Advanced repeated structuring and learning procedure to detect acute myocardial ischemia in serial 12-lead ECGs

Sbrollini, A.; Haar, C.C. ter; Leoni, C.; Morettini, M.; Burattini, L.; Swenne, C.A.

Physiological Measurement

PAPER

OPEN ACCESS

Advanced repeated structuring and learning procedure to detect acute myocardial ischemia in serial 12-lead ECGs

Agnese Sbrollini[1] , C Cato ter Haar[2,3], Chiara Leoni[1], Micaela Morettini[1], Laura Burattini[1] and Cees A Swenne[2,*]

1 Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy
2 Cardiology Department, Leiden University Medical Center, Leiden, the Netherlands
3 Cardiology Department, Amsterdam University Medical Center, Amsterdam, the Netherlands
* Author to whom any correspondence should be addressed

E-mail: a.sbrollini@staff.univpm.it, c.c.terhaar@amsterdammumc.nl, chiara.leoni.940@gmail.com, m.morettini@univpm.it, l.burattini@univpm.it and c.a.swenne@lumc.nl

Abstract

Objectives. Acute myocardial ischemia in the setting of acute coronary syndrome (ACS) may lead to myocardial infarction. Therefore, timely decisions, already in the pre-hospital phase, are crucial to preserving cardiac function as much as possible. Serial electrocardiography, a comparison of the acute electrocardiogram with a previously recorded (reference) ECG of the same patient, aids in identifying ischemia-induced electrocardiographic changes by correcting for interindividual ECG variability. Recently, the combination of deep learning and serial electrocardiography provided promising results in detecting emerging cardiac diseases; thus, the aim of our current study is the application of our novel Advanced Repeated Structuring and Learning Procedure (AdvRS&LP), specifically designed for acute myocardial ischemia detection in the pre-hospital phase by using serial ECG features. *Approach.* Data belong to the SUBTRACT study, which includes 1425 ECG pairs, 194 (14%) ACS patients, and 1035 (73%) controls. Each ECG pair was characterized by 28 serial features that, with sex and age, constituted the inputs of the AdvRS&LP, an automatic constructive procedure for creating supervised neural networks (NN). We created 100 NNs to compensate for statistical fluctuations due to random data divisions of a limited dataset. We compared the performance of the obtained NNs to a logistic regression (LR) procedure and the Glasgow program (Uni-G) in terms of area-under-the-curve (AUC) of the receiver-operating-characteristic curve, sensitivity (SE), and specificity (SP). *Main Results.* NNs (median AUC = 83%, median SE = 77%, and median SP = 89%) presented a statistically (*P* value lower than 0.05) higher testing performance than those presented by LR (median AUC = 80%, median SE = 67%, and median SP = 81%) and by the Uni-G algorithm (median SE = 72% and median SP = 82%). *Significance.* In conclusion, the positive results underscore the value of serial ECG comparison in ischemia detection, and NNs created by AdvRS&LP seem to be reliable tools in terms of generalization and clinical applicability.

1. Introduction

According to the World Health Organization, cardiovascular diseases (CVDs) are the leading cause of death globally (Choudhary *et al* 2021, Ng 2016). An estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths (WHO 2021). The most common CVD is atherosclerosis: the formation of cholesterol-rich plaques within the inner arterial walls. Rupture, ulceration, fissure, or atherosclerotic plaque erosion introduces thrombogenic factors into the bloodstream, resulting in intraluminal thrombus formation. Partial or complete occlusion of the blood vessel by the thrombus causes acute ischemia and subsequent cell death in the vascular bed distal from the occlusion site. When occurring in the coronary circulation, this scenario

is called acute coronary syndrome (ACS) (Ibanez *et al* 2018, Collet *et al* 2021). A myocardial infarction develops within hours when the occlusion does not resolve spontaneously and there is no timely effective intervention (such as percutaneous coronary intervention, coronary artery bypass grafting, antithrombotic or fibrinolytic medication). Myocardial infarction is often detrimental to cardiac pump function and may lead to heart failure and to arrhythmias that may be lethal. Hence, fast and correct triage and immediate access to adequate therapy are essential. E.g. transporting the patient to a proper intervention center may occur in parallel with activating a catheterization room, thus bypassing delays by the formal hospital admission or emergency room procedures (Ibanez *et al* 2018, Collet *et al* 2021).

Other causes of ischemia, ACS, and possible subsequent infarction in which a condition other than coronary plaque instability plays a role, are situations in which an imbalance between myocardial oxygen supply and demand occurs due to, e.g. hypotension, hypertension, tachyarrhythmias, bradyarrhythmias, anemia, hypoxemia, coronary artery spasm, spontaneous coronary artery dissection, coronary embolism, and coronary microvascular dysfunction (Thygesen 2018).

Usually, ischemia-related symptoms like chest discomfort trigger the ACS patient or those surrounding them to call for emergency medical assistance. However, chest discomfort or chest pain-equivalent symptoms like dyspnea, epigastric pain, and pain in the left arm are not uniquely associated with ACS. The list of differential diagnoses includes various cardiac, pulmonary, vascular, gastrointestinal, orthopedic, and psychological conditions, of which several are reasons for hospital admission, and some (e.g. aorta dissection) are emergencies. In practice, only a limited fraction (order of magnitude 10%) of the patients, who are urgently transported to a hospital because of chest discomfort, are in the hospital diagnosed as ACS patients (Cotterill *et al* 2015, Thygesen 2018, Al-Zaiti *et al* 2022, Tsao *et al* 2022).

Upon the arrival of the medical emergency services, an ECG is routinely made, and its interpretation plays an essential role in the triage decision regarding the treatment that the patient should receive. When the sensitivity of the ECG interpretation for acute ischemia/ACS is too low, it leaves patients without adequate treatment; when the specificity is too low, the low prevalence of ACS in the group of patients with chest discomfort causes flooding of emergency medical facilities, and patients are at risk of being subjected to unnecessary invasive diagnostics or risk-bearing treatments. The traditional ECG criteria lack sensitivity for ACS (Al-Zaiti *et al* 2022). Proper medical decision-making at this triage stage thus requires an effective ECG diagnosis with sufficient sensitivity without losing specificity (ter Haar *et al* 2020a).

