



Universiteit  
Leiden  
The Netherlands

## Cardiac age detected by machine learning applied to the surface ECG of healthy subjects: creation of a benchmark

Wall, H.E.C. van der; Hassing, G.J.; Doll, R.J.; Westen, G.J.P. van; Cohen, A.F.; Selder, J.L.; ... ; Gal, P.

### Citation

Wall, H. E. C. van der, Hassing, G. J., Doll, R. J., Westen, G. J. P. van, Cohen, A. F., Selder, J. L., ... Gal, P. (2022). Cardiac age detected by machine learning applied to the surface ECG of healthy subjects: creation of a benchmark. *Journal Of Electrocardiology*, 72, 49-55. doi:10.1016/j.jelectrocard.2022.03.001

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3420602>

**Note:** To cite this publication please use the final published version (if applicable).



## Cardiac age detected by machine learning applied to the surface ECG of healthy subjects: Creation of a benchmark

Hein E.C. van der Wall, MSc<sup>a,b,\*</sup>, Gert-Jan Hassing, MSc<sup>a</sup>, Robert-Jan Doll, PhD<sup>a</sup>, Gerard J.P. van Westen, PhD<sup>b</sup>, Adam F. Cohen, MD PhD<sup>a,b,c</sup>, Jasper L. Selder, MD PhD<sup>d</sup>, Michiel Kemme, MD PhD<sup>d</sup>, Jacobus Burggraaf, MD PhD<sup>a,b,c</sup>, Pim Gal, MD PhD<sup>a,c</sup>

<sup>a</sup> Centre for Human Drug Research, The Netherlands

<sup>b</sup> Leiden Academic Centre for Drug Research, The Netherlands

<sup>c</sup> Leiden University Medical Center, The Netherlands

<sup>d</sup> Amsterdam University Medical Center, The Netherlands

### ARTICLE INFO

#### Keywords:

Aging  
ECG  
Machine learning  
Healthy volunteers  
Artificial intelligence

### ABSTRACT

**Objective:** The aim of the present study was to develop a neural network to characterize the effect of aging on the ECG in healthy volunteers. Moreover, the impact of the various ECG features on aging was evaluated.

**Methods & results:** A total of 6228 healthy subjects without structural heart disease were included in this study. A neural network regression model was created to predict age of the subjects based on their ECG; 577 parameters derived from a 12-lead ECG of each subject were used to develop and validate the neural network; A tenfold cross-validation was performed, using 118 subjects for validation each fold. Using SHapley Additive exPlanations values the impact of the individual features on the prediction of age was determined. Of 6228 subjects tested, 1808 (29%) were females and mean age was 34 years, range 18–75 years. Physiologic age was estimated as a continuous variable with an average error of  $6.9 \pm 5.6$  years ( $R^2 = 0.72 \pm 0.04$ ). The correlation was slightly stronger for men ( $R^2 = 0.74$ ) than for women ( $R^2 = 0.66$ ). The most important features on the prediction of physiologic age were T wave morphology indices in leads V4 and V5, and P wave amplitude in leads AVR and II.

**Conclusion:** The application of machine learning to the ECG using a neural network regression model, allows accurate estimation of physiologic cardiac age. This technique could be used to pick up subtle age-related cardiac changes, but also estimate the reversing of these age-associated effects by administered treatments.

© 2022 Published by Elsevier Inc.

### Introduction

Surface electrocardiograms (ECGs) are used frequently in routine clinical care, but also in investigational studies examining the effects of pharmacological and non-pharmacological treatments on the heart. Readout measures include the RR interval, PR interval, QRS duration and (corrected) QT interval. The ECG has long offered valuable insights into cardiac and non-cardiac health and disease, its interpretation requires considerable human expertise. Typically, the pharmacological treatment effects are mediated by recognized channels on the cardiac surface [1]. However, there are cardiac effects that require a longer period of time to become visible on the surface ECG, such as aging induced cardiac fibrosis, and it is largely unknown if these subtle effects can be visualized on a surface ECG [2,3]. Advanced AI methods, such as deep-learning convolutional neural networks, have enabled rapid,

human-like interpretation of the ECG, while signals and patterns largely unrecognizable to human interpreters can be detected by multilayer AI networks with precision, making the ECG a powerful, non-invasive biomarker [4].

