Chapter 5
Advances in the Analysis of Electrocardiogram in Context of Mass Screening: Technological Trends and Application of AI Anomaly Detection

Illya Chaikovsky and Anton Popov

Abstract Electrocardiography is still the most wide-spread method of functional diagnosis. The chapter has been targeted towards the debate on evolution and current attitude on the heart failure screening electrocardiography, reviewing the clinical practices of applying remote electrocardiogram (ECG) recording gadgets, the quantity and origin of data possible to be collected with ECG gadgets having various number of sensors using different modern methods of mathematical transformation of ECG signal, i.e. fourth generation ECG analysis. Accent has been made towards the application of the modern machine learning method – anomaly detection to heart activity analysis. Anomaly detection is one of the machine learning methods which identifies the data samples who deviate from some concept of normality. Such samples represent novelty, or outliers in the dataset, and often carry important information. As an example of application of anomaly detection in biomedical signal analysis, the problem of identifying the subtle deviations from the population norm based on the ECG is presented. The time-magnitude features derived from six leads of Signal Averaged ECG are used in the Isolation Forest anomaly (IFA) detector to quantify the distance of the single ECG from the cluster of normal controls. Input data to the IFA technique consists of diverse tree amounts as well as several pollution factors. For comparison, five different groups were examined: patients with proven coronary artery diseases, military personnel with mine-explosive injuries, COVID-19 survivors, and two subgroups involving participants of widespread-screening in one of the countryside areas in Ukraine.

I. Chaikovsky (✉)
Glushkov Institute of Cybernetics of National Academy of Science of Ukraine, Kyiv, Ukraine

A. Popov
Electronic Engineering Department, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine

Faculty of Applied Sciences, Ukrainian Catholic University, Lviv, Ukraine

S. Mian Qaisar et al. (eds.), *Advances in Non-Invasive Biomedical Signal Sensing and Processing with Machine Learning*, https://doi.org/10.1007/978-3-031-23239-8_5

5.1 Introduction

Increasingly engaging patient participation into diagnosis and treatment decision-making (patient empowerment) is one of the main trends in modern medicine [1], supported, among other things, at the level of legislative initiatives in the healthcare system. Modern information technologies are a prerequisite for the implementation of this trend. An important part of these technologies is the increasing distribution of aesthetic hand-size gadgets for electrocardiography tests, which users may apply to some extent themselves only, without their doctor's supervision. It is more correct to call these devices as combined software - hardware systems, because. How they all have the appropriate software. The progress of microelectronics and the development of the Internet, especially "cloud" services, make these devices affordable for any user and determine their annual sales of hundreds of thousands (possibly millions) of pieces, and the devices of most manufacturers are available globally, in any country, including number in Ukraine. The analysts at Global Industry Analyst Inc. report the US portable electrocardiograph market alone being around \$1.1 billion. The annual growth of this market is at least 6% [2]. Thus, the individual use of portable electrocardiographic devices is becoming a significant social phenomenon. It can be said that with the help of such devices spontaneous screening of heart diseases occurs, far exceeding in scale any screening programs using classical 12-lead electrocardiography.

I must say that in addition to the obvious advantages, this trend also carries a certain danger. The possibilities of partial lead electrocardiography are, of course, significantly limited compared to classical electrocardiography, which is often not recognized by users who do not have professional medical knowledge. This certainly applies to screening opportunities for various heart conditions. It is also important to note that the value of even routine 12-lead electrocardiography in screening for heart diseases, primarily coronary artery disease, i.e. its ability to increase predictive accuracy over traditional risk factors is currently under intense debate. Special attention in this context is paid to the use of artificial intelligence (AI) for the analysis of the electrocardiogram (ECG). One of the main driving elements of AI in medical imaging is an aim for higher effectiveness of healthcare. The amount of medical data grows at a disproportionately high rate compared to the number of physicians available. Innovations in portable and wearable medical devices don't only bring a certain amount of useful data, but at the same time provide new opportunities for screening of certain diseases, as they can increase the number of monitoring timeframes. In this context it is extremely important to analyze slight ECG changes that are not obvious by applying common visual and/or automatic electrocardiogram interpretation as well as to develop metrics which are valid not only for 12-leads ECG but also for ECG with limited number of leads [3]. The electrocardiogram (ECG) has proved a valuable data source for AI studies [4]. Current chapter focuses on contemporary electrocardiography role opinions analysis in heart disease screening, to present portable electrocardiographic gadgets operation experience, to discuss data quantity and entity to be collected applying electrocardiography gadgets having varying lead numbers.

In addition, the results of our own research are presented, the approach based on AI method is proposed in order to define the subtle deviations of the ECG from the population norm to use in diagnosis and prognostic systems.

5.2 Evolution of Views on the Role of the Electrocardiogram in Assessing the Risk of Major Adverse Cardiovascular Events

There is no doubt that assessing an individual's risk for serious cardiovascular events is extremely important.

Many different models have been developed to assess the overall risk of developing coronary artery disease and other cardiovascular diseases. Probably the first of them was the Framingham risk scale, which includes 7 parameters [5]. The SCORE (Systematic Coronary Risk Evaluation) scale, developed on the basis of European Institutions studies results [6], is widely known.

Currently, these scales are a reliable tool for determining the likelihood of developing cardiovascular events in the next 5–10 years in patients with existing CVD and in individuals without clinical manifestations of cardiovascular pathology. At the same time, an individual profile of risk factors and concomitant cardiovascular conditions is measured in order to determine the need, tactics and intensity of clinical intervention.

However, despite all the obvious advantages, these scales, which are widely used and included in clinical guidelines, also have a number of limitations. A reflection of this fact is the emergence in recent years of new scales such as JBS3 risk score and MESA risk score, which include 15 and 12 parameters, respectively.

The most obvious problem is the practical implementation of certain interventions in people without clinical symptoms of cardiovascular diseases. How much cost will this entail? Is it possible to follow the recommendations even with a highly developed healthcare system?

According to S. Shalnova and O. Vikhireva [7], if we apply the SCORE scale to the adult population of Norway, the country with the highest life expectancy, it turns out that among 40-year-old Norwegians, every fifth woman and most men have a high risk of cardiovascular disease. Vascular diseases. At the age of 50, these figures increase to 39.5% and 88.7%, respectively. For age 65, 84.0% of women and 91.6% of men fall into the high-risk category. From a practical point of view, the strict implementation of the European recommendations seems to be very difficult even for such a prosperous country as Norway, especially in countries with a much less developed system of insurance medicine, such as Ukraine.

Individual risk is determined based on the results of extrapolation of cohort observations. It is clear that the probability of matches with a small set of features included in the risk scales for specific individuals may be low.

Naturally, risk scale estimation has no reason to be personalized, i.e. to add the conventional venture factors to personal physiologically crucial parameters obtained with equipped techniques.

Variety of all instrumental techniques makes electrocardiography undoubtedly outstanding. 12-lead electrocardiography has been used in epidemiological studies since the late 40s of the last century, in fact, from the very first steps in the epidemiology of non-communicable diseases. Especially for epidemiological studies, a method was created for measuring the elements of the electrocardiogram and describing its changes, called the Minnesota code, which makes the analysis of the electrocardiogram unified, and therefore suitable for analyzing a large amount of data. Subsequently, this classification system has been improved more than once, and related systems have appeared, such as Novacode, MEANS and some others. In accordance with these analysis systems, all ECG changes are classified into a certain number of diagnostic classes, each change can be recognized as small (minor) or large (major).

Since the late 1970s, articles based on large sample analyzes and long-term (usually 10 to 30 years) follow-up periods have been regularly published in the most reputable journals, demonstrating the value of both large and small ECG changes (in according to the Minnesota code or similar classifier) as separate warning items of lethal and non-lethal cardiovascular accidents [8–13]. There was no doubt that the use of electrocardiographic predictors increased the predictive accuracy of SCORE and other risk scales.

However, in 2011, a team of authors [14] published a systematic review commissioned by the US Commission on Preventive Services Tasks (USPSTF), which analyzed 62 studies, including almost 174,000 participants with a follow-up period from 3 to 56 years.

The key questions addressed by the authors of these guidelines are: Does electrocardiographic screening in asymptomatic individuals lead to a more accurate classification into groups having heightened, moderate, or neglectable coronary artery disease risks compared with traditional (Framingham) risk factors, what are the benefits of screening with compared to not doing it in terms of CHD outcome, whether doing screening is harmful.