To diagnose myocardial ischemia, clinical guidelines (O'Gara *et al* 2013, Ibanez *et al* 2018, Collet *et al* 2021) recommend evaluating the ECG of the patient, interpreting its ECG features and, specifically, investigating signs of ST elevation/depression at the J point. Considering that distributions of ECG features related to myocardial ischemia overlap with normal values (Rijnbeek *et al* 2014) and present high intra-individual variability (due to already existing heart disease and related ECG abnormalities like old myocardial infarctions, bundle branch block or hypertrophy), the pre-hospital ECG is not easily interpretable. Serial electrocardiography, i.e. comparing the acute ECG and a previously obtained non-acute ECG of the same patient, can help identify within-subject modifications of ECG features (Sbrollini *et al* 2019) that can be useful to identify myocardial ischemia (ter Haar 2020a, 2020b).

A combination of deep learning, i.e. a family of machine-learning algorithms that imitates human reasoning by using advanced and highly non-linear statistical tools (Mathews 2019, Iman *et al* 2021), and serial electrocardiography provided promising results. Specifically, neural networks (NNs) created by our Repeated Structuring and Learning Procedure (RS&LP) were able to detect emerging cardiac diseases (Marinucci *et al* 2020, Sbrollini *et al* 2018, 2019, 2021, 2023). Thus, the present study aims to present the application of a novel version of the RS&LP, Advanced Repeated Structuring and Learning Procedure (AdvRS&LP), specifically designed for acute myocardial ischemia detection in the pre-hospital phase using uniquely serial ECG features.

2. Methods

2.1. Database

Data were collected during the SUBTRACT study (ter Haar *et al* 2020a), jointly performed by the Amsterdam Medical Center (AMC) and the Leiden University Medical Center (LUMC), a retrospective observational study that aimed to evaluate the diagnostic value of serial electrocardiography for the detection of acute myocardial ischemia in the pre-hospital phase.

The SUBTRACT study protocol was approved by the medical ethical committees (METCs) of the AMC and the LUMC and by the boards of directors of other participating medical centers, based on the two academic METCs' approval (ter Haar *et al* 2020a).

Eligible patients were found by searching the administrative databases of the emergency medical services in consecutive order for patients who needed urgent ambulance transport to one of the participating hospitals

Table 1. Ischemia categories, numbers of patients, and case-control assignments in the study group.

Ischemia category	Number (%) of patients	Cases/controls
Presumably ischemic	166 (12%)	Cases
Probably ischemic	28 (2%)	Cases
Uncertain	196 (13%)	Excluded
Probably non-ischemic	66 (5%)	Controls
Presumably non-ischemic	969 (68%)	Controls
	1425 (100%)	

because of chest discomfort or chest pain-equivalent symptoms and in whom at least one ambulance ECG of sufficient technical quality had been made. Ambulance ECGs with poor signal quality, without a regular supraventricular rhythm or with atrial flutter, or that could not be processed by the University of Glasgow (Uni-G) ECG Analysis Program (Macfarlane *et al* 2005), e.g. in case of suspected lead interchange, were not analyzed. When multiple analyzable ECGs of one ambulance ride were available, the first ECG was selected for further analysis.

Then, we checked the administrative and ECG databases of all cooperating hospitals in the Leiden and Amsterdam regions for a previous ECG recording of the same patient that had electively been made (hence, not in an acute state).

Patients were excluded when no earlier elective ECG could be found and in case of insufficient information, e.g. due to death before a reliable diagnosis could be established. Patients with a major cardiac event, e.g. open-heart surgery or myocardial infarction, between the time instants at which the reference and ambulance ECGs were recorded, were also excluded. Finally, if a patient had multiple ambulance visits by the emergency medical services during the study period, only data regarding the most recent visit were included.

A total of 1425 patients were included (736 males and 689 females; 67 ± 14 years). See ter Haar *et al* (2020a) for the clinical characteristics of these patients (ter Haar *et al* 2020a). For each patient, two 10-s 12-lead ECGs were available for analysis, one ambulance ECG (AECG) made in the acute situation and one electively made prior ECG that functions as a reference ECG (RECG). The AECG was recorded with LIFEPAK 12 (Physio-Control) using the Mason-Likar electrode configuration (Mason and Likar 1966). The RECGs were, depending on the hospital in which they were made, recorded by various electrocardiographs (GE, Schiller, Mortara, Siemens/Dräger) using the standard electrode configuration of the diagnostic resting ECG.

For each patient, the presence or absence of myocardial ischemia at the moment of the AECG recording was assessed by evaluating all available clinical data (i.e. laboratory values and clinical diagnosis) obtained from the medical records of the hospital into which the patient was admitted after the ambulance ride (ter Haar *et al* 2020a). As a result of this assessment, each AECG was assigned to one of five categories, namely presumably ischemic, probably ischemic, uncertain, probably non-ischemic, and presumably non-ischemic. In the current study, patients classified as presumably or probably ischemic were considered cases, while patients classified as presumably or probably non-ischemic were considered controls. Patients classified as uncertain were excluded from the analysis. The resulting study group comprised 194 (14%) cases and 1035 (73%) controls. Table 1 gives a breakdown of the ischemia categories, numbers of patients, and case-control assignments in the study group.

2.2. ECG analysis and feature extraction

AECGs were mathematically converted to comply with the standard 12-lead electrode configuration (Man *et al* 2008). Then, the AECGs and the RECGs were processed by the LEADS program (Draisma *et al* 2005). Briefly, LEADS first converts the standard 12-lead ECG into a vectorcardiogram using the Kors matrix (Draisma *et al* 2005). Then, LEADS detects the heartbeats in the recording, computes the averaged predominant beat, and automatically identifies its QRS-complex onset, J-point, and T-wave offset. These landmarks in time were visually inspected and, if necessary, manually corrected by two independent clinicians. Finally, each ECG pair was characterized by the 28 serial features, listed in table 2. Figure 1 gives an example of serial feature extraction.