There has been a number of recent investigations regarding the prediction of physiological age using medical records, vital signs and laboratory data, or epigenetic changes [5–7]. The likelihood of having a “normal” ECG decreases with age. The most common findings are left ventricular hypertrophy pattern, leftward axis deviation and QRS widening [8]. Some of these abnormalities were significantly associated with all-cause death [9,10]. These investigations also indicated the existence of a gap between predicted physiological age and actual chronological age. Exploration of this gap is clinically important as a serious gap difference has been shown to be associated with higher risks of all-cause mortality, cardiovascular disease, obesity, earlier menopause, and frailty [6,11–15]. Various previous studies have already shown that the 12-lead ECG can be a reliable tool to estimate physiological aging [6,11–20].

\* Corresponding author at: Centre for Human Drug Research, Zernikedreef 8, 2333 CL Leiden, The Netherlands.

E-mail address: [Hvdwall@chr.nl](mailto:Hvdwall@chr.nl) (H.E.C. van der Wall).

Previous studies have applied artificial intelligence to the raw ECG data, allowing estimation of physiologic ECG age, which was found to reflect aging and comorbidities [21]. However, these algorithms were based on large hospital datasets, thus including patients that may have disease-induced abnormalities in their ECGs, which makes the outcome difficult to interpret when applied to a healthy volunteer. Therefore, the aim of the present analysis was to develop a neural network in healthy volunteers to characterize the effect of aging on the ECG.

## Methods and materials

### Population

All data were collected at the Centre for Human Drug Research in Leiden, the Netherlands, a clinical research organization specialized in early phase drug development studies. Data collected during the mandatory medical screening to verify study eligibility for enrolment in the early phase drug development studies as a volunteer between 2010 and 2019 were included in the present analysis. The medical screening consisted of a single visit to the clinical unit where a detailed anamnesis, a physical examination, vital signs including blood pressure, temperature, weight and height measurement, body mass index calculation, and a twelve-lead ECG were recorded. Ethical approvals from the Medical Ethical Review Committee for the included studies were acquired and informed consent documents were signed by the volunteers prior to any data collection. The present study was performed in accordance to local regulations. All activities were performed in accordance with applicable standard operating procedures.

### Data collection for the model

ECG parameters of 6228 subjects with an age between 18 and 75 years were included in the present study. All subjects that were used in this dataset were considered healthy, none of them had known cardiovascular risk factors, and all ECGs were considered normal, or abnormal but without clinical significance. The ECG reviews were performed manually, using standard MUSE cardiology terms. From each subject ECG, 574 features were extracted by the MUSE system. Additionally, gender was used as a feature. The age of the subjects was rounded in whole years. At least ten ECGs were available for each age.

In supplementary Tables 1 and 2, 54 features present in most leads and other ECG features used for the machine learning model are shown, respectively. In addition, gender of each subject was also included in the model.

### Data pre-processing and selection

As validation set two subjects of each age were kept apart as final test set. The rest of the data was used as the training set.

To create a balanced training set the Synthetic Minority Oversampling Technique (SMOTE) algorithm was applied on the training set to create 'synthetic' subjects for the less populated age groups based on the values in the concerning age groups [22].

### Machine learning

A neural network was used as a machine learning model. The keras module v. 2.4.3 in python 3.8.5 was used to build a model. Before training, internal cross validation (three-fold) within the training set was used to optimize the model. The network was optimized for number of layers, number of nodes per layer, activation function per layer for each layer and learning rate. A batch size of 300 was used. The number of epochs (defined as the number of cycles through the full training dataset) for internal validation was determined based on validation performance in the internal validation set. The number of epochs for final validation was based on the median of the optimal number of epochs

for the internal cross validations. This process of optimization, training, and validation was repeated 10 times with different training and test sets. The optimal models were evaluated on the test set with the  $R^2$  score and mean absolute error. We also evaluated the model performance with respect to gender.

To gain insight into the impact of the individual features on the predicted age, each fold SHapley Additive exPlanations (SHAP) values were calculated [23] based on the training set. The importance of the features was validated by means of permutation importance (defined as the decrease in a model score when a single feature value is randomly shuffled) [24].

