ_{V6}, $STHL_{II}$, $STHL_{V3}$, $\|\Delta ST_{V1}\|$	4	59/30	82	59/77	88	30/47	61
†$STHL_{V6}$, VHF_{P2}, $RMSSD_{P2}$, $\|\Delta STc_{V3}\|$, LF_{P2}, VLF_{P2}	6	56/31	86	56/82	93	31/51	71
‡VHF_{P2}, $STHL_{V5}$, VLF_{P2}, ΔQ_{II}, ΔQRS_{d}, ΔST_{V6}	6	54/28	85	54/68	90	28/40	76

N = number of variables; *Depo* = depolarisation indexes; *Repo* = repolarisation indexes; *Repo/HR* = HR-adjusted repolarisation indexes. HRV = HRV indexes

*Classification variables selected by stepwise method to discriminate between ischaemic and low-risk

†Classification variables obtained in MATEO *et al.* (2001*b*) to discriminate between ischaemic and non-ischaemic

‡Most significant variables of depolarisation, repolarisation and HRV to discriminate between ischaemic and low-risk

low-risk and asymptomatic volunteer were considered as a single group against ischaemic. The most remarkable increments (from 6% to 10%) in *EX* appeared in the (*Depo*, *Repo*, *Repo/HR*), (*HRV time*), (*HRV*) and (*Depo*, *Repo*, *Repo/HR*, *HRV*) sets.

3.2.3 *Low-risk against asymptomatic volunteer:* Finally, we attempted to discriminate between the low-risk and the asymptomatic volunteer groups by means of all variable sets. Classification results are shown in Table 6.

Some variable sets were unable to discriminate between the low-risk and the asymptomatic volunteer groups, such as repolarisation (without HR information) and depolarisation indexes. On the other hand, ST/HR hysteresis indexes achieved an *EX* of 74% and (*Repo*, *Repo/HR*) and (*Depo*, *Repo*, *Repo/HR*) sets of approximately 72% and 74%, respectively. Time-domain HRV indexes exhibited a slightly higher discriminatory ability than frequency-domain indexes ($EX = 72\%$ against $EX = 68\%$).

4 Discussion

4.1 *Robust index estimation method*

In this work we have presented a robust method to estimate automatically repolarisation and depolarisation indexes on noisy exercise test ECGs. It achieved a reduction in the estimation error of the indexes *STHL*, ΔQ, ΔR and ΔS of 14.7% (5 μV), 22.4% (8 μV), 9.6% (9 μV) and 5.9% (6 μV) in mean, respectively, and of 15.9% (10 μV), 2.3% (3 μV), 2.7% (4 μV) and 2.5% (4 μV) in standard deviation, respectively. Note that the reduction in the ΔQ estimation error may seem larger than that in ΔR and ΔS as it is expressed as a percentage of the index value, influenced by the corresponding wave amplitude. Concerning the *STHL* index, it was shown in BAILÓN *et al.* (2002) that the major improvement of the method was not provided by its error reduction but by the possibility of using the processed ST/HR diagram for monitoring, as the trace was much neater.

Another possible approach to evaluate the robust estimation method performance would have been to use a pharmacological stress test ECG as the reference signal, as recordings from drug infusion stress are much more noise free than those from exercise stress.

Despite the slight improvement achieved in ECG index estimation in simulated records, when real exercise test recordings were analysed, an improvement in classification results was appreciated with respect to MATEO *et al.* (2001*b*) after the robust estimation method was applied to the repolarisation and depolarisation indexes. The robust index estimation method described in Section 2.2.2 clearly outperformed the diagnostic accuracy of ECG indexes to stratify ischaemic and non-ischaemic patients (see Table 3).