The authors did not find convincing evidence that analysis of the 12-lead electrocardiogram improves the accuracy of classification into risk groups, nor evidence that the implementation of electrocardiographic screening positively affects the outcome of coronary heart disease or the appointment of a risk-reducing drug vascular treatment events (statins, aspirin). On the other hand, there are no studies demonstrating direct or indirect harm caused by electrocardiographic screening. There are only works containing general arguments about the undesirability of coronary angiography or other functional tests associated with visualization of the myocardium, when there are no sufficient grounds for this. This review formed the basis of the USPSTF recommendations [15]. The general conclusion of these recommendations is that annual routine electrocardiographic examination should not be performed in asymptomatic individuals who are at low risk on the basis of traditional risk scales. With regard to individuals belonging to groups of medium or high risk, the authors were unable to draw a definite conclusion about the balance of benefits and potential harm in the implementation of ECG screening. This strategy is also consistent with the findings of other recommendations and extensive studies published in recent years [16–18]. It must be said that these recommendations and studies have the highest evidentiary power according to modern systems for assessing the quality of recommendations for diagnostic tests, such as GRADE [19], because they are based on a meta-analysis of a large amount of data and rely on an assessment of how much a diagnostic test improves the outcome of a disease or predicts the onset of that disease.

It is interesting to note that the latest European Society of Cardiology guidelines in a field of cardiovascular disease prevention [20] advise treating individuals with a low risk of cardiovascular events, as defined by SCORE, but with evidence of LV hypertrophy on ECG, as individuals at medium risk. For other ECG syndromes, this recommendation is not given.

Separately, electrocardiographic screening should be considered in people of older age groups, and vice versa, in young people, in the latter case, in order to detect structural heart diseases and prevent sudden cardiac death, especially in people experiencing great physical and psychological overload, such as athletes and staff power structures. As for the elderly, the study by R. Auer et al. is very informative [21]., which was published after the publication of the mentioned USPSTF recommendations. In this work, on significant material (more than two thousand participants, follow-up period of 8 years), it was shown that in people aged 70–80 years, electrocardiographic signs, compared with traditional risk factors, significantly increase the predictive accuracy of screening, compared with screening only applying conventional venture factors. The most complete and context-relevant are the recommendations of the American Heart Association (AHA) published in 2014. The AHA supports such screening in principle but is not pushing for it to be mandatory. It is emphasized that significant resources are needed to conduct mandatory annual screening of all competing athletes, so the decision on the need for such screening is delegated to the local level. This view is broadly shared by the US National Institutes of Health and the European Society of Cardiology Sports Cardiology Working Group. At the same time, the International Olympic Committee more specifically recommends an electrocardiographic examination of all competitive athletes at least every two years.

Regarding law enforcement personnel, national programs of mandatory electrocardiographic screening are carried out mainly in military aviation, especially in relation to candidates for entry into the ranks of the air force (USA, Israel, Italy, etc.). For military personnel of other military branches, electrocardiographic screening is usually required only after 40 years. In younger military personnel, it is carried out only if there is evidence, for example, a burdened family history. In Ukraine, in accordance with the orders of the heads of law enforcement agencies, an electrocardiographic examination is a mandatory part of the annual medical examination, regardless of age.

In addition to electrocardiographic screening, carried out based on such long-established systems for analyzing 12-channel electrocardiograms as the Minnesota code and its derivatives (Novacode, MEANS), works have recently appeared that describe screening using new tools for analyzing electrocardiograms. Such efforts have, in our opinion, sufficient grounds. At one time, the ECG indicators included in the above classifiers of electrocardiographic signs were selected not because of their special clinical significance or exceptional physiological nature, but due to the existence at that time of extensive studies of clinical and electrocardiographic relationships in relation to these parameters. Therefore, nothing prevents other electrocardiographic signs from being investigated in the context of their value for electrocardiographic screening.

These works can be divided into two categories: one is represented by studies in which routine amplitude-time indicators are calculated, but their unusual combinations are used. An example of such works are, for example, studies of the Froelicher group [22]. In our opinion, the work of E. Gorodeski et al. [23] is of considerable interest, which shows the ability of a decision rule that includes 6 demographic and 14 electrocardiographic amplitude-time parameters selected from more than four hundred indicators to forecast the all-round reasons for women's mortality in post-menopausal period. It is important to note that in accordance with the Minnesota code, the electrocardiogram in 12 leads in all examined (more than 33 thousand observations) was normal. The follow-up period in this work was 8 years.

Another category includes works that use characteristics that are relatively complex compared to the usual amplitude-time indicators and parameters that require signal conversion using computer technology.

Such methods are currently known as the fourth generation of electrocardiography, where the first generation of electrocardiography used to be a quasi-manual assessment of the time-vs-amplitude ECG data including the eye-related analysis of plotted electrocardiographic curves, the second one included the automatically-processed measurement of the time-vs-amplitude ECG data and, similar to prior generation, graphs optical analysis. Third descent uses both automated measurement and electrocardiographic diagnostics which is followed by stating a syndromic electrocardiographic summary.

Having this said, it is obvious that the automated tools taken from both the second and the third ECG generations just copy and ease the functions of a person - a doctor of functional diagnostics. The crucial contribution of the fourth generation is the software-related assessment of the information given by these analysis methods going beyond the visual data analysis.

An example of such work is a multicenter study by T. Shlegel et al. [24], which demonstrated the advantages of a multi-parameter scale that, in addition to routine amplitude-time parameters, includes technologies such as high-frequency analysis of the QRS complex, analysis of T-wave morphology, and some others, in screening for coronary heart disease, hypertrophy left ventricle, left ventricular systolic dysfunction.

It must be said that studies belonging to the latter category have a rather weak evidentiary value in accordance with the already mentioned GRADE scale, since they provide only diagnostic accuracy (sensitivity, specificity) and no data on a clinically important outcome of the disease. It has been clear that a long-term follow-up stage is necessary in evaluating the outcome.

Let us separately discuss the ethnic difference in the parameters of a normal electrocardiogram. The ethnic and even more so racial difference in the quantitative indicators of the electrocardiogram of a healthy person has long attracted the attention of researchers. Knowing the exact limits of the ranges of normal values in different ethnic groups is necessary in order to decide which electrocardiogram is abnormal and to what extent. Ethnic and racial differences in several main amplitude-time parameters of the electrocardiogram, as well as criteria based on the assessment of the amplitudes of left ventricular hypertrophy, were studied in detail [25, 26]. There are papers describing the influence of racial differences on the effectiveness of electrocardiographic decision rules based on artificial intelligence [27]. However, multilateral, universal electrocardiogram analysis systems such as Minnesota coding, to the best of our knowledge, have not been investigated in terms of ethnic and racial variability. In addition, the problem of the influence of ethnic and racial differences on the prognostic value of certain electrocardiographic parameters is still waiting for its researcher.

5.3 The Systems of Electrocardiographic Leads, Electrocardiogram with Limited Number of Leads for Heart Disease Screening

This section briefly looks back to the development stages of the currently widespread accepted scheme for recording and analyzing the 12-lead ECG. It has been known that those 3 leads having a bipolar limb design have been exploited in clinically-performed electrocardiography, shaping the Einthoven triangle in the frontal plane, for several decades. These leads today are called standard. In 1942, Goldberger proposed enhanced unipolar limb leads that supplement the three standard leads from the point of view of areal heart electrical activity analysis, despite they are also located at the frontal plain and not equationally independent from the standard leads. Before that, in the 1930s, F. Wilson suggested 6 chest unipolar leads, which entered the practice with varied success by the end of the 1940s. The clinical contribution of the invention was broken down by Wilson himself in an iconic 1948 article [28]. These leads' value is proven by their unique feature to detect and differentiate the QRST complex pathological changes, as it was reported by the group of modern cardiography inventors. They often diagnose pathology when the limb leads remain unchanged or uninformative". The special significance of these leads in the diagnosis of ischemia in the anterior septal region of the left ventricle is emphasized. Thus, the Wilson-invented leads, which came into practice, we repeat, from the end of the 1940s of the last century, increased the sensitivity of the electrocardiography method to local changes in the myocardium and, especially, in relation to changes in the anterior wall. For time being, the standard was the ECG in 3, and later in 6 leads from the limbs.

This information is provided here to re-emphasize the importance of clearly articulating the scope and limitations of portable electrocardiographs having a certain quantity of leads.

The development of miniaturized miniaturized aesthetic hand-size gadgets for electrocardiography tests, which users may apply to some extent themselves only, without their doctor's supervision is a part of a broader tendency, given the name of POST (point-of-care testing), that among the patients briefly stands for a medical test carried out directly anywhere, not attending the doctors'.