## Results

The clinical characteristics of the 6228 included subjects are displayed in Table 1. The study population was divided into ten chronological age groups of 6 years, starting from the age of 18 years. Each age group contained at least 194 subjects, and younger age groups comprised up to 2282 subjects. A total of 1808 (29%) volunteers were female.

The relation between the (predicted) physiologic age and the chronological age was assessed in 10 sets of 116 subjects. In Fig. 2a, the relation between predicted physiologic age and chronological age of all 10 test sets is shown. The average relationship of the models showed an  $R^2$  of  $0.72 \pm 0.04$  (mean  $\pm$  SD). The mean absolute error of all predictions was  $6.9 \pm 5.5$  years. The average predicted physiologic age was 0.3 years younger than the average chronological age of the subjects. The median deviation of all predicted ages was 5.6 years from the actual age, indicating that half of the predictions was within the range of 5.6 years of chronological age.

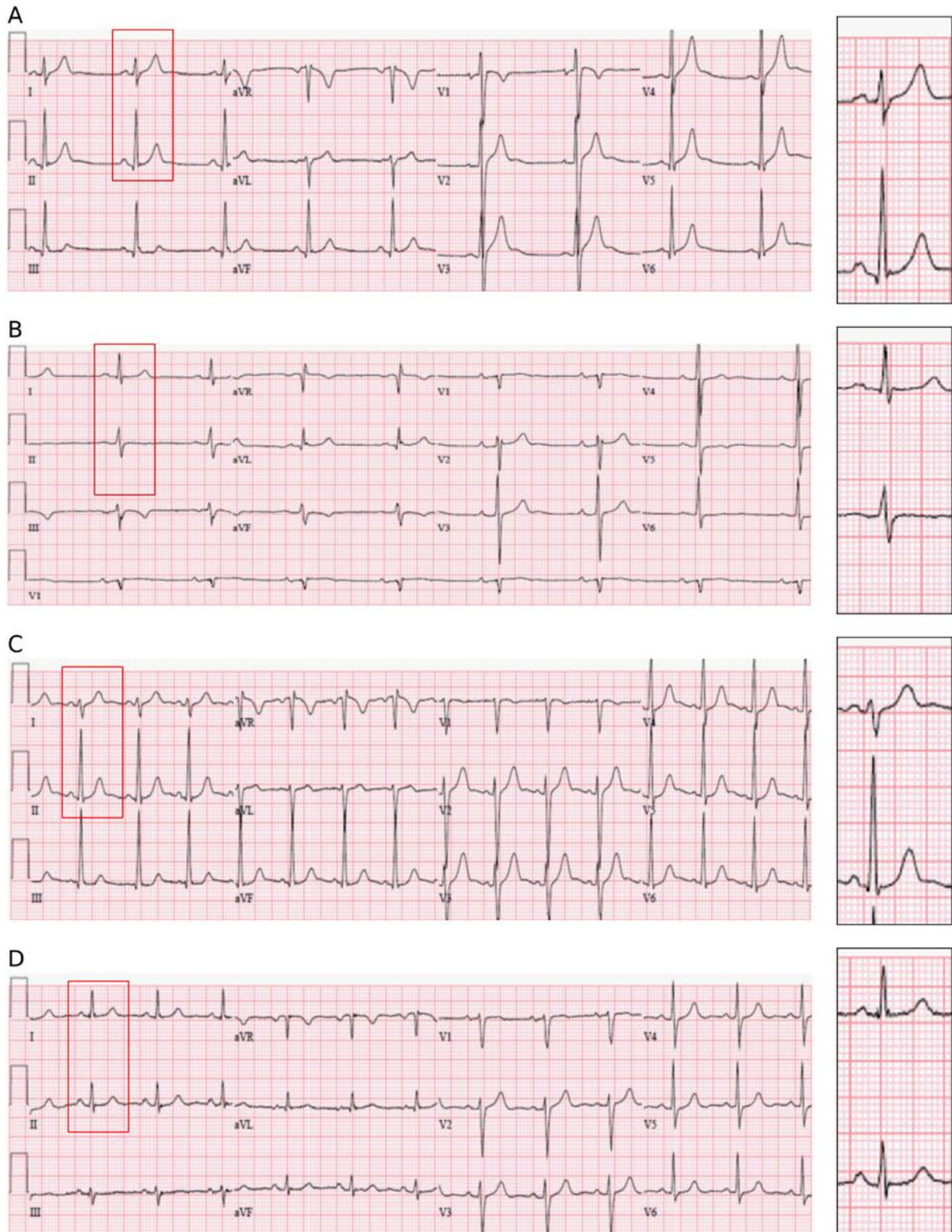
ECG examples of a young 18 year old male (1A) and an elderly 74 year old male (1B) are shown in Fig. 1. Fig. 1C shows an ECG of a young 19 year old female and Fig. 1D shows an ECG of an elderly 74 year old female subject. Several differences between the young and older healthy subjects were discernable. In elderly persons the heart rate was lower, the T wave had a lower (absolute) amplitude in leads I,II,III,AVR, and AVL and the P-wave duration seemed shorter. However, these ECG differences showed considerable variations in the healthy population.