2.3. Advanced repeated structuring and learning procedure

The AdvRS&LP is an improved version of the RS&LP (Sbrollini *et al* 2019, 2021). The procedure requires that the database is divided into a learning dataset (in our current study, 70% of the entire database; 141 cases and 720 controls) and a testing dataset (in our current study, 30% of the entire database; 53 cases and 315 controls). Moreover, the learning dataset has to be further divided into a training dataset (in our current study, 80% of the learning dataset; 113 cases and 576 controls) and a validation dataset (in our current study, 20% of the learning

Table 2. Serial features computed from each reference ECG (RECG)—ambulance ECG (AECG) pair.
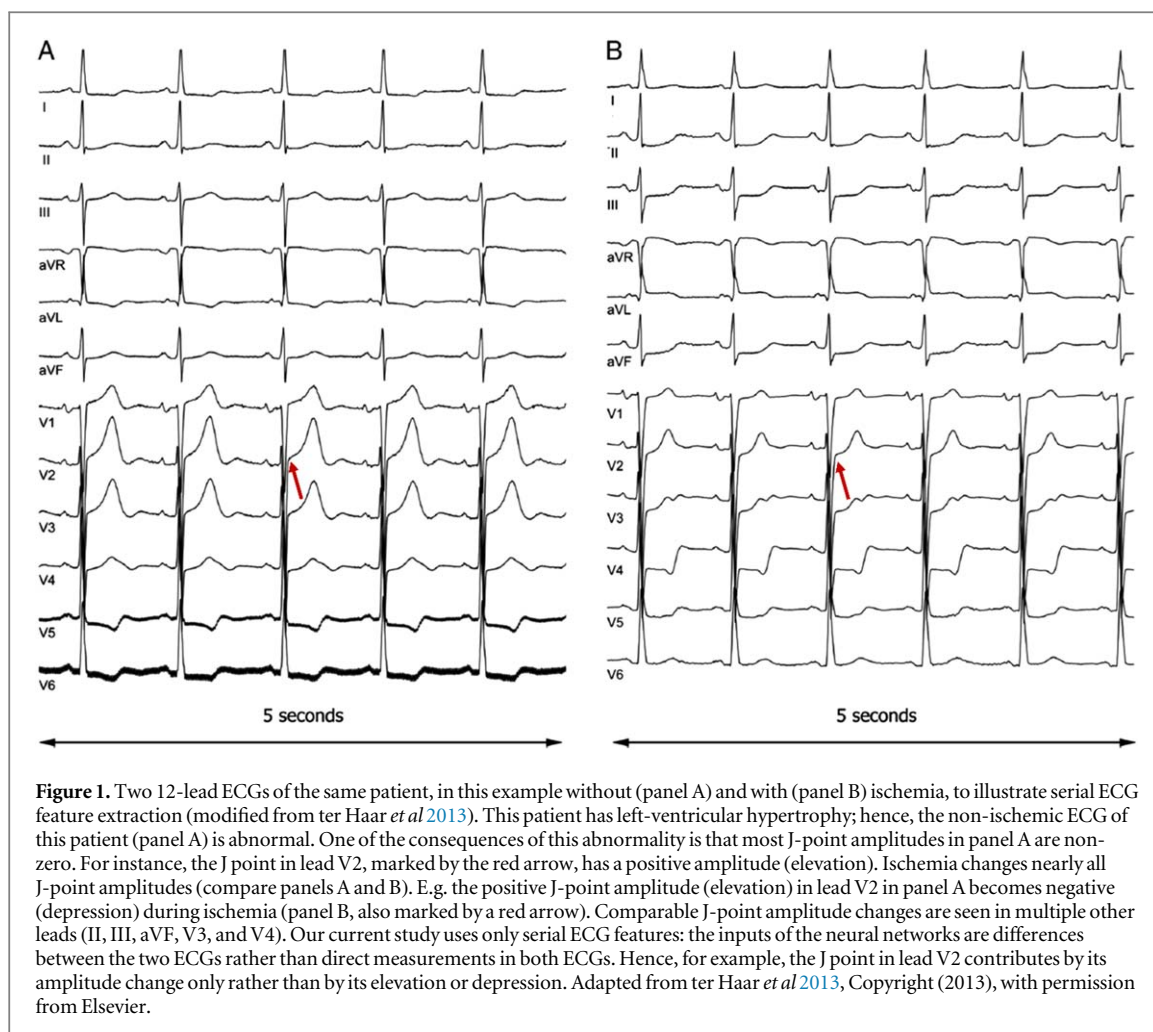
	Acronym (unit)	Description
1	QRSD (ms)	The difference between the QRS durations of the AECG and RECG
2	QRSaD (ms)	The absolute value of the difference between the QRS durations of the AECG and RECG
3	MQRSD (μV)	The difference between the magnitudes of the maximal QRS vectors of the AECG and RECG
4	MQRSaD (μV)	The absolute value of the difference between the magnitudes of the maximal QRS vectors of the AECG and RECG
5	IQRSD (mV·ms)	The difference between the magnitudes of the QRS-integral vectors of the AECG and RECG
6	IQRSaD (mV·ms)	The absolute value of the difference between the magnitudes of the QRS-integral vectors of the AECG and RECG
7	QRSCD (%)	The difference between the QRS complexities of the AECG and RECG
8	QRSCaD (%)	The absolute value of the differences between the QRS complexities of the AECG and RECG
9	JD (μV)	The absolute value of the difference between the magnitudes of the J-point vectors of the AECG and RECG
10	SDJ8 (μV)	The summed absolute values of the relative J-point shifts considering the 8 independent leads
11	SDJ12 (μV)	The summed absolute values of the relative J-point shifts considering the 12 standard leads
12	MTD (μV)	The difference between the magnitudes of the maximal T vectors of the AECG and RECG
13	MTaD (μV)	The absolute value of the difference between the magnitudes of the maximal T vectors of the AECG and RECG
14	ITD (mV·ms)	The difference between the magnitudes of the T-integral vectors of the AECG and RECG
15	ITaD (mV·ms)	The absolute value of the difference between the magnitudes of the T-integral vectors of the AECG and RECG
16	TCD (%)	The difference between the T-wave complexities of the AECG and RECG
17	TCaD (%)	The absolute value of the difference between the T-wave complexities of the AECG and RECG
18	TSD (%)	The difference between the T-wave symmetries of the AECG and the RECG
19	TSaD (%)	The absolute value of the difference between T-wave symmetries of the AECG and RECG
20	NPTD (adi)	The difference between the number of leads with positive T waves in the AECG and RECG
21	NTPC (adi)	The number of leads that present a T-wave polarity change between AECG and RECG
22	QTD (ms)	The difference between the QT intervals of the AECG and RECG
23	QTaD (ms)	The absolute value of the difference between the QT intervals of the AECG and RECG
24	VGD (μV)	The difference between the magnitudes of the ventricular gradients of the AECG and RECG
25	SAD (°)	The difference between the QRS-T spatial angles of the AECG and RECG
26	SAaD (°)	The absolute value of the difference between the QRS-T spatial angles of the AECG and RECG
27	HRD (bpm)	The difference between the heart rates of the AECG and RECG
28	HRaD (bpm)	The absolute value of the difference between the heart rates of the AECG and RECG.

dataset; 28 cases and 144 controls). Finally, the procedure requires that the case-control prevalence ratio is maintained in all datasets.