The pioneers at this home-available gadgets area were, actually, the household automatic measurers of the blood pressure, which worldwide production was started by OMRON in 1988, having had their distribution begun 20–25 years ago.

Later, in the 90's, the personal blood testers also became available, primarily glucose level detectors. Pluggable electrocardiographs with just several electrocardiographic leads represent the next POST-tools wave that has been produced since the very beginning of the twenty-first century.

The first such devices for mass use were apparently CheckMyHeart electrocardiographs (Great Britain).

There are currently dozens of different types of portable electrocardiograph devices on the market for individual use. Basically, they represent a sole-channel electrocardiographs having finger contacts. We list some of them, namely those that are most often mentioned in the relevant reviews: AfibAlert (USA), AliveCor/Kardia (USA), DiCare (China), ECG Check (USA), HeartCheck Pen (Canada), InstantCheck (Taiwan), MD100E (China), PC-80 (China). REKA E 100 (Singapore), Zenicor (Sweden), Omron Heart Scan (Japan), MDK (The Netherlands).

All listed devices have one or more international technical certificates (ISO, CE, FDA) and are registered as electrocardiographic devices. AliveCor/Kardia and ECG Check devices are structurally integrated with smartphones, other devices are specialized electrocardiographic attachments for mobile devices (smartphone, tablet, laptop) capable of registering an ECG signal and transmitting it over a distance without distortion. Most devices are available on the open market without restrictions, manufacturers of some models (REKA E 100, ECG Check) declare that they are intended primarily for distribution by prescription. In this case, their purchase (or temporary use) is usually covered by health insurance. Many devices (for example, Zenicor) have a Web service that allows you to immediately bring the registered electrocardiogram in one lead to the doctor. Without exception, all single-channel devices have extensive and categorical disclaimers (i.e. disclaimers) explaining that for any, even the most minor, symptoms of heart disease, you should immediately consult a doctor, not relying only on the results of automatic ECG analysis in one lead.

The leading area of these gadgets' clinical use is the arrhythmic phenomenon registration, as it has been stated by the producing entities, primarily atrial fibrillation. In this regard, the experience of AliveCor/Kardia is very interesting. The company was founded in 2011. A year later, the company received the first certificate from the US Food and Drug Administration. The wording of the intended use of the certified product includes recording, displaying, storing, and transmitting a single-lead electrocardiogram using an iPhone-integrated device. In 2013, AliveCor/Kardia received a second FDA certification confirming that the electrocardiogram obtained with its device is fully equivalent to a single lead electrocardiogram recorded by a standard commercial electrocardiograph. These certificates have given the right to professional cardiologists to prescribe the AliveCor/Kardia device, combined with the iPhone. patients. In 2014, the company received permission from the FDA to freely sell its devices without intermediaries to anyone. At this stage, the device was still used only for recording, visualization and wireless transmission of the electrocardiogram, without the function of its automatic analysis. However, a few months later, an FDA certificate was obtained for the algorithm for automatic diagnosis of atrial fibrillation, which was implemented in the software. From that moment on, the user of the technology began to receive immediate feedback on the presence or absence of this type of arrhythmia. Finally, the latest FDA approval the company has received is a certification for the so-called Normal Detector, an algorithm that tells the user if their electrocardiogram is normal or abnormal, not only in relation to atrial fibrillation, but in relation to rhythm regularity in general.

Thus, modern miniature electrocardiographic software and hardware systems are developing in the direction of complicating the built-in algorithms for analyzing and interpreting an electrocardiogram - from devices designed only for recording a single-channel electrocardiogram and transmitting it to a specialist, to peculiar communicators that immediately and directly provide the user with more or less significant information about the state of his heart.

The detection of previously undiagnosed arrhythmias is one of the most natural areas for widespread use of miniature electrocardiographic devices.

Among these arrhythmias, atrial fibrillation seems to take the first place, due to its prevalence, social significance, and the relative reliability of automatic algorithms for diagnosing atrial fibrillation on a single ECG lead.

The already mentioned company AliveCor/Kardia has apparently conducted the most extensive screening studies on this topic.

Also, large-scale screening was carried out using a single-channel portable cardiograph MDK. The results of these and other similar studies convincingly prove the clinical and cost-effectiveness of atrial fibrillation screening using specialized electrocardiographic attachments for mobile devices. It is important that there is no need to organize special events for such screening. Screening is carried out during a routine visit, for example, to a family doctor, or even simply at home, if each screening participant is provided with a mobile device with an electrocardiographic attachment during the study.

Other types of screening (detection of structural heart diseases or risk factors for their occurrence), carried out using electrocardiographic gadgets having a certain lead quantity, are hardly mentioned in known studies. We can only recall the already old work [29], which shows the predictive value of T-wave flattening in the first lead in relation to myocardial infarction in young and middle-aged men. However, the sample and follow-up period presented in this paper are small and have not been confirmed in later and larger studies.

For a number of years, Ukraine's NAS Glushkov Institute of Cybernetics has been developing original electrocardiographic devices. The" philosophy" of these devices software relies on subtle ECG alterations analysis, non-obvious to the classical visual and/or automatic electrocardiogram representation, that are invisible based on conventional optical and/or automated electrocardiogram reading. The proprietary metric is developed, which allows to define the "distance" from each current electrocardiogram to the gender and age benchmark. We interpret this distance as an additional risk factor for MACE as well as a qualifier for distant future events.

Hierarchical classification of miniature electrocardiographic devices has been suggested according to two criteria – the ability of ECG signal acquisition and software capabilities (Table 5.1).

Obviously, the class of any electrocardiographic device can be defined as the sum of the scores on both of the above criteria.

The role of routine 12-lead electrocardiography in classical epidemiological studies in cardiology is being intensively discussed by the scientific community. However, the increasing use of miniature electrocardiographic devices without the participation of medical personnel will remain one of the main technological trends in the coming years. An inevitable, perhaps most important, part of this trend is the screening component, ie. identification of various, including non-trivial, electrocardiographic signs that will be interpreted by users (mostly clinically healthy people) as signs of a heart disease or indicators of an increased risk of such a disease in the future. As shown above, there is indeed a need for new, individualized indicators of increased cardiovascular complaints chances, taken from physiological response

Table 5.1 Classification of miniature electrocardiographic devices

Level	
	According to the possibilities of ECG signal acquisition
1	Only the 1st standard ECG lead
2	3 standard ECG leads (consecutive)
3	3 standard leads (parallel
4	All 6 leads from the extremities
	According to the capabilities of the software
1	Only ECG visualization, measurement of some of the simplest amplitude - time parameters of ECG and parameters of heart rate variability (HRV)
2	Ability to immediately assess the functional state and its trends based on the analysis of subtle changes in the shape, magnitude and duration of waves and intervals of ECG signal
3	Ability to immediately assess the functional state based on the formation of a multilateral integrated indicator. Elements of ECG analysis of the fourth generation

Table source: Ilya Chaikovsky

analysis of a particular person, in addition to traditional risk scales. Methods for analyzing the value of periodically emerging new risk factors for cardiovascular diseases were recently summarized by M.A. Hlatky et al. [30]. As with the GRADE scale, the main criterion is the predictive value for a clinically important outcome of the disease. It is interesting that considerable attention is paid to complex indicators, consisting of several particular ones.

However, long-term follow-up studies are only the final phase of investigating the value of a new risk marker. The conclusions of such studies become available only after many years of the "life" of a new indicator. For quite a long time, it can (and, due to the technological trend described above, will) be used without strict scientific evidence of its value. Therefore, when creating such new indicators, it is necessary to strive to make them as multilateral and complex as possible. So, with regard to new electrocardiographic markers that can be obtained using portable devices for individual use, this means using the maximum possible number of electrocardiographic leads in portable devices and a prognostic conclusion based not on any particular indicator, but on a combination of several indicators.

5.4 The Generations of ECG Analysis, some Modern Approaches Based on Mathematical Transformation of ECG Signal

As it was mentioned above, we propose to divide the methods of electrocardiogram analysis into 4 generations.

Methods of the fourth generation, in turn, can be divided into 2 groups. The first group consists of approaches that are based only on improved methods of data analysis, more informative criteria and biomarkers, electrocardiogram registration is carried out in the usual way. These 2-nd group of methods includes new technical means of signal recording.

All these methods have a common pathophysiological basis: all of them are aimed at assessing the electrical homogeneity of the myocardium by various means. In this case, the greater the heterogeneity of the myocardium from an electrical point of view, in other words, the greater the dispersion of the generated transmembrane action potentials in amplitude and length, the greater the likelihood of serious cardiovascular events.