The average predicted age of all subject is presented in Fig. 2b. The average predicted age of the 20 subjects per chronological age had a mean absolute error of  $3.4 \pm 3.0$  years ( $R^2 = 0.93$ ). For subjects between 30 and 60 years old the mean absolute error of the average predicted age per chronological age was  $1.6 \pm 1.1$  years.

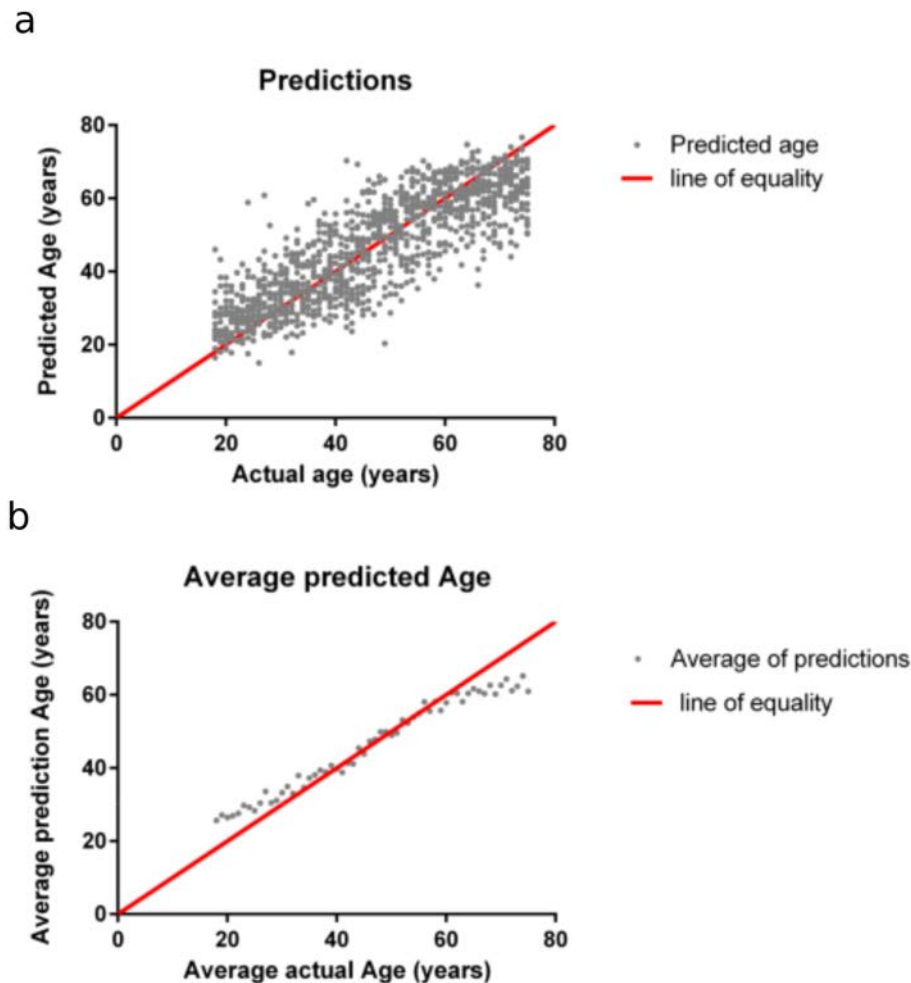
In order to study gender differences, the predicted physiological ages of the male and female subjects in the test sets were separated and are presented in Fig. 3. The predicted ages of the male subjects were more accurate ( $R^2 = 0.74$ ) than the predictions of the female subjects ( $R^2 = 0.64$ ). The mean absolute error in women of the predictions was  $7.5 \pm 5.9$  years, significantly higher than that in men ( $6.8 \pm 5.3$  years,  $p = 0.03$ ).