AdvRS&LP creates supervised deep learning NNs with the number of inputs (IN) equal to the number of features (in this study, IN equals 30, namely the 28 serial features plus sex and age), and the number of outputs (OUT) equal to the number of classes (in this study, OUT equals 2, namely the case and control classes). Panel A of figure 2 shows a block scheme of the AdvRS&LP. It is an iterative procedure that facilitates a gradual structural and performance improvement of a NN (initially composed of one neuron in one hidden layer) during learning.

Specifically, during each structuring cycle, the actual NN is extended by adding additional (AD) neurons in an existing hidden layer or by creating an additional hidden layer consisting of AD neurons, which number is initially equal to the number of inputs. Only AD neurons are initialized by random weights and biases (ranging between -1 and $+1$ and characterized by a sigmoid activation function), while the weights and biases of the neurons composing the actual NN remain unchanged. The initialization is considered suitable only if the performance of the initialized NN is better than that of the actual NN after one single epoch; otherwise, the current initialization of the AD neurons is rejected. Initialized NNs are trained by the scaled-conjugate-gradients algorithm (Møller 1993), and the early stopping criterion prevents overfitting (Prechelt 2012). Moreover, we considered the inverse of class prevalence as input to compensate for class disproportionality (King and Zeng 2003). The performance of all trained NNs is compared with the performance of the actual NN, and the NN with the lowest validation error is considered the best. If the best NN is not equivalent to the actual NN, the best NN is promoted as the actual NN, and the procedure starts anew (example in figure 2, panel B). Otherwise, if the best NN corresponds to the actual NN, the actual NN is maintained unchanged, the number of AD neurons decreases with one neuron, and the procedure starts anew considering the modified number of AD neurons (example in figure 2, panel C). The procedure stops if none of the structures is suitable after initialization (example in figure 2, panel D) or if the number of AD neurons is equal to zero (example in figure 2, panel E).

The AdvRS&LP was repeated 100 times to prevent overfitting, constructing 100 final NNs. The NN with the highest area under the curve (AUC) of the receiver operating characteristic (ROC) in the learning dataset was

Figure 1. Two 12-lead ECGs of the same patient, in this example without (panel A) and with (panel B) ischemia, to illustrate serial ECG feature extraction (modified from ter Haar *et al* 2013). This patient has left-ventricular hypertrophy; hence, the non-ischemic ECG of this patient (panel A) is abnormal. One of the consequences of this abnormality is that most J-point amplitudes in panel A are non-zero. For instance, the J point in lead V2, marked by the red arrow, has a positive amplitude (elevation). Ischemia changes nearly all J-point amplitudes (compare panels A and B). E.g. the positive J-point amplitude (elevation) in lead V2 in panel A becomes negative (depression) during ischemia (panel B, also marked by a red arrow). Comparable J-point amplitude changes are seen in multiple other leads (II, III, aVF, V3, and V4). Our current study uses only serial ECG features: the inputs of the neural networks are differences between the two ECGs rather than direct measurements in both ECGs. Hence, for example, the J point in lead V2 contributes by its amplitude change only rather than by its elevation or depression. Adapted from ter Haar *et al* 2013, Copyright (2013), with permission from Elsevier.

considered the optimal NN. We created 100 NNs to compensate for statistical fluctuations due to random data divisions of a limited dataset.

2.4. Statistical analysis

All computations were performed in Matlab (Mathworks, version R2022b) and in RStudio (RStudio, version 2021.09.2).

Univariate AUC and its 95% confidence intervals (95% CI) of the ROC of each serial feature were computed, considering the entire database and the learning and testing datasets. The serial feature that provided the highest learning univariate AUC was considered the best serial feature.

The performance of the NNs was assessed by computing the ROC of both learning and testing datasets and, accordingly, the AUC and 95% CI. Finally, the value of sensitivity (SE) and the value of specificity (SP) were selected on the ROC. NNs performance was compared with the results obtained with the logistic regression procedure (LR) used by ter Haar *et al* (2020a) and with case/control classification obtained by the Uni-G algorithm (Macfarlane *et al* 2005).

Distributions of AUC obtained from the NNs were compared with those obtained by LR. Finally, SE and SP distributions obtained with the NNs were compared to those obtained by LR and by the Uni-G algorithm. Distributions are reported as 50th [25th; 75th] percentiles and statistical comparisons were performed by the Wilcoxon rank-sum test. The statistical significance (*P* value) was 0.05.