It should be noted that one of the main areas of application of fourth generation technologies is the screening of cardiac diseases.

The feasibility of electrocardiographic screening for the most common heart diseases, especially coronary heart disease, in terms of price / effect ratio is intensively discussed. Early detection of myocardial ischemia by resting electrocardiography or ambulatory electrocardiography in individuals over 40 years without symptoms of coronary heart disease, with atypical chest pain, or with stable low-grade angina reduces the risk of severe cardiovascular events. However, the cost of a mass survey is very significant [31]. Thus, to increase the effectiveness of electrocardiographic screening, it is necessary to solve two tasks: on the one hand to reduce the cost and simplify the procedure of examination and interpretation of results, on the other hand to increase the sensitivity of the test.

fourth generation electrocardiography is aimed at solving both problems.

The most famous of the methods of the first group of fourth-generation electrocardiography technologies is the so-called "signal-averaged electrocardiogram". From the semantics point of view, a signal – averaged electrocardiogram is an electrocardiogram obtained by averaging several electrocardiographic complexes in order to improve signal quality. The averaging procedure is used in many modern methods of electrocardiogram analysis, but in the scientific literature, unfortunately, the narrow meaning of the term signal-averaged electrocardiogram (synonymous with high-resolution ECG) has taken root. This is usually called the analysis of late potentials, i.e. time and spectral study of low-amplitude and high-frequency signals in the final part of the QRS complex and the initial part of the ST segment. This method is widely used, there are consensus documents of various cardiological societies (such as the American College of Cardiology) [32]. The value of the analysis of late potentials to determine the risk of ventricular tachycardia in patients with myocardial infarction has been proven. Less strong, but still sufficient evidence suggests that this method is also useful for determining the risk of ventricular tachycardia in patients with non-coronary cardiomyopathy [33]. An improvement of this method is the analysis of the late potentials of P-wave, which is used to assess the risk of paroxysms of atrial fibrillation [34].

However, it can be said that recent analysis of late potentials has been replaced by newer and more modern methods of electrocardiogram analysis, which have not yet been tested as widely as late potential analysis but are even more promising in terms of their value in determining cardiovascular events.

We consider the morphology (shape) analysis of the electrocardiogram T-wave as potentially interesting and electrophysiologically reasonable. Mathematically, this method is the decomposition of the electrocardiographic signal by values at special (singular) points with the analysis of the main components [35]. In the framework of this method, the so-called coefficient of complexity of the T-wave is calculated.

This coefficient indicates the extent to which the electrocardiogram T-wave shape can be described by a simple dipole model of the electrocardiogram source and, accordingly, the contribution of more complex sources. The higher this coefficient, the more heterogeneous the myocardium in electrical terms.This T-wave morphology consideration has proved to be a good predictor of the risk of cardiovascular events in the general population [36], among young athletes [37] and for myocardial infarction patients. Other, simpler approaches to the consideration of the electrocardiogram T-wave shape, also exist. Thus here, it is necessary to mention the ECG analysis in the phase space. If the system is described by two variables, then the phase space has two dimensions, and each variable corresponds to one dimension. In this case, the phase space is a phase plane, i.e. a rectangular coordinate system, along the axes of which the values of the two variables are being plotted. Having said that, the technology nature states that at each point of the initial electrocardiographic response in the time domain, its first derivative is estimated by quantitative methods, and all further processing is carried out on the phase plane. This approach to ECG analysis has been used for a long time, at least since the late 70's of last century [38]. The Cybernetic Center of the National Academy of Sciences of Ukraine has been developing a different approach to ECG phase analysis for several years. This treatment involves dividing the phase trajectory into separate cardiac cycles, selecting trajectories with the same morphology, trajectories being averaged in phase space with subsequent "reference cycle" assessment on the average phase trajectory.

Thus, it is the indicators of the shape of the average phase trajectory that are evaluated, and the trajectories of ectopic cardiocomplexes are not considered, while in the other above-mentioned works on ECG phase analysis, on the contrary, they are the subject of analysis. We have proposed a number of quantitative indicators for the analysis of the shape of the average phase trajectory, the most sensitive of which was the symmetry index, i.e. the ratio of the maximum velocity at the rising segment of the T-wave to the maximum velocity at the descending segment of T-wave (at positive T) or the ratio of the maximum velocity on the descending segment of the T-wave to the maximum speed on the ascending segment of the T-wave (with negative T).

The diagnostic value of this approach to ECG analysis in many clinical situations has been demonstrated, including the analysis of only first electrocardiographic lead [39].

Also, a certain reserve of increasing diagnostic informativeness in the assessment of repolarization patterns may be linked to the evolution of mathematical description models of the ST-segment and T-wave by several approximation functions [40].

Another promising approach is the so-called high-frequency analysis of the QRS complex. It consists in calculating the signal power in the band 150–250 Hz in the middle of the QRS complex part of ECG. It has been shown that a decrease in this indicator is a reliable predictor of myocardial ischemia, both in acute coronary syndrome and in chronic asymptomatic ischemia [41].

It is reasonable to mention a simple and clear approach based on the calculation of the spatial angle between the vertices of the QRS complex and the T wave of the electrocardiogram. This parameter is essentially an improved ventricular gradient of Wilson, known since 1934. In recent years, large-scale studies have shown that this simple indicator is a strong predictor of cardiovascular events and mortality in the general population, and especially among women [42].

At the end of the last and the beginning of our century, such an electrocardiographic indicator of myocardial electrical homogeneity was widely used as the spatial variance of the QT interval, i.e. the difference between the longest and shortest QT interval in 12 leads. Recently, this indicator has been criticized, but certainly has not yet exhausted its usefulness [43].

The approach developed by L. Titomir [44], called dipole electrocardiotopography (DECARTO), is interesting and well-founded. This is a method of visual display and analysis of information obtained using 3 orthogonal leads. It is a quasi-mapping of the electrical process in the ventricles of the heart based on orthogonal ECG, based on the use of a model of ventricular depolarization wave, which is reflected by the electrical vector of the heart. The components of this vector are proportional to the corresponding signals of orthogonal leads. At each point in time, the depolarization front is projected onto a spherical quasipericardium (a sphere centered in the geometric center of the ventricles that covers the heart) in the form of a spherical segment. The main area of application of this method is acute coronary syndrome, the prognosis of short-term and long-term results of treatment of acute myocardial infarction.

Finally, a separate subgroup should include methods based not on the analysis of indirect electrocardiographic signals, but rather the variability of certain characteristics of individual cardio-complexes over a period of time. It is necessary to distinguish these methods from the analysis of heart rate variability, when analyzing not the parameters of your own cardiocycle, but only R-R intervals.

There are many methods and estimates of variability of certain elements of the cardiac signal from complex to complex. This is an analysis of the variability of the amplitude of the T-wave at the microvolt level, and some others. The most common of these is the analysis of the duration of the variability of the QT interval (QTV). This indicator is also used to assess the risk of life-threatening ventricular arrhythmias in patients diagnosed with heart diseases [45].

This review is far from being complete. There are other modern methods of electrocardiogram analysis, the authors of which insist on their high efficiency. In all this diversity, the clinician can easily “drown”. Therefore, information technologies should be developed that summarize the data obtained with several modern computerized methods of electrocardiogram analysis and offer the doctor an integral coefficient that shows the probability of a heart disease or cardiovascular event. In this regard, it should be noted the results obtained by the Laboratory of Functional Diagnostics of the US National Aerospace Administration (NASA) in Houston.

In recent years, much attention of researchers is attracted by electrocardiogram analysis using artificial intelligence methods. As a matter of fact, ECG pays benefits for AI applications deep learning. The ECG is highly accessible and provides iterable unprocessed data which is available for digital storing and transferring. Another feature is totally automated ECG representation, when the accurate study applications use huge ECG data banks and sets of clinical data, cooperated with cutting-edge computer abilities, are demonstrating the usefulness of the AI-engaged ECG, the detection apparatus of ECG signs and structures unseen to the man’s eye. These structures are able to spot the cardiac disease, such as left ventricular (LV) systolic dysfunction, silent atrial fibrillation (AF) and hypertrophic cardiomyopathy (HCM), but might also reflect systemic physiology, such as a person’s age and sex or their serum potassium levels etc. [46] Of course, the use of AI for ECG analysis in the context of predicting various heart diseases is of particular interest. In the work [47] it is shown that the ECG-AI model based solely on information extracted from ECG independently predicts HF with accuracy comparable to existing FHS and ARIC risk calculators. There are several other high-quality works on this topic.