**Table 1**  
Age and gender characteristics of the 6228 healthy subjects.

Subject age (years)	N	% female
18–23	2282	29
24–29	1563	26
30–35	449	20
36–40	247	17
41–46	245	24
46–52	339	35
53–57	241	38
58–63	194	42
63–69	393	40
69–75	275	35



**Fig. 1.** ECG samples of young and elderly male and female healthy subjects. A: ECG of a young 18 year old male subject. B: ECG of an elderly 74 year old male subject. C: ECG of a young 19 year old female subject. D: ECG of an elderly 74 year old female subject. to the right of each EKG, the heartbeat in leads I and II marked in red is enlarged.



**Fig. 2.** Relationship between predicted) physiologic age and chronological age for 1180 healthy adults (10 test sets, Fig. 2a, top). The average predictions for each age are shown in Fig. 2b (bottom).

Fig. 4 shows the SHAP values of the 40 most important ECG features used in the prediction model. So, the impact of each individual feature on the model output and physiologic aging can be seen. Some of the most important features on the prediction of physiologic age were T top abnormalities in leads V4 and V5, P top amplitude in leads AVR and II and atrial rate.

An increase of P peak amplitude in lead II for example, indicates a younger physiological age (a long red bar to the left). A longer PR interval both indicate an older physiologic age (longer red bar to the right). A higher atrial rate indicates a younger physiologic age (large red bar to the left). The impact of gender was only of minor importance with SHAP values ranging from  $-1.2$  to  $0.9$ . The order of the feature permutation importance is similar to the order of the SHAP values, confirming the impact of the features.

## Discussion

In this study we developed machine learning models that allow accurate prediction of physiologic cardiac age of healthy subjects based on 12-lead surface ECG parameters. Using a neural network we were able to estimate the age of a healthy subject with an error of 6.9 years and to analyze the impact of the ECG features. As the models were trained using only healthy subjects, we can assume that the predicted actual age is equal to the cardiac age. The created models of the present study may serve as a benchmark for testing the effects of new pharmacological drugs on potential decline or improvement of physiologic health of the heart.

## Application of machine learning

Attia et al. recently sought to determine whether the application of machine learning algorithms, including convolutional neural networks, to a large ECG patient data set would be capable of predicting age and sex reported by patients, independent of additional clinical data [21]. They further investigated whether discrepancies between ECG age and chronological age might be a marker of physiological health. When the convolutional neural network-predicted age exceeded a patient's actual age by at least 7 years, there was a higher incidence of cardiovascular comorbidities, potentially suggesting that the convolutional neural network-predicted age from 12-lead ECGs may correlate with physiological health. Their findings suggested that physiological age is distinct from chronological age, and may have useful clinical applications. For example, if a patient's biologic age is 60 but their ECG age predicts that they are 70, it may indicate underlying cardiovascular disease and potential risk. A limitation of their study was, as also recognized by the authors, that all individuals included were patients, and thus an ECG was obtained for a certain clinical indication. It was questioned by the authors whether their results are similarly accurate among an ostensibly healthy population is unknown, and revalidation in such a cohort will therefore be critical.

The same holds true for the study by Hirota et al., who studied biological age, physiological age, and all-cause mortality by 12-lead ECG in patients without structural heart disease [25]. Their data showed that the gap between ECG-predicted physiological and biological age allowed estimation of increased risk of all-cause mortality. Although