3. Results

Table 3 reports the AUC and 95% confidence intervals of the univariate serial feature classification in the entire database and the AUC distributions of the univariate serial feature classifications in 100 learning and testing dataset realizations. Feature SDJ12 provided the highest univariate AUC in the entire database (83%) and also in both the learning (82%) and testing datasets (84%).
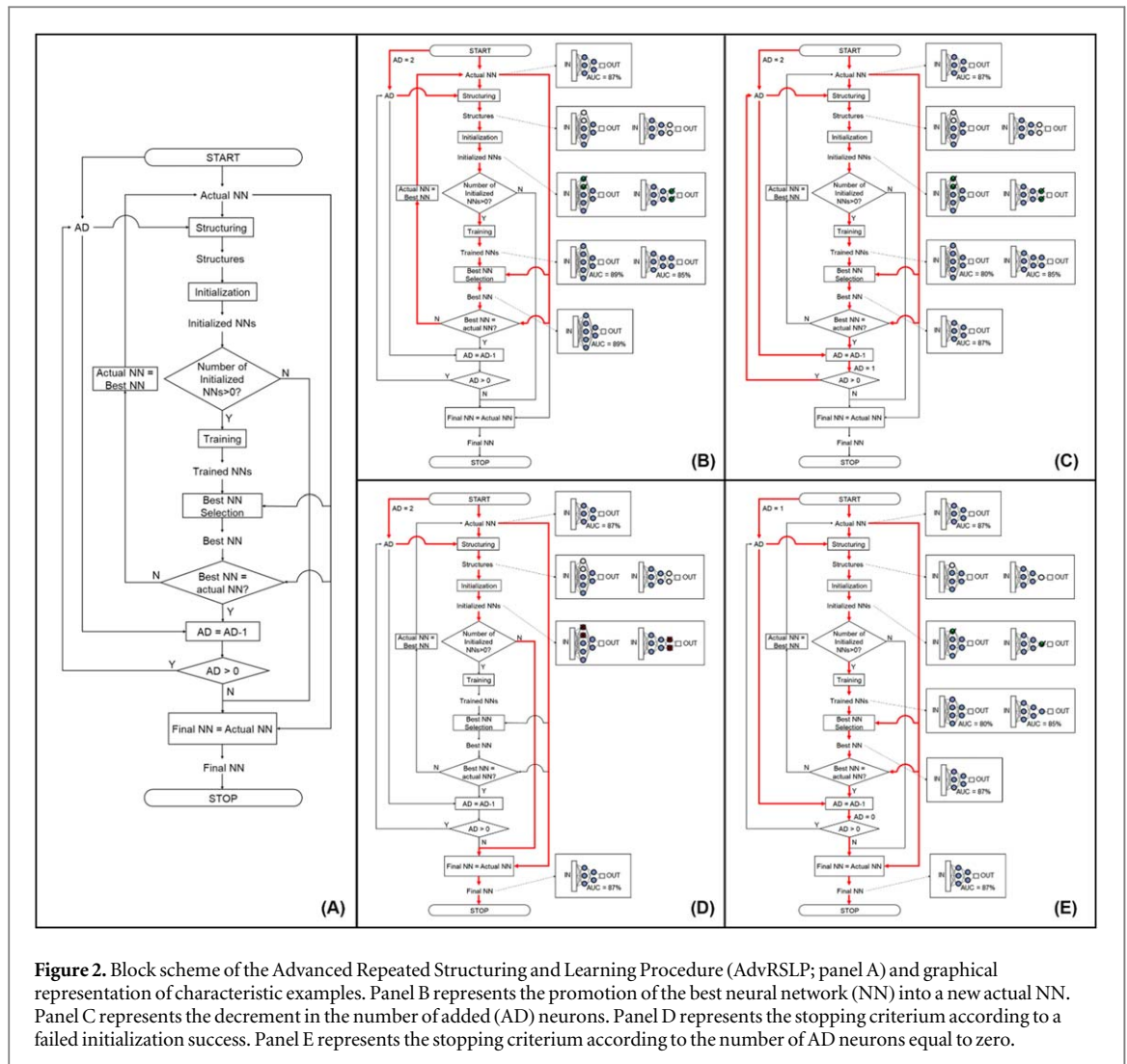
Figure 2. Block scheme of the Advanced Repeated Structuring and Learning Procedure (AdvRSLP; panel A) and graphical representation of characteristic examples. Panel B represents the promotion of the best neural network (NN) into a new actual NN. Panel C represents the decrement in the number of added (AD) neurons. Panel D represents the stopping criterium according to a failed initialization success. Panel E represents the stopping criterium according to the number of AD neurons equal to zero.

Table 4 summarizes the performances of the NNs, the LR, and the Uni-G algorithm. Figure 3 depicts the ROC curves of the NNs (panel A) and the LR (panel B) for testing datasets. Red dots represent the SE and SP of the Uni-G algorithm; median SE and median SP are represented by green squares. NN architectures are variable: the least complex architecture presents 47 neurons in one hidden layer, and the most complex architecture shows four hidden layers with 61 neurons in the first layer, 28 neurons in the second layer, 113 neurons in the third layer, and 33 neurons in the fourth layer. NNs performed statistically better than LR in both learning (NNs: median AUC = 90%; LR: median AUC = 87%; *P* value < 0.05) and testing (NNs: median AUC = 83%; LR: median AUC = 80%; *P* value < 0.05) as reported in table 4. The NNs (learning: median SE = 91% and median SP = 91%; testing: SE = 77% and SP = 89%) also had a significantly better performance than LR (learning: median SE = 67% and median SP = 81%; testing: median SE = 67% and median SP = 81%) and Uni-G algorithm (learning: median SE = 65% and median SP = 80%; testing: median SE = 72% and median SP = 82%) in terms of SE and SP in both the learning and testing datasets (table 4).

4. Discussion

In patients presenting with chest pain, reliable ECG-based detection of acute myocardial ischemia at first medical contact is crucial for the correct triage, which is essential for optimal preservation of cardiac function. This paper presents the Advanced Repeated Stucturing and Learning Procedure (an improved version of the Repeated Stucturing and Learning Procedure—Sbrollini *et al* 2019, 2021) as a deep learning method able to handle the detection of acute myocardial ischemia in ambulance ECGs by serial comparison with a previous non-ischemic reference ECG of the same patient.

For this study, we used the data collected during the multicenter SUBTRACT study (ter Haar *et al* 2020a). These data are rare and valuable because of the following reasons:

Table 3. The area under the curve (AUC) and 95% confidence intervals of the univariate serial feature classification in the database and AUC distributions (50th [25th; 75th] percentiles) of the univariate serial feature classifications in 100 learning and testing dataset realizations.