Also, we would especially like to emphasize the work in which it is demonstrated that the difference between individual biological age, defined by AI ECG, and chronological age is an independent predictor of all-cause and cardiovascular mortality. Discrepancies between these possibly reflect disease independent biological aging [48].

It should be noted that the further development of fourth generation electrocardiography is impossible without mathematical and computer modeling. However, models of heart electrical activity which are set with high levels of the object spatial distribution have considerable gaps with empirical models mostly used to define the links of electrophysiological phenomena in a heart at the visceral level with ECG changes. The latter primarily involves the task of understanding the ECG pathology changes mechanisms in myocardial ischemia due to the rise in its electrical inhomogeneity and gradually its unsteadiness. Given the inconsistency of a few sets of testing volumes, this context complicates the new algorithm evolution for diagnosing ischemia in its early stages, as well as methods for quantifying some manifestations of ECG disorders ofheart muscle activation and repolarization applying the only model.

The development of computer technology and information technology has given a new impetus to electrocardiography. The electrocardiographic signal, which is easily recorded and digitized, allegedly "provokes" doctors and mathematicians to cooperate under the motto of Galileo Galilei: "Measure everything that can be measured, make measurable everything that has not been measured before." The result of this collaboration is the creation of new effective methods of electrocardiographic diagnosis, which over time will find a place in every clinic and doctor's office and may replace traditional electrocardiography.

5.5 Anomaly Detection in ECG Using Machine Learning Approach

Many devices can provide ECG in clinical controlled conditions, as well as during exercise and in 24/7 regime during everyday life activities. Comparatively wide availability of the ECG recordings, both clinically and from wearable devices supports its usage as a basis of the online diagnosis tool. Performing ECG anomaly detection tries to serve as the prediction and prevention tool for dangerous health conditions associated with heart malfunctioning.

In ECG the anomalous behavior may be represented as the irregular heart rhythms or heartbeats with unusual time-magnitude parameters. Despite many deviations from the conditional normal ECG have been described in the literature, their combination, or the tiny changes at the beginning of the disease development, may not be obvious, be rare events, or be hidden. Therefore, application of rule-based anomaly detection may be less effective than use of the data-driven machine learning approach.

Anomaly detection is one of the fundamental tasks in data mining and consists of identifying the objects which considerably deviate from some concept of normality in each dataset [49]. Depending on the context, such objects can be irregular, unexpected, rare, or simply “strange”. Due to the fuzzy and application-dependent definitions for “deviation” and “normality”, there exist a lot of anomaly detection algorithms. The algorithms of anomaly detection may be roughly subdivided into two major classes: unsupervised and supervised, depending on the availability of the labeled data. Excellent description of the anomaly detection techniques can be found in [50].

Unsupervised methods do not use the information about the labels of the data (normal or anomalous) in the training set. This group of methods contains model-based methods (relying on the description of the data generation model) and proximity-based methods (which use the distances between data points in the feature space). In supervised anomaly detection, the prior information about the labels in the training dataset is available. In this case, many general-purpose classification methods can potentially be used. But in case of application to anomaly vs. normal object classification, this is less used than unsupervised, because the number of anomalies is usually limited, and often the anomalies are not available beforehand.

Many methods of anomaly detection have been applied to the ECG analysis, focusing on the different types of animal characteristics. In [51], the review of most widely used approaches is presented. Authors of [52] applied Support Vector Machine classifier after wavelet-based extraction of heart rate variability features to of arrhythmic beat classification. Multilinear principal component analysis is used to process ECG for extracting disease-altered patterns, followed by anomaly detection using deep support vector data description in [53]. The proposed framework achieves superior performance in detecting abnormal ECG patterns, such as atrial fibrillation, right bundle branch block, and ST-depression. Recently, deep learning networks training is getting more attention for ECG anomaly detection development. In [54] a novel hybrid architecture consisting of Long Short Term Memory cells and Multi-Layer Perceptrons. Simultaneous training pushes the overall network to learn descriptive features complementing each other for making decisions, which led to the average classification accuracy of 97% across several records in the ECG database.

5.6 Isolation Forest Anomaly Detection for Quantifying the Deviation of Signal Averaged ECG from Population Norm

5.6.1 Isolation Forest Unsupervised Anomaly Detection

Isolation Forest (IF) [55] is one of the most popular methods for anomaly detection [56]. To define if the vector of features representing the ECG under analysis is an anomaly or not in the group of vectors, unsupervised learning is used for isolating the anomalies. The idea is that the anomalies are easier to separate from the cluster than non-anomalous vectors because they lie on the outskirts of the cluster. To isolate a vector in space, the IF algorithm recursively selects the axis in a feature space, and then randomly splits this axis by selecting the value between the minimum and maximum of the corresponding feature values. After several partitions, the coordinate of the single vector appears to be separated. The number of partitions required to isolate each vector is used to compute the score of anomalies of that vector. If the vector is located far from the rest of the vectors, the number of partitions required to isolate it is quite small, since the coordinates of the vector deviate substantially from the coordinates of other vectors. This method is applied to the identification of tiny changes in ECG from several groups of subjects.

5.6.2 Subjects Data

Five different groups were examined [57]: healthy subjects with no reported cardiovascular problems (Normal Controls, NC), subjects with proven coronary artery diseases (CID), subjects recovered from COVID-19, military personnel with mine-explosive injuries (Combatants), and two subgroups of participants of mass-screening in one of the rural region of Ukraine. Subgroup 1 consisted of persons, who died during five- years follow-up (all-cause mortality), subgroup 2 - persons, who didn't die during this period.

Signal averaged ECG (SAECG) considering a group of 181 people (males, aged from 18 to 28) is used in this study. Originally the data contained ECGs recorded in six ECG leads (I, II, III, aVR, aVL, aVF). From each of six SAECG leads, 34 features were extracted:

- Durations of P, Q, R, S peaks, and QRS complex,
- Durations of PR, QR, TR intervals,
- Duration of JT segment, duration from J to the top of T peak, duration from the top to the end of T peak,
- Amplitudes of P, Q, R, S, T peaks, and J wave,
- Mean magnitude over the ST segment, the magnitude at the end of ST-segment,
- Area under P peak, area under the QRS complex,

- Areas under the T peak from the beginning to the top, and from the top to the end,
- Ultimate derivative values are at ascending/descending parts of T peak.

In total, 204 features from each lead are calculated to characterize multichannel SAECG.

5.6.3 *Quantification of the Distance to the Norm*

After SAECG features are extracted, each ECG can be presented as a vector in the 204-dimensional feature space. The value of every feature is the coordinate of the particular ECG in that space with respect to the corresponding axes. In the case of having the group of ECG with similar characteristics, the corresponding feature vectors will form in the cluster in the space. If the particular ECG is located far from the cluster, this might indicate that their features are distinct from those of the cluster members. The vector of ECG which is similar to the group of ECGs forming the cluster will be located within the cluster.

In this work, the concept of outlier/inlier is proposed to be used for detecting the deviations of the ECG from the group of other ECGs. To define whether the particular ECG is an outlier or not, the Isolation Forest anomaly detector is used.

The procedure to use IF in defining the deviation of the current ECG from the group of the norm is as follows:

1. For the group of normal ECGs, SAECG is obtained, and the features of SAECGs are extracted.
2. Train the IF anomaly detector using the group of normal ECGs.
3. For the new ECG under analysis, pass it through the IF and obtain the anomaly score.

The negative values of the anomaly score indicate that the ECG is an anomaly; this is interpreted as the substantial deviation of the ECG from the norm. Additionally, the absolute value of the anomaly score can be used as a degree of deviation from the normal group. The bigger the absolute value is, the more distant the SAECG is located from the population norm, and therefore the difference between the current ECG and the normal group is more significant.

5.6.4 *Experiment Results*

In Fig. 5.1 the distribution of subjects from NC, CID, and COVID groups are presented. As vertical axes, the projection on the main eigenvector from PCA decomposition of the feature matrix is provided for reference, to construct the 2D plot. Table 5.2 contains mean and standard deviation values of the distance from the centers of the cluster for different groups.