The ischaemic against non-ischaemic classification issue presented some limitations. HR information was used to

define the non-ischaemic group (see Section 2.1) and was also considered to compute some of the ECG indexes (see Section 2.2.1). As a result, those indexes may have biased extra classification capabilities. Furthermore, the non-ischaemic group was not homogeneous nor representative of the usual non-ischaemic population undergoing an exercise test, as it included asymptomatic healthy volunteers from the Army together with non-ischaemic patients (identified by exercise test results). This was the reason to define the low-risk group in this study and to keep the non-ischaemic group only for purposes of robust estimation method validation.

4.2 *ECG indexes diagnostic performance*

In this work, we have studied the CAD diagnostic performance of several ECG indexes, reported in medical literature. In MATEO *et al.* (2001*b*), the classification problem of ischaemic against non-ischaemic was studied. Its main limitations were the biased extra classification capabilities of some indexes and the inhomogeneity of the non-ischaemic group, as explained in the preceding Section. In this work, we proposed a new classification problem that avoided those limitations by defining two groups that were as homogeneous as possible, and independent of any information extracted from the exercise test, e.g. the maximum HR. The two groups considered were ischaemic and low-risk (refer to Section 2.1).

In this study population (ischaemic against low-risk), the diagnostic performance of ST level ($SE = 54\%$, $SP = 71\%$) was in the range of those reported in other publications ($SE = 68 \pm 16\%$, $SP = 77 \pm 17\%$) (GIANROSSI *et al.*, 1989).

Around 1990, Michaelides *et al.* proposed a new index measured on the depolarisation period, the Athens QRS score, which included exercise-induced changes in Q-, R- and S-wave amplitudes, to improve the diagnostic exercise test value (MICHAELIDES *et al.*, 1990). In our study, depolarisation indexes showed a better diagnostic performance than ST indexes ($SE = 78\%$, $SP = 81\%$ against $SE = 65\%$, $SP = 66\%$), which is in accordance with the results published in MICHAELIDES *et al.* (1990), VANCAMPEN *et al.*, (1996) and TOTH *et al.* (2001). However, alternative combinations of QRS amplitude variables showed a better performance in our study population than the Athens QRS score, as previously reported in MATEO *et al.* (2001*b*) and GARCÍA *et al.* (2000).

Since Detrano *et al.* suggested, in 1986, the addition of HR information to ST indexes to increase stress test accuracy in CAD diagnosis (DETRANO *et al.*, 1986), different HR-adjusted ST-based indexes have been proposed. Among them, the ST/HR hysteresis appeared the most competent (LEHTINEN *et al.*, 1996). This was corroborated by classification results for ST/HR hysteresis in our study population ($SE = 89\%$, $SP = 75\%$), supporting the idea that the recovery phase of the exercise test provides valuable diagnostic information on cardiovascular disease (OKIN *et al.*, 1989; MORSHEDI-MEIBODI *et al.*, 2002; WATANABE *et al.*, 2001). The whole HR-adjusted ST-based indexes improved *EX* by approximately 20% over ST indexes in our study population (86% against 65%), which is in accordance with previous works (DETRANO *et al.*, 1986; MORISE and DUVAL, 1995; HERPIN *et al.*, 1995; OKIN and KLIGFIELD, 1995*b*; LEHTINEN *et al.*, 1996; LEHTINEN, 1999; MORISE, 1997). In HERPIN *et al.* (1995), different HR-adjusted ST-based indexes, not including ST/HR hysteresis, were studied, achieving a higher specificity than standard ST measurement but not better sensitivity. In contrast, in our study both sensitivity and specificity were improved by HR-adjusted ST indexes ($SE = 90\%$, $SP = 79\%$ against $SE = 65\%$, $SP = 66\%$), suggesting that ST/HR hysteresis can increase the sensitivity of the exercise test. Moreover, classification results obtained by ST/HR hysteresis were similar to those achieved by all repolarisation indexes together.