		Entire database	Learning datasets	Testing datasets
	Cases/controls	194/1035	141/720	53/315
1	*QRSD* (ms)	61 [57; 66]	61 [56; 66]	63 [58; 68]
2	*QRSaD* (ms)	50 [46; 54]	47 [42; 52]	57 [52; 62]
3	*MQRSD* (μV)	47 [43; 52]	48 [43; 53]	45 [40; 50]
4	*MQRSaD* (μV)	55 [51; 60]	55 [50; 60]	56 [51; 62]
5	*IQRSD* (mV·ms)	56 [52; 61]	58 [53; 64]	51 [46; 57]
6	*IQRSaD* (mV·ms)	57 [53; 62]	56 [51; 61]	60 [55; 65]
7	*QRSCD* (%)	62 [58; 67]	62 [57; 68]	62 [57; 68]
8	*QRSCaD* (%)	62 [57; 66]	61 [55; 66]	65 [59; 70]
9	*JD* (μV)	80 [76;84]	79 [74; 84]	81 [76; 85]
10	*SDJ8* (μV)	82 [78; 86]	81 [76; 85]	85 [81; 89]
11	*SDJ12* (μV)	83 [79; 86]	82 [78; 87]	84 [80; 88]
12	*MTD* (μV)	59 [55; 64]	60 [54; 65]	59 [54; 65]
13	*MTaD* (μV)	62 [57; 66]	61 [55; 66]	63 [58;69]
14	*ITD* (mV·ms)	58 [53; 63]	56 [51; 62]	62 [57; 68]
15	*ITaD* (mV·ms)	59 [55; 64]	59 [54; 65]	59 [54; 65]
16	*TCD* (%)	53 [43; 51]	49 [44; 54]	43 [38; 49]
17	*TCaD* (%)	55 [51; 60]	55 [49; 60]	56 [50; 61]
18	*TSD* (%)	46 [41; 50]	44 [39; 49]	50 [45; 55]
19	*TSaD* (%)	58 [53;62]	58 [53; 64]	56 [51; 62]
20	*NPTD* (adi)	43 [39; 48]	43 [38; 48]	44 [39; 49]
21	*NTPC* (adi)	62 [57; 66]	62 [56; 67]	62 [57; 68]
22	*QTD* (ms)	45 [41; 50]	45 [40; 50]	46 [41; 51]
23	*QTaD* (ms)	57 [52; 61]	56 [51; 62]	57 [52; 63]
24	*VGD* (μV)	64 [60; 69]	66 [60; 71]	62 [56; 67]
25	*SAD* (°)	53 [49; 58]	52 [47; 58]	55 [50; 61]
26	*SAaD* (°)	60 [56; 65]	61 [56; 67]	58 [52; 63]
27	*HRD* (bpm)	55 [50; 59]	56 [50; 61]	54 [48; 59]
28	*HRaD* (bpm)	56 [51; 60]	56 [51; 61]	56 [50; 61]

Table 4. Performance of the neural networks (NN), the logistic regression (LR), and the Uni-G algorithm. The areas under the curve (AUC) are reported as median and 95% confidence intervals (95% CI). Distributions of sensitivity (SE) and specificity (SP) are reported as 50th [25th; 75th] percentiles.
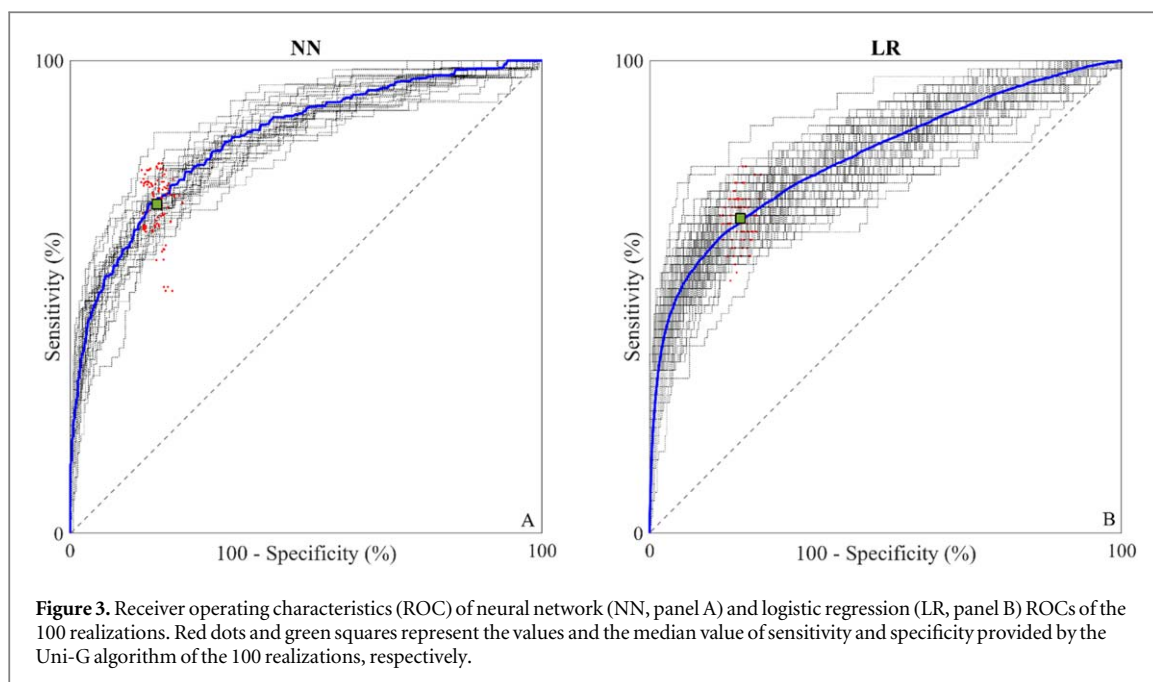
		NN	LR	Uni-G
Learning dataset	AUC [95% CI] (%)	90[a] [88; 91]	87 [86; 87]	—
	SE (%)	91[a,b] [84; 93]	67 [66; 68]	65 [64; 68]
	SP (%)	91[a,b] [88; 92]	81 [66; 68]	80 [79; 81]
Testing dataset	AUC [95% CI] (%)	83[a] [82; 85]	80 [78; 82]	—
	SE (%)	77[a,b] [73; 84]	67 [64; 69]	72 [65; 74]
	SP (%)	89[a,b] [88; 90]	81 [80; 82]	82 [80; 83]

[a] *P* value <0.05 when comparing the distributions of NNs with those obtained by LR.
[b] *P* value <0.05 when comparing the distributions of NNs with those obtained by the Uni-G algorithm.

- These data could only be collected because of the joint participation of ambulance services, academic and non-academic hospitals in the Leiden and Amsterdam regions, who granted access to their administrative, clinical, and ECG databases.

- For each patient, the database contains an acute ambulance ECG as well as a previously electively recorded ECG made under non-ischemic conditions, often found in another hospital than the facility to which the patient was acutely admitted after the ambulance ride.

Figure 3. Receiver operating characteristics (ROC) of neural network (NN, panel A) and logistic regression (LR, panel B) ROCs of the 100 realizations. Red dots and green squares represent the values and the median value of sensitivity and specificity provided by the Uni-G algorithm of the 100 realizations, respectively.

- For each patient, detailed and complete clinical information is available about the hospital admission period following ambulance transportation. Based on these findings, a nuanced categorization of patients could be made as either a case patient (presumably or probably ischemic during the ambulance ECG recording), a control patient (presumably or probably non-ischemic during the ambulance ECG recording), or undecided (lack of arguments to decide whether ischemia was present during the exact moment of the recording of the ambulance ECG or not).