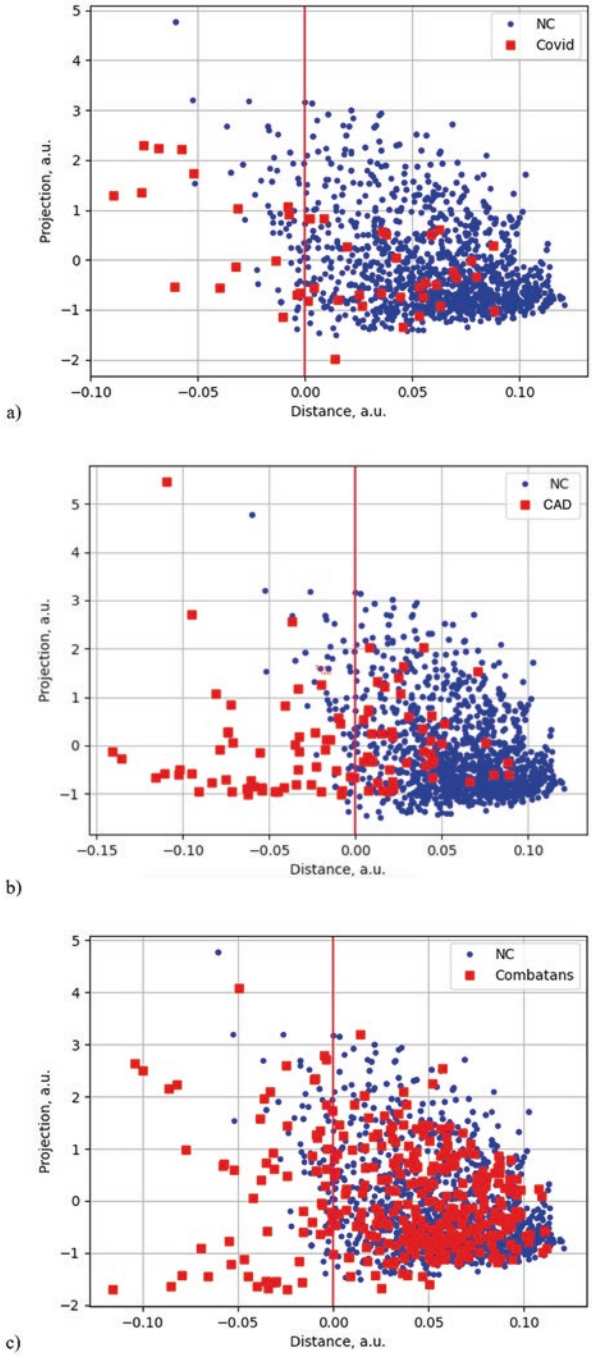

Fig. 5.1 The distribution of subjects from COVID group (**a**), CAD group (**b**), and Combatants group (**c**) with respect to NC subjects. (Image source: Illya Chaikovsky, Anton Popov)

Table 5.2 Mean and standard deviation of the anomaly score

Group name	Mean ± std. of the anomaly score
Healthy volunteers	0.0583 ± 0.0318
CAD	−0.0129 ± 0.0522
Wounded combatants	0.0389 ± 0.0431
COVID patients	0.0157 ± 0.0481
Subjects, died within 5 years of follow-ups (all-cause mortality)	−0.0102 ± 0.0601
Subjects, didn't die within 5 years of follow-ups	0.0204 ± 0.0498

Table source: Illya Chaikovsky, Anton Popov

As one can see, the distribution of distances and relative amount of outliers differs for the different groups, suggesting the receptiveness of the proposed method to the characteristics of the original SAECG. In case of COVID subjects, the positions are distributed quite equally, and points representing the COVID patients have both positive and negative distance values. This suggests that ECG of COVID subjects does not contain subtle deviations from the NC group. In contrast, many of CAD subjects have a negative distance from the NC group, which suggests the difference between the SAECG characteristics due to the changes of heart activity.

When estimating distance between the studied groups and normal controls it was found that the largest distance takes place between healthy volunteers' group and CAD patients group and group of subjects who died within 5 years of follow-ups (all-cause mortality). This may suggest further development of the mortality predictive score based on the outlier detection. The minimal distance from NC was detected in the Combatants group.

As already mentioned, the role of artificial intelligence methods in ECG analysis is increasing significantly. Of course, the disadvantages of this method should not be underestimated [58, 59]. Among these shortcomings, the need to analyze a large amount of data is usually noted, which requires a high qualification of the researchers. Often there is a need to update the used models during the study or after deployment due to the shift in the data. Another important problem is overfitting, i.e. a situation where the algorithm over-adjusts to the training sample, reaching maximum accuracy in it, but its performance is much worse on other populations. Solving this problem requires "fine tuning" of the training process.

However, the biggest problem hindering the widespread implementation of AI methods in the practice of cardiologists seems to be the "black box nature" of the findings obtained using AI. This means that clinicians have a poor understanding of the logical chain of inferences that lead to certain results, and therefore do not trust them. To solve this problem, in recent years, a special line has been developing in the theory and practice of machine learning methods, which is called explainable machine learning [60]. Of course, this approach also has its limitations.

The above limitations, especially the "black box nature", probably lead, among other things, to the fact that AI-assisted ECG analysis methods are used mainly for solving more specific problems, and are rarely used to determine the global risk of death from all causes. Only a few works on this topic are known. In the work of A.A. Mahayni and co-authors [61] used AI to predict long-term mortality after cardiac surgery. They used preoperative ECGs from subjects with detected left ventricular ejection fraction to train convolutional neural network for binary classification, and demonstrated increased prognostic value in the prediction of long-term mortality in patients undergoing cardiac surgical procedures. Another work [62] used a massive amount of the 12-lead resting-state ECGs to train deep neural network and predict the 1-year mortality with high efficacy. It was demonstrated that even the ECG, which is interpreted as normal by a qualified electrocardiologist, may provide important prognostic information to the AI algorithm and help to build correct predictions.

We, within the framework of our metric, interpret the distance from each current electrocardiogram to the threshold as a risk factor for long-term outcomes .

The novelty of our approach is the ability to specify the groups of SAECGs, from which we want to find the deviations. This enables adjustment of the risk analysis to the cohorts of people who inherently have specific characteristics of heart electrical activity, such as sportsmen, children, etc. We demonstrated that the same framework could be used to catch the tiny deviations from the group norm, and quantify its value, which may be used as a prognostic feature of the risk. The advantages of the employed Isolation Forest anomaly detector is its ability to work with smaller sample sizes, which is useful for the case of having relatively small groups to compare the ECG with. At the same time, IF can scale to handle the extremely large data volume and is robust to the presence of the irrelevant features. Together with its ability to directly isolate the anomalies in the dataset and provide a quantified score of the distance from the cluster, it is best suitable for the detection of subtle changes in the SAECGs with respect to the group norm.

5.7 Teaching Assignment

1. Describe the evolution and current opinion regarding the role of the electrocardiogram in assessing the venture for major adverse cardiovascular accidents.
2. Describe the basic characteristics and classification of the miniature electrocardiographic devices.
3. Highlight the generations of ECG analysis, describe the basic characteristics of each.
4. Describe your reflection of the adjustments of the risk factors to different groups of subjects, proposed in this chapter.
5. Describe the main idea of the interpretation of the distance from the group of feature vectors, proposed in this chapter.

5.8 Conclusion

In the nearest time there will certainly be more cases for delegating the patients with rights to test their heart state. Electrocardiographic tools with mobile phone size and simple access to analyzing cloud servers are becoming crucial for giving the patients personalized medicine approach.

Electrocardiograms analysis cloud platforms with few leads are going to develop gradually from oversimplified processing of just several rhythm disturbances to more sophisticated analysis, services, and diagnosis. The decision-making developments involving artificial intelligence as their base will give estimations for everyone. Their goals are severe cardiovascular risks both in the general population and, especially in certain cohorts, such as those with diabetes, pre-diabetes, and heart failure patients. The attempts are to be made to cope with the most powerful barrier to the engagement of machine learning logics – their black-box character, i.e., the difficulty, especially for clinicians, to perceive and moreover trust the representation of data leading to the diagnosis. Based on electrocardiogram and heart rate variability analysis, individualized recommendations will be made regarding frequency, duration, intensity, type, and total amount of physical activity, as well as detailed dietary recommendations.

On the other hand, the classic 12-lead electrocardiogram will remain the most frequently used technology in clinical cardiology for the long time. The progress in biomedical computing and signal processing, and the available computational power offer fascinating new options for electrocardiogram analysis relevant to all fields of cardiology. Successful implementation of artificial intelligence technologies for the analysis of routine 12-channel electrocardiogram is a critical step to further increase the power of electrocardiography. In this respect, the huge amount of digitized electrocardiogram recordings from affordable and widespread devices enables the growth of artificial intelligence data-driven machine learning approaches. The sophisticated algorithms requiring the training data will be more available both on the devices and as the cloud-based services, providing the automated diagnosis, prediction, and prevention of cardiovascular diseases and supporting human well-being.