Toth *et al.* found, in 2001, that QRS score was unrelated to exercise-induced ST-segment displacement and therefore complementary (TOTH *et al.*, 2001). On the contrary, in our study, addition of depolarisation indexes to the ST and HR-adjusted ST-based indexes did not improve the diagnostic value ($SE = 90\%$, $SP = 76\%$) with respect to the previous ones, suggesting that depolarisation information could be redundant with repolarisation indexes, at least when HR information is considered.

Finally, since the late 1990s, some studies have pointed out the potential relationship between ischaemia and HRV. In previous works (MATEO *et al.*, 2001*a*; *b*; DILAVERIS *et al.*, 1998), HRV during exercise test was analysed and compared in patients with CAD and in healthy subjects. In MATEO *et al.* (2001*b*), it was shown that HRV indexes improved the diagnostic value of the test, compared with ST indexes. In the present work, HRV indexes obtained the best diagnostic performance of all variable sets. Frequency-domain indexes were more accurate than time-domain indexes ($SE = 89\%$, $SP = 87\%$ against $SE = 74\%$, $SP = 77\%$), and the combination of both achieved classification results ($SE = 94\%$, $SP = 92\%$) similar to those obtained by exercise echocardiography ($SE = 85\%$, $SP = 84\%$ (WARNIK *et al.*, 1996)) or nuclear imaging ($SE = 90\%$, $SP = 90\%$ (ISKANDRIAN *et al.*, 1997)).

In Table 5, it may seem strange that the *EX* obtained with all ECG indexes (82%) is lower than that achieved with only HRV indexes (93%). The reason is that some repolarisation and/or depolarisation indexes could not be measured on all patients (mainly owing to poor signal quality, which makes automatic wave detectors fail), therefore reducing the number of cases included in the discriminant analysis in comparison with the HRV index analysis. Classification results obtained from varying-size data sets must be analysed cautiously, as diagnostic performance from the smaller ones could be optimistically biased. Consequently, the number of cases for which a variable set could be determined should be considered together with diagnostic performance to judge the practical clinical value of a variable set. In our study, HRV exhibited the best compromise *diagnostic performance/available number of cases*.

The most significant ECG indexes obtained in this study to discriminate between the ischaemic and the low-risk groups differed from those reported in the previous study (MATEO *et al.*, 2001*b*) to discriminate between the ischaemic and the non-ischaemic group. However, some main features remained unchanged.

First, *aVL* appeared as the most relevant lead to measure ST level at exercise peak. Significant ST-based indexes collected information from pseudo-orthogonal leads: *aVL*, *RV*4 and *V*6 remained constant in both works. Information from lead *RV*4 may be recovering those cases with occlusion in the right coronary artery, as suggested by MICHAELIDES *et al.* (1996). When HR information was added to ST indexes, leads *V*6, *V*1 and *II* were relevant in both studies.

The most significant depolarisation indexes were measured in different leads but represented, in both studies, information mainly from the frontal plane of the body (*II*, *aVL*, *aVR* and *V*1, and *aVF*, *V*6, *I* and *II*, respectively). QRS width was selected as a discriminatory variable in both cases, thus supporting the findings in MICHAELIDES *et al.* (1993).

The most relevant HRV indexes were nearly the same in both studies: *VHF* and *VLF* at exercise peak, *HF* at recovery, and *SDNN* at rest. To the authors' knowledge, the *VHF* index was first analysed in exercise test recordings in MATEO *et al.* (2001*a*). The *VHF* index could be related to the sympathetic tone, as it is only significant at high HR and during exercise. It had the highest discriminant power among all HRV indexes, being in agreement with MATEO *et al.* (2001*a*), where it was shown that

VHF power increased during exercise and that it was significantly higher in the non-ischaemic population.

When all indexes were jointly considered, a recombination of variables and leads occurred, resulting in totally different selected variables in both studies. This suggests that more than one variable set could represent all the relevant diagnostic information of an exercise test.