- The collected data are representative of the patient mix that is contacting the EMS because of chest pain and in whom ACS has to be ruled in or ruled out without any filtering by specific ECG criteria (e.g. STEMI criteria) for the case patients and with overwhelmingly more control patients, partly healthy, partly with various pathologies, sometimes even interfering with their ECG.

In ACS, the ECG can dynamically change with time, which could contribute to the diagnosis; thus, ECG monitoring would be helpful in this respect. However, current clinical practice is to make 10 s ECGs, which is sufficient to catch the situation of the moment. Sometimes, multiple 10-s ECGs are made during one ambulance ride. In combination with the symptoms during these ECG recordings, we could have attempted to select the presumably most diagnostic ECG. However, no detailed information about the patient's symptoms during the ECG recordings was available in our retrospective observational study. Hence, in patients with more than a single ambulance ECG, we decided to analyze the first one, assuming that this ECG is the most decisive in the triage process and the decision to which healthcare facility the patient has to be transported.

The cardinalities of the classes (cases and controls) are strongly unbalanced. Machine learning algorithms, such as NNs and LR, usually get the inverse of class prevalence as additional input to compensate for this disproportionality (King and Zeng 2003). This is current practice in healthcare applications, where the numbers of cases are often less than the number of controls. In our present study, we have applied the same strategy, but in future projects, we aim to create a procedure independent of cardinality imbalance.

To facilitate the clinical interpretability of the results, our procedure used serial ECG features that were selected because they have specific physiological or pathophysiological evidence. Moreover, serial features were chosen considering the recent outcomes of serial ECG analysis with LR (ter Haar *et al* 2020a), which revealed its usefulness for the detection of acute cardiac ischemia. The best serial feature is SDJ12, the sum of the absolute values of the AECG versus RECG J-amplitude displacements in all 12 leads. Indeed, this feature is in accordance with the clinical guidelines (O'Gara *et al* 2013, Levine *et al* 2016, Ibanez *et al* 2018). However, considering the results of the comparison between our NN and the best serial feature, electrocardiographic myocardial ischemia detection cannot be performed by using only J-point features.

Our new method is an advanced version of the RS&LP, an algorithm for NN construction (Marinucci *et al* 2020, Sbrollini *et al* 2018, 2019, 2021, 2023). Initially, the method added one neuron for each epoch, limiting the NN construction procedure. First of all, adding one neuron at each epoch substantially increases the computational time of the learning phase, which may take days to create a single NN (Sbrollini *et al* 2023).

Moreover, adding one neuron at a time may limit the training performance. Indeed, let us consider the hyperplane of solutions (with the principal components as inputs). The training procedure may guide the NN towards a local minimum in case of an insufficient number of epochs or an insufficient number of neurons in an insufficient number of layers. Considering that the early-stopping validation criterion is efficient in the evaluation of the number of epochs, an ideal constructive algorithm must explore as many NN architectures as possible to reach the global minimum in the hyperplane of solutions. For these reasons, AdvRS&LP presents a variable number of added neurons that varies with the NN learning performance: initially, the number of added neurons after each learning-and-validation cycle equals the number of features to span the entire hyperplane of solutions, while it decreases when the performance of the NN approaching stability. At the end of the procedure, a final low number of added neurons refines the NN performance.

The second main difference between the AdvRS&LP and the original RS&LP is the decision to remove some structural rules. Originally, the RS&LP created NN that should have a pyramidal structure (the number of neurons in a given layer must be equal to or lower than the number of neurons in the previous layer), and the number of layers should be lower than four. Differently, AdvRS&LP promotes the growth of each layer independently from the others and does not limit the number of layers to three. Despite the high freedom in structuring, the AdvRS&LP is limited during the initialization of newly added neurons by early-stopping validation criterion and by selecting the best NN according to the validation error to prevent generalization loss. Our results confirmed the utility of the generalization rules applied to AdvRS&LP. Indeed, the learning performance is not very high (median AUC on the learning dataset close to 90%) due to these rules, but it guarantees the generalization of the NN (median AUC on the testing dataset of 83%).

Compared with LR and with the Uni-G algorithm (table 4), the NNs created by the AdvRS&LP yielded the best performance: they provided statistically significantly higher AUC, SE, and SP values in both the learning and testing datasets. The results of NNs and Uni-G algorithm were computed on the same data divisions, while the results of LR were computed on different random data divisions (of the same database, however) because the random data divisions used in the study by ter Haar *et al* (2020a) could not be replicated. However, considering the high number of data divisions (100 realizations), the comparison is statistically reliable.

The Uni-G algorithm yielded a median SE of 72% and a median SP of 82%. This was the result of the analysis of the ambulance ECGs alone. It is a result that is expected not to differ too much from that of an experienced cardiologist (Willems *et al* 1991). Our current study, combining deep learning and serial ECG analysis, yields an SE of 77% and an SP of 89%: a statistically better performance, particularly better SE. Likely, this improvement is partly because of the artificial intelligence-based approach of our current method and partly because of the serial comparison with the reference ECG. Having statistically better results in the testing dataset confirms the superiority of the NN created by AdvRS&LP in terms of generalization and clinical applicability.

This study aimed to diagnose ischemia by analyzing differences between the non-ischemic reference ECG and the ambulance ECG of EMS patients only. The positive results underscore the value of serial comparison of ECG in ischemia detection. Of note, we have not used the direct features of the ambulance ECG that are normally used in the first place to detect ischemia. In the following studies, we will add this information to obtain the maximal possible performance of the current AdvRS&LP algorithm on the basis of an acute ambulance ECG and a historical reference ECG.

Limitations of the clinical application of this method in the real clinical scenario have to be found in the availability of a previous ECG of the same patient that can be used as a reference. We expect that new services, such as a unique accessible cloud storage service for ECG recordings of a region, will facilitate the fast accessibility of the data, the identification of a reference ECG of the patient, and the possibility of using our NN in real-life clinical scenarios.

Data availability statement

The data cannot be made publicly available upon publication due to legal restrictions preventing unrestricted public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Ethical statement

The SUBTRACT study protocol was approved by the medical-ethical committees (MECs) of the AMC (date of approval September 22, 2016, MEC file ID W 16_290 # 16.344) and the LUMC (date of approval February 15, 2017, MEC file ID G16.116), and by the boards of directors of the other participating hospitals and emergency medical services (ter Haar *et al* 2020a), based on the two academic hospitals METCs' approval.