References

1. J. Calvillo, I. Román, L.M. Roa, How technology is empowering patients. A literature review. 2013. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.893.3296&rep=rep1&type=pdf
2. Single lead ECG equipment market size. [cited 16 Feb 2021]. Available: https://www.grandviewresearch.com/industry-analysis/single-lead-ecg-equipment-market
3. I. Chaikovsky, Electrocardiogram scoring beyond the routine analysis: Subtle changes matters. Expert Rev. Med. Devices **17**(5), 379–382 (2020)
4. Z.I. Attia, D.M. Harmon, E.R. Behr, P.A. Friedman, Application of artificial intelligence to the electrocardiogram. Eur. Heart J. **42**(46), 4717–4730 (2021)

5. K.M. Anderson, P.W. Wilson, P.M. Odell, W.B. Kannel, An updated coronary risk profile. A statement for health professionals. Circulation **83**, 356–362 (1991)
6. R.M. Conroy, K. Pyörälä, A.P. Fitzgerald, S. Sans, A. Menotti, G. De Backer, D. De Bacquer, P. Ducimetière, P. Jousilahti, U. Keil, I. Njølstad, R.G. Oganov, T. Thomsen, H. Tunstall-Pedoe, A. Tverdal, H. Wedel, P. Whincup, L. Wilhelmsen, Graham IM; SCORE project group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. Eur. Heart J. **24**(11), 987–1003 (2003)
7. S.A. Shal'nova, O.V. Vikhireva, Otsenkasummarnogoriskaserdechno-sosudistykhzabolevaniy. Ratsional'nayafarmakoterapiya v kardiologii. 2005; 1. Available: https://cyberleninka.ru/article/n/7138553
8. Pooling Project Research Group. Relationship of blood pressure, serum cholesterol, smoking habit, relative weight and ECG abnormalities to incidence of major coronary events: Final Report of the Pooling Project.. American Heart Association; 1978
9. W.B. Kannel, K. Anderson, D.L. McGee, L.S. Degatano, M.J. Stampfer, Nonspecific electrocardiographic abnormality as a predictor of coronary heart disease: The Framingham study. Am. Heart J. **113**, 370–376 (1987)
10. D. De Bacquer, G. De Backer, M. Kornitzer, H. Blackburn, Prognostic value of ECG findings for total, cardiovascular disease, and coronary heart disease death in men and women. Heart **80**, 570–577 (1998)
11. M.L. Daviglus, Y. Liao, P. Greenland, A.R. Dyer, K. Liu, X. Xie, C.F. Huang, R.J. Prineas, J. Stamler, Association of nonspecific minor ST-T abnormalities with cardiovascular mortality: The Chicago Western Electric Study. JAMA **281**(6), 530–536 (1999)
12. P. Greenland, X. Xie, K. Liu, L. Colangelo, Y. Liao, M.L. Daviglus, A.N. Agulnek, J. Stamler, Impact of minor electrocardiographic ST-segment and/or T-wave abnormalities on cardiovascular mortality during long-term follow-up. Am. J. Cardiol. **91**(9), 1068–1074 (2003)
13. P. Denes, J.C. Larson, D.M. Lloyd-Jones, R.J. Prineas, P. Greenland, Major and minor ECG abnormalities in asymptomatic women and risk of cardiovascular events and mortality. JAMA **297**, 978–985 (2007)
14. Chou R, Arora B, Dana T, Fu R, Walker M, Humphrey L. Screening asymptomatic adults with resting or exercise electrocardiography: a review of the evidence for the U.S. preventive services task force. Centre for Reviews and Dissemination (UK); 2011
15. Moyer VA. Screening for Coronary Heart Disease With Electrocardiography: U.S. Preventive Services Task Force Recommendation Statement. Annals of Internal Medicine. 2012
16. AAFP recommendations for preventive services guideline. In: AAFP [Internet]. [cited 28 Feb 2020]. Available: https://www.aafp.org/online/etc/medialib/aafp_org/documents/clinical/CPS/rcps08-2005.Par.0001.File.tmp/October2012SCPS.pdf
17. P. Greenland, J.S. Alpert, G.A. Beller, E.J. Benjamin, M.J. Budoff, Z.A. Fayad, E. Foster, M.A. Hlatky, J.M. Hodgson, F.G. Kushner, M.S. Lauer, L.J. Shaw, S.C. Smith Jr., A.J. Taylor, W.S. Weintraub, N.K. Wenger, A.K. Jacobs, S.C. Smith Jr., J.L. Anderson, N. Albert, C.E. Buller, M.A. Creager, S.M. Ettinger, R.A. Guyton, J.L. Halperin, J.S. Hochman, F.G. Kushner, R. Nishimura, E.M. Ohman, R.L. Page, W.G. Stevenson, L.G. Tarkington, C.W. Yancy, American College of Cardiology Foundation; American Heart Association, 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: A report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines. J. Am. Coll. Cardiol. **56**(25), e50–e103 (2010)
18. A. Groot, M.L. Bots, F.H. Rutten, H.M. den Ruijter, M.E. Numans, I. Vaartjes, Measurement of ECG abnormalities and cardiovascular risk classification: A cohort study of primary care patients in the Netherlands. Br. J. Gen. Pract. **65**, e1–e8 (2015)
19. British Medical Journal Publishing Group, GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ **336**, 1106 (2008). https://doi.org/10.1136/bmj.a139
20. Committee for Practice Guidelines ESC. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012) The Fifth Joint Task Force of the European Society

of Cardiology and American Heart Association. Eur. Heart J. 2012. Available: https://academic.oup.com/eurheartj/article-abstract/33/13/1635/488083
21. R. Auer, D.C. Bauer, P. Marques-Vidal, J. Butler, L.J. Min, J. Cornuz, S. Satterfield, A.B. Newman, E. Vittinghoff, N. Rodondi, Health ABC Study, Association of major and minor ECG abnormalities with coronary heart disease events. JAMA **307**(14), 1497–1505 (2012)
22. S.Y. Tan, G.W. Sungar, J. Myers, M. Sandri, V. Froelicher, A simplified clinical electrocardiogram score for the prediction of cardiovascular mortality. Clin. Cardiol. **32**, 82–86 (2009)
23. E.Z. Gorodeski, H. Ishwaran, U.B. Kogalur, E.H. Blackstone, E. Hsich, Z.M. Zhang, M.Z. Vitolins, J.E. Manson, J.D. Curb, L.W. Martin, R.J. Prineas, M.S. Lauer, Use of hundreds of electrocardiographic biomarkers for prediction of mortality in postmenopausal women: The Women's Health Initiative. Circ. Cardiovasc. Qual. Outcomes **4**(5), 521–532 (2011)
24. T.T. Schlegel, W.B. Kulecz, A.H. Feiveson, E.C. Greco, J.L. DePalma, V. Starc, B. Vrtovec, M.A. Rahman, M.W. Bungo, M.J. Hayat, T. Bauch, R. Delgado, S.G. Warren, T. Núñez-Medina, R. Medina, D. Jugo, H. Arheden, O. Pahlm, Accuracy of advanced versus strictly conventional 12-lead ECG for detection and screening of coronary artery disease, left ventricular hypertrophy and left ventricular systolic dysfunction. BMC Cardiovasc. Disord. **10**, 28 (2010)
25. A.L. Arnold, K.A. Milner, V. Vaccarino, Sex and race differences in electrocardiogram use (the National Hospital Ambulatory Medical Care Survey). Am. J. Cardiol. **88**, 1037–1040 (2001)
26. P.M. Okin, J.T. Wright, M.S. Nieminen, S. Jern, A.L. Taylor, R. Phillips, V. Papademetriou, L.T. Clark, E.O. Ofili, O.S. Randall, L. Oikarinen, M. Viitasalo, L. Toivonen, S. Julius, B. Dahlöf, R.B. Devereux, Ethnic differences in electrocardiographic criteria for left ventricular hypertrophy: The LIFE study. Losartan intervention for endpoint. Am. J. Hypertens. **15**(8), 663–671 (2002)
27. P.A. Noseworthy, Z.I. Attia, L.C. Brewer, S.N. Hayes, X. Yao, S. Kapa, P.A. Friedman, F. Lopez-Jimenez, Assessing and mitigating bias in medical artificial intelligence: The effects of race and ethnicity on a deep learning model for ECG analysis. Circ. Arrhythm. Electrophysiol. **13**(3), e007988 (2020)
28. F.N. Wilson, The clinical value of chest leads. Br. Heart J. **10**(2), 88–91 (1948)
29. M.L. Løchen, K. Rasmussen, P.W. Macfarlane, E. Arnesen, Can single-lead computerized electrocardiography predict myocardial infarction in young and middle-aged men? J. Cardiovasc. Risk **6**(4), 273–278 (1999)
30. M.A. Hlatky, P. Greenland, D.K. Arnett, C.M. Ballantyne, M.H. Criqui, M.S. Elkind, A.S. Go, F.E. Harrell Jr., Y. Hong, B.V. Howard, V.J. Howard, P.Y. Hsue, C.M. Kramer, M.C. JP, S.L. Normand, C.J. O'Donnell, S.C. Smith Jr., P.W. Wilson, American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council, Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. Circulation **119**(17), 2408–2416 (2009)
31. W. Stanford, Screening of coronary artery disease: Is there a cost-effective way to do it? Am. J. Cardiol Imaging **10**(3), 180–186 (1996)
32. ACC Consensus Document on Signal-Averaged Electrocardiography. JACC **27**(1), 238–249 (1996)
33. J.F. Robillon, J.L. Sadoul, D. Jullien, P. Morand, P. Freychet, Abnormalities suggestive of cardiomyopathy in patients with type 2 diabetes of relatively short duration. Diabete Metab. **20**(5), 473–480 (1994)
34. F. Cecchi, A. Montereggi, I. Olivotto, P. Marconi, A. Dolara, B.J. Maron, Risk for atrial fibrillation in patients with hypertrophic cardiomyopathy assessed by signal averaged P wave duration. Heart **78**(1), 44–49 (1997)
35. F. Extramiana, A. Haggui, P. Maison-Blanche, R. Dubois, S. Takatsuki, P. Beaufils, A. Leenhardt, T-wave morphology parameters based on principal component analysis reproducibility and dependence on T-offset position. Ann. Noninvasive Electrocardiol. **12**(4), 354–363 (2007)