We attempted to stratify the ischaemic and low-risk groups using the classification variables obtained in MATEO *et al.* (2001*b*) to discriminate between the ischaemic and non-ischaemic groups. Classification results did not change significantly, thus supporting our previous suggestion that diagnostic information could be collected by different sets of indexes. Only in the (*Depo*, *Repo*, *Repo/HR*) and the (*HRV*) sets was diagnostic performance slightly lower. This could indicate, either that stepwise selected variables were highly dependent on the study population in those sets, or that the low-risk and non-ischaemic groups presented significant differences in those indexes.

In the last step of this work, we studied the electrocardiographic similarities between the asymptomatic healthy volunteers from the Army and the usual non-ischaemic patients. To this end, we used the ischaemic against non-ischaemic and the ischaemic against low-risk stepwise selected variables to discriminate between the ischaemic group and a group composed of low-risk plus asymptomatic volunteer and between low-risk and asymptomatic volunteer.

A slight increase in diagnostic performance was observed when the low-risk and asymptomatic volunteer groups were jointly considered against the ischaemic group, showing that asymptomatic volunteer electrocardiographic features were at least as different from ischaemic ones as low-risk characteristics.

When considering the low-risk against asymptomatic volunteer groups, some variable sets were unable to discriminate between them, such as repolarisation (without HR information) and depolarisation indexes, meaning that these features were not significantly different in both groups. However, HRV indexes and HR-adjusted ST-based indexes achieved an *EX* of around 73%. This suggested that these indexes were significantly different, not only between the ischaemic and low-risk groups but also between the low-risk and asymptomatic volunteer groups.

A few limitations of the present study should be addressed. The first one is the influence of referral bias: the conventional interpretation of the exercise ECG probably affected the decision to proceed with coronary angiography. This limitation is unavoidable in ECG studies using coronary angiography as the gold standard, because it is impractical in clinical routine to investigate all patients with coronary angiography, regardless of the results of the preceding exercise test.

Another limitation of the work is the unbalanced number of males and females in each of the study population groups, as it has been shown that exercise-induced electrocardiographic changes are gender-dependent (OKIN and KLIGFIELD, 1995*a*). Gender-specific analysis might improve the diagnostic accuracy of the exercise test in the detection of CAD.

Finally, comparing HR-adjusted ST-based indexes' diagnostic performance reported in different works is not easy, as they are highly dependent on the ST-segment measurement point (OKIN *et al.*, 1991*a*), which varies widely among the different works.

5 Conclusions

The CAD diagnostic performance of several exercise ECG indexes was analysed: ST/HR hysteresis improved the diagnostic accuracy of the classical exercise test, based only on ST-level measurements; depolarisation indexes did not add significant information to HR-corrected repolarisation indexes, but might be redundant; finally, HRV indexes exhibited a potential value in CAD diagnosis, especially by means of the VHF power at stress peak.

To deal with high noise levels in exercise ECG recordings, mainly at stress peak, we developed a robust method to automatically estimate repolarisation and depolarisation indexes. When this method was applied, the CAD diagnostic performance in the ischaemic/non-ischaemic population was significantly improved (increase in both *SE* and *SP*).

The improvement achieved in stress test diagnostic performance would reduce the number of uncomfortable, and expensive and unnecessary interventions performed, such as coronary angiography and other techniques, and would allow clinical efforts to be focussed on CAD-risk patients. Nevertheless, the inclusion of these ECG indexes in routine exercise test trials still needs further studies in prospective populations.

Acknowledgments—This work was supported by projects TIC 2001-2167-CO2-02 from MCYT and P075/2001 from CONSIDDGA (Spain).

References

BADILINI, F., ZAREBA, W., TITLEBAUM, E., and MOSS, A. (1996): 'Analysis of ST segment variability in Holter recordings', in 'Noninvasive electrocardiology: clinical aspects of Holter monitoring. Frontiers in cardiology' (W.B. Saunders Company Ltd, UK, 1996), pp. 357–372

BAILÓN, R., OLMOS, S., SERRANO, P., GARCÍA, J., and LAGUNA, P. (2002): 'Robust measure of ST/HR hysteresis in stress test ECG recordings', in 'Computers in cardiology', vol. 29 (IEEE Computer Society Press, 2002), pp. 329–332

BAYÉS, A. (1992): 'Tratado de electrocardiografía clínica' (Ed.) (Científico-Técnica, Barcelona, Spain, 1992)

BILCHICK, K., FETICS, B., DJOUKENG, R., FISHER, S., FLETCHER, R., SINGH, S., NEVO, E., and BERGER, R. (2002): 'Prognostic value of heart rate variability in chronic congestive heart failure (Veterans Affairs' Survival Trial of Antiarrhythmic Therapy in Congestive Heart Failure)', *Am. J. Cardiol.*, **90**, pp. 24–28

BRUCE, R., KUSUMI, F., and HOSMER, D. (1973): 'Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease', *Am. Heart J.*, **85**, pp. 546–562

D'AGOSTINO, R., RUSSELL, M., HUSE, D., ELLISON, C., SILBERSHATZ, H., WILSON, P., *et al.* (2000): 'Primary and subsequent coronary risk appraisal: new results from the Framingham study', *Am. Heart J.*, **139**, pp. 272–281

DETRANO, R., SALCEDO, E., PASSALCQA, M., and FRIIS, R. (1986): 'Exercise electrocardiographic variables: a critical appraisal', *J. Am. Coll. Cardiol.*, **8**, pp. 836–847

DILAVERIS, P., ZERVOPOULOS, G., MICHAELIDES, A., SIDERIS, S., PSOMADAKI, Z., GIALAFOS, E., GIALAFOS, J., and TOUTOUZAS, P. (1988): 'Ischemia-induced reflex sympathoexcitation during the recovery period after maximal treadmill exercise testing', *Clin. Cardiol.*, **21**, pp. 585–590

GARCÍA, J., SERRANO, P., BAILÓN, R., GUTIÉRREZ, E., DEL RIO, A., CASASNOVAS, J., FERREIRA, I., and LAGUNA, P. (2000): 'Comparison of ECG-based clinical indexes during exercise test', in 'Computers in cardiology', vol. 27 (IEEE Computer Society Press, 2000), pp. 833–836

GIANROSSI, R., DETRANO, R., MULVIHILL, D., LEHMANN, K., DUBACH, P., COLOMBO, A., MCARTHUR, D., and FROELICHER, V. (1989): 'Exercise-induced ST depression in the diagnosis of coronary artery disease: A meta-analysis', *Circulation*, **80**, pp. 87–98

HAMPEL, F., RONCHETTI, E., ROUSSEEUW, P., and STAHEL, W. (1986): 'Robust statistics' (John Wiley & Sons, New York, USA)

HERPIN, D., FERRANDIS, J., BORDERON, P., GAUDEAU, B., RAGOT, S., and DEMANGE, J. (1995): 'Comparison of the diagnostic accuracy of different methods of measurement of heart rate-adjusted ST-segment depression during exercise testing for identification of coronary artery disease', *Am. J. Cardiol.*, **76**, pp. 1147–1151

HOKE, M., ROSS, B., WICKESBERG, R., and LÜTKENHÖNER, B. (1984): 'Weighted averaging: theory and application to electric response audiometry', *Electroencephal. Clin. Neurophysiol.*, **57**, pp. 579–584

ISKANDRIAN, A., DECKER, W. V., GUPTA-BALA, S., HEO, J., ACIO, E., and NALLAMOTHU, N. (1997): 'Nuclear cardiac imaging' (Medical Monograph Series, Office of Continuing Medical Education, Drexel University College of Medicine), **7**

JAGER, F., MOODY, G., ANTOLIC, G., MASIC, D., and MARK, R. (1997): 'Sympatho-vagal correlates of transient ischemia in ambulatory patients', in 'Computers in cardiology' (IEEE Computer Society Press, 1997), pp. 387–390

KLIGFIELD, P., and OKIN, P. (1999): 'Heart rate adjustment of ST depression in patients with coronary disease and negative standard exercise test', *J. Electrocardiol.*, **32**, pp. 193–197

LAGUNA, P., JANÉ, R., and CAMINAL, P. (1994): 'Automatic detection of wave boundaries in multilead ECG signals: validation with the CSE database', *Comput. Biomed. Res.*, **27**, pp. 45–60

LEHTINEN, R., SIEVÄNEN, H., VIIK, J., TURJANMAA, V., NIEMELÄ, K., and MALMIVUO, J. (1996): 'Accurate detection of coronary artery disease by integrated analysis of the ST-segment depression/heart rate patterns during the exercise and recovery phases of the exercise ECG test', *Am. J. Cardiol.*, **78**, pp. 1002–1006

LEHTINEN, R. (1999): 'ST/HR hysteresis. Exercise and recovery phase ST depression/heart rate analysis of the exercise ECG', *J. Electrocardiol.*, **32**, pp. 198–203

MATEO, J., and LAGUNA, P. (2000): 'Improved heart rate variability time-domain signal construction from the beat occurrence times according to the IPFM model', *IEEE Trans. Biomed. Eng.*, **47**, pp. 985–996

MATEO, J., SERRANO, P., BAILÓN, R., GARCÍA, J., FERREIRA, A., DEL RIO, A., FERREIRA, I., and LAGUNA, P. (2001*a*): 'Heart rate variability measurements during exercise test may improve the diagnosis of ischemic heart disease'. Proc. 23rd Int. Conf., IEEE-EMBS Society, Istanbul (CD-ROM)

MATEO, J., SERRANO, P., BAILÓN, R., OLMOS, S., GARCÍA, J., DEL RÍO, A., FERREIRA, I., and LAGUNA, P. (2001*b*): 'ECG-based clinical indexes during exercise test including repolarization, depolarization and HRV', in 'Computers in cardiology', vol. 28 (IEEE Computer Society Press, 2001), pp. 309–312

MATEO, J., and LAGUNA, P. (2003): 'Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal', *IEEE Trans. Biomed. Eng.* **50**, pp. 334–343

MICHAELIDES, A., TRIPOSKIADIS, F., BOUDOULAS, H., SPANOS, A., PAPADOPOULOS, P., KOUROUKLIS, K., and TOUTOUZAS, P. (1990): 'New coronary artery disease index based on exercise-induced QRS changes', *Am. Heart J.*, **120**, pp. 292–302

MICHAELIDES, A., RYAN, J., FOSSEN, D. V., POZDERAC, R., and BOUDOULAS, H. (1993): 'Exercise-induced QRS prolongation in patients with coronary artery disease: a marker of myocardial ischemia', *Am. Heart J.*, **126**, pp. 1320–1325

MICHAELIDES, A., PSOMADAKI, Z., AGGELI, K., GEORGIADES, G., PITSAVOS, C., SEFERLIS, C., and TOUTOUZAS, P. (1996): 'Best detection of coronary artery disease and identification of the significantly narrowed coronary artery(ies), using a new technique in exercise testing', *J. Am. Coll. Cardiol.*, **27** p. 129A (932)

MOODY, G., and MARK, R. (1982): 'Development and evaluation of a 2-lead ECG analysis program', in 'Computers in cardiology' (IEEE Computer Society Press, 1982), pp. 39–44

MORISE, A. (1997): 'Accuracy of heart rate-adjusted ST segments in populations with and without posttest referral bias', *Am. Heart J.*, **134**, pp. 647–655

MORISE, A., and DUVAL, R. (1995): 'Diagnostic accuracy of heart rate-adjusted ST segments compared with standard ST-segment criteria', *Am. J. Cardiol.*, **75**, pp. 118–121

MORSHEDI-MEIBODI, A., LARSON, M., LEVY, D., O'DONNELL, C., and VASAN, R. (2002): 'Heart rate recovery after treadmill exercise testing and risk of cardiovascular disease events (The Framingham Heart Study)', *Am. J. Cardiol.*, **90**, pp. 848–852

OKIN, P., AMEISEN, O., and KLIGFIELD, P. (1989): 'Recovery-phase patterns of ST segment depression in the heart rate domain. Identification of coronary artery disease by the rate-recovery loop', *Circulation*, **80**, pp. 533–541

OKIN, P., BERGMAN, G., and KLIGFIELD, P. (1991*a*): 'Effect of ST segment measurement point on performance of standard and heart rate-adjusted ST segment criteria for the identification of coronary artery disease', *Circulation*, **84**, pp. 57–66

OKIN, P., BERGMAN, G., and KLIGFIELD, P. (1991*b*): 'Heart rate adjustment of the time-voltage ST segment integral: identification of coronary disease and relation to standard and heart rate-adjusted ST segment depression criteria', *J. Am. Coll. Cardiol.*, **18**, pp. 1487–1492

OKIN, P., and KLIGFIELD, P. (1995*a*): 'Gender-specific criteria and performance of the exercise electrocardiogram', *Circulation*, **92**, pp. 1209–1216

OKIN, P., and KLIGFIELD, P. (1995*b*): 'Heart rate adjustment of ST segment depression and performance of the exercise electrocardiogram: a critical evaluation', *J. Am. Coll. Cardiol.*, **25**, pp. 1726–1735

TASK FORCE OF ESC AND NASPE, T. (1996): 'Heart rate variability. Standards of measurement, physiological interpretation, and clinical use', *Eur. Heart J.*, **17**, pp. 354–381

TOTH, A., MARTON, Z., CZOPF, L., KESMARKY, G., HALMOSI, R., JURICSKAY, I., HABON, T., and TOTH, K. (2001): 'QRS score: a composite index of exercise-induced changes in the Q, R, and S waves during exercise stress testing in patients with ischemic heart disease', *Ann. Noninvas. Electrocardiol.*, **6**, pp. 310–318

TSUJI, H., VENDITTI, F., MANDERS, E., EVANS, J., LARSON, M., FELDMAN, C., and LEVY, D. (1994): 'Reduced heart rate variability and mortality risk in an elderly cohort. The Framingham Heart Study', *Circulation*, **90**, pp. 878–883

TSUJI, H., LARSON, M., VENDITTI, F., MANDERS, E., EVANS, J., FELDMAN, C., and LEVY, D. (1996): 'Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study', *Circulation*, **94**, pp. 2850–2855

VANCAMPEN, C., VISSER, F., and VISSER, C. (1996): 'The QRS score: a promising new exercise score for detecting coronary artery disease based on exercise-induced changes of Q-, R- and S-waves: a relationship with myocardial ischaemia', *Eur. Heart J.*, **17**, pp. 699–708

WARNIK, W., ROSS, J., and KARALIS, D. (1996): 'Stress echocardiography in clinical practice' (Medical Monograph Series, Office of Continuing Medical Education, Drexel University College of Medicine), **6**

WATANABE, J., THAMILARASAN, M., BLACKSTONE, E., THOMAS, J., and LAUER, M. (2001): 'Heart rate recovery immediately after treadmill exercise and left ventricular systolic dysfunction as predictors of mortality. The case of stress echocardiography', *Circulation*, **104**, pp. 1911–1916.

Author's biography

RAQUEL BAILÓN obtained her degree in Telecommunication Engineering from the University of Zaragoza, Spain, in 2001. Currently, she is a PhD student in the Department of Electronics Engineering and Communications of the same university, supported by a Spanish Government grant. Her main research interests lie in biomedical signal processing, especially in the electrocardiogram analysis for purposes of diagnosis.