ORCID iDs

Agnese Sbrollini ⬤ https://orcid.org/0000-0002-9152-7216

References

Al-Zaiti S, Macleod R, van Dam P, Smith S W and Birnbaum Y 2022 Emerging ECG methods for acute coronary syndrome detection: Recommendations & future opportunities *J. Electrocardiol.* **74** 65–72

Choudhary J, Chiu S, Bhugra P, Bikdeli B, Supariwala A, Jauhar R and Chatterjee S 2021 Clinical implications of the ISCHEMIA trial: Invasive versus conservative approach in stable coronary disease *Curr. Cardiol. Rep.* **23** 1–9

Collet J P *et al* 2021 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation *Eur. Heart J.* **42** 1289–367

Cotterill P G, Deb P, Shrank W H and Pines J M 2015 Variation in chest pain emergency department admission rates and acute myocardial infarction and death within 30 days in the Medicare population *Acad. Emergency Med.* **22** 955–64

Draisma H H M, Swenne C A, van de Vooren H, Maan A C, Hooft van Huysduynen B, van der Wall E E and Schalij M J 2005 LEADS: An interactive research oriented ECG/VCG analysis system *Comput. Cardiol.* **32** 515–8

Ibanez B *et al* 2018 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC) *Eur. Heart J.* **39** 119–77

Iman M, Arabnia H R and Branchinst R M 2021 Pathways to Artificial General Intelligence: A brief overview of developments and ethical issues via artificial intelligence, machine learning, deep learning, and data science *Adv. Artif. Intell. Appl. Cogn. Comput.* **1** 73–87

King G and Zeng L 2003 Logistic regression in rare events data *J. Stat. Software* **8** 1–27

Levine G N *et al* 2016 2016 ACC/AHA Guideline Focused Update on Duration of Dual Antiplatelet Therapy in Patients With Coronary Artery Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines: An Update of the 2011 ACCF/AHA/SCAI Guideline for Percutaneous Coronary Intervention, 2011 ACCF/AHA Guideline for Coronary Artery Bypass Graft Surgery, 2012 ACC/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease, 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction, 2014 AHA/ACC Guideline for the Management of Patients With Non–ST-Elevation Acute Coronary Syndromes, and 2014 ACC/AHA Guideline on Perioperative Cardiovascular Evaluation and Management of Patients Undergoing Noncardiac Surgery *Circulation* **134** e123–55

Macfarlane P W, Devine B and Clark E 2005 The University of Glasgow (Uni-G) ECG analysis program *Comput. Cardiol.* **32** 451–4

Man S-C, Maan A C, Kim E, Draisma H H M, Schalij M J, van der Wall E E and Swenne C A 2008 Reconstruction of standard 12-lead electrocardiograms from 12-lead electrocardiograms recorded with the Mason-Likar electrode configuration *J. Electrocardiol.* **41** 211–9

Marinucci D, Sbrollini A, Marcantoni I, Morettini M, Swenne C A and Burattini L 2020 Artificial neural network for atrial fibrillation identification in portable devices *Sensors* **20** 1–14

Mason R E and Likar I 1966 A new system of multiple-lead exercise electrocardiography *Am. Heart J.* **71** 196–205

Mathews S M 2019 Explainable artificial intelligence applications in NLP, biomedical, and malware classification: A literature review *Intelligent Computing-proc. of the Computing Conf.* pp 1269–92

Møller M F 1993 A scaled conjugate gradient algorithm for fast supervised learning *Neural Netw.* **6** 525–33

Ng K, Steinhubl S R, DeFilippi C, Dey S and Stewart W F 2016 Early detection of heart failure using electronic health records *Circ.: Cardiovascular Quality Outcomes* **9** 649–58

O'Gara P T *et al* 2013 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines *Circulation* **127** e362–425

Prechelt L 2012 Early stopping—but when? *Lecture Notes in Computer Science* (Berlin, Heidelberg: Springer) (*Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*) vol 7700, pp 53–67

Rijnbeek P R *et al* 2014 Normal values of the electrocardiogram for ages 16–90 years *J. Electrocardiol.* **47** 914–21

Sbrollini A *et al* 2019 Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: A deep-learning approach *Biomed. Eng. Online* **18** 1–17 15

Sbrollini A, Barocci M, Mancinelli M, Paris M, Raffaelli S, Marcantoni I, Morettini M, Swenne C A and Burattini L 2023 Automatic diagnosis of newly emerged heart failure from serial electrocardiography by repeated structuring & learning procedure *Biomed. Signal Process. Control* **79** 104185

Sbrollini A, de Jongh M, Cato ter Haar C W, Treskes R, Man S, Burattini L and Swenne C A 2018 Serial ECG analysis: Absolute rather than signed changes in the spatial QRS-T angle should be used to detect emerging cardiac *Pathol. Comput. Cardiol.* **45** 1–4

Sbrollini A, Marcantoni I, Morettini M, Swenne C A and Burattini L 2021 Repeated structuring & learning procedure for detection of myocardial ischemia: A robustness analysis *2021 43rd Annual Int. Conf. of the IEEE Engineering in Medicine & Biology Society (EMBC)* pp 467–70

ter Haar C C *et al* 2020a An initial exploration of subtraction electrocardiography to detect myocardial ischemia in the pre-hospital setting *Ann. Noninvasive Electrocardiol.* **25** e12722

ter Haar C C *et al* 2020b Prevalence of ECGs exceeding thresholds for st-segment-elevation myocardial infarction in apparently healthy individuals: The role of ethnicity *J. Am. Heart Assoc.* **9** e015477

ter Haar C C, Maan A C, Warren S G, Ringborn M, Horáček B M, Schalij M J and Swenne C A 2013 Difference vectors to describe dynamics of the ST segment and the ventricular gradient in acute ischemia *J. Electrocardiol.* **46** 302–11

Thygesen K, Alpert J S, Jaffe A S, Chaitman B R, Bax J J, Morrow D A and White H D 2018 Fourth universal definition of myocardial infarction *Eur. Heart J.* **40** 237–69

Tsao C W *et al* 2022 Heart disease and stroke statistics - 2022 Update: A report from the American Heart Association *Circulation* **145** e153–639

Willems J L *et al* 1991 The diagnostic performance of computer programs for the interpretation of electrocardiograms *New Engl. J. Med.* **325** 1767–73

World Health Organization (WHO) 2021 Cardiovascular diseases (CVDs) www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds); consulted on 17th July 2023