36. P.M. Okin, R.B. Devereux, R.R. Fabsitz, E.T. Lee, J.M. Galloway, B.V. Howard, Principal component analysis of the T wave and prediction of cardiovascular mortality in American Indians: The strong heart study. Circulation **105**(6), 714–719 (2002)
37. A. Pelliccia, F.M. Di Paolo, F.M. Quattrini, C. Basso, F. Culasso, G. Popoli, R. De Luca, A. Spataro, A. Biffi, G. Thiene, B.J. Maron, Outcomes in athletes with marked ECG repolarization abnormalities. N. Engl. J. Med. **358**(2), 152–161 (2008)
38. C.N. Mead, S.M. Moore, K.W. Clark, B.F. Spenner, L.J. Thomas Jr., A detection algorithm for multiform premature ventricular contractions. Med. Instrum. **12**(6), 337–339 (1978)
39. I.A. Chaikovsky, L.S. Fainzilberg, *Medical aspects of the use of the FASAGRAPH device in clinical practice and at home* (Kyiv, IRT Center ITIS, 2009), p. 74
40. O.V. Baum, I.A. Chaĭkovskiĭ, L.A. Popov, V.I. Voloshin, L.S. Faĭnzil'berg, M.M. Budnik, Electrocardiographic image of myocardial ischemia: Real measurements and biophysical models. Biofizika **55**(5), 925–936 (2010)
41. T.T. Schlegel, W.B. Kulecz, J.L. DePalma, A.H. Feiveson, J.S. Wilson, M.A. Rahman, M.W. Bungo, Real-time 12-lead high-frequency QRS electrocardiography for enhanced detection of myocardial ischemia and coronary artery disease. Mayo Clin. Proc. **79**(3), 339–350 (2004)
42. P. Rautaharju, C. Kooperberg, J. Larson, A. LaCroix, Electrocardiographic abnormalities that predict coronary heart disease events and mortality in postmenopausal women. Circulation **113**, 473–480 (2006)
43. M. Malik, V.N. Batchvarov, Measurement, interpretation and clinical potential of QT dispersion. J. Am. Coll. Cardiol. **36**(6), 1749–1766 (2000)
44. L.I. Titomir, N.E. Barinova, *Electrocardiographic mapping* (Methodological Guide, Moscow, 2006) 51 p
45. W.L. Atiga, H. Calkins, J.H. Lawrence, G.F. Tomaselli, J.M. Smith, R.D. Berger, Beat-to-beat repolarization lability identifies patients at risk for sudden cardiac death. J. Cardiovasc. Electrophysiol. **9**(9), 899–908 (1998)
46. K.C. Siontis, P.A. Noseworthy, Z.I. Attia, P.A. Friedman, Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. Nat. Rev. Cardiol. **18**(7), 465–478 (2021)
47. O. Akbilgic, L. Butler, I. Karabayir, P.P. Chang, D.W. Kitzman, A. Alonso, L.Y. Chen, E.Z. Soliman, ECG-AI: Electrocardiographic artificial intelligence model for prediction of heart failure. Eur. Heart J. Digit. Health **2**(4), 626–634 (2021)
48. A.O. Ladejobi, J.R. Medina-Inojosa, M. Shelly Cohen, Z.I. Attia, C.G. Scott, N.K. Le Brasseur, B.J. Gersh, P.A. Noseworthy, P.A. Friedman, S. Kapa, F. Lopez-Jimenez, The 12-lead electrocardiogram as a biomarker of biological age. Eur. Heart J. Digit. Health **2**(3), 379–389 (2021)
49. L. Ruff, J.R. Kauffmann, R.A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T.G. Dietterich, K.R. Müller. A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE. 2021 Feb 4
50. C.C. Aggarwal, *Outlier analysis* (Springer, Cham, 2017), p. 466
51. H. Li, P. Boulanger, A survey of heart anomaly detection using ambulatory electrocardiogram (ECG). Sensors **20**(5), 1461 (2020)
52. C. Venkatesan, P. Karthigaikumar, P. Anand, S. Satheeskumaran, R. Kumar, ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications. IEEE Access **6**, 9767–9773 (2018)
53. H. Zhou, C. Kan, Tensor-based ECG anomaly detection toward cardiac monitoring in the internet of health things. Sensors **21**(12), 4173 (2021)
54. G. Sivapalan, K.K. Nundy, S. Dev, B. Cardiff, D. John, ANNet: a lightweight neural network for ECG anomaly detection in IoT edge sensors. IEEE Trans. Biomed. Circuits Syst. **16**(1), 24–35 (2022)
55. F. T. Liu, K. M. Ting and Z. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, 2008, pp. 413–422. https://doi.org/10.1109/ICDM.2008.17

56. F.T. Liu, K.M. Ting, Z.-H. Zhou, Isolation-based anomaly detection. ACM Trans. Knowl. Discov. Data (TKDD) **6**(1), 1–39 (2012)
57. I. Chaikovsky, A. Popov, D. Fogel, A. Kazmirchyk, Development of AI-based method to detect the subtle ECG deviations from the population ECG norm. Eur. J. Prev. Cardiol. **28**(Supplement_1), zwab061–zwab229 (2021)
58. T. Reichlin, R. Abächerli, R. Twerenbold, M. Kühne, B. Schaer, C. Müller, C. Sticherling, S. Osswald, Advanced ECG in 2016: Is there more than just a tracing? Swiss Med. Wkly. **146**, w14303 (2016)
59. R. Cuocolo, T. Perillo, E. De Rosa, L. Ugga, M. Petretta, Current applications of big data and machine learning in cardiology. J. Geriatr. Cardiol. **16**(8), 601–607 (2019)
60. J. Petch, S. Di, W. Nelson, Opening the black box: The promise and limitations of explainable machine learning in cardiology. Can. J. Cardiol. **38**(2), 204–213 (2022)
61. A.A. Mahayni, Z.I. Attia, J.R. Medina-Inojosa, M.F.A. Elsisy, P.A. Noseworthy, F. Lopez-Jimenez, S. Kapa, S.J. Asirvatham, P.A. Friedman, J.A. Crestenallo, M. Alkhouli, Electrocardiography-based artificial intelligence algorithm aids in prediction of long-term mortality after cardiac surgery. Mayo Clin. Proc. **96**(12), 3062–3070 (2021)
62. S. Raghunath, A.E. Ulloa Cerna, L. Jing, D.P. VanMaanen, J. Stough, D.N. Hartzel, B.K. Fornwalt, Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat. Med. **26**(6), 886–891 (2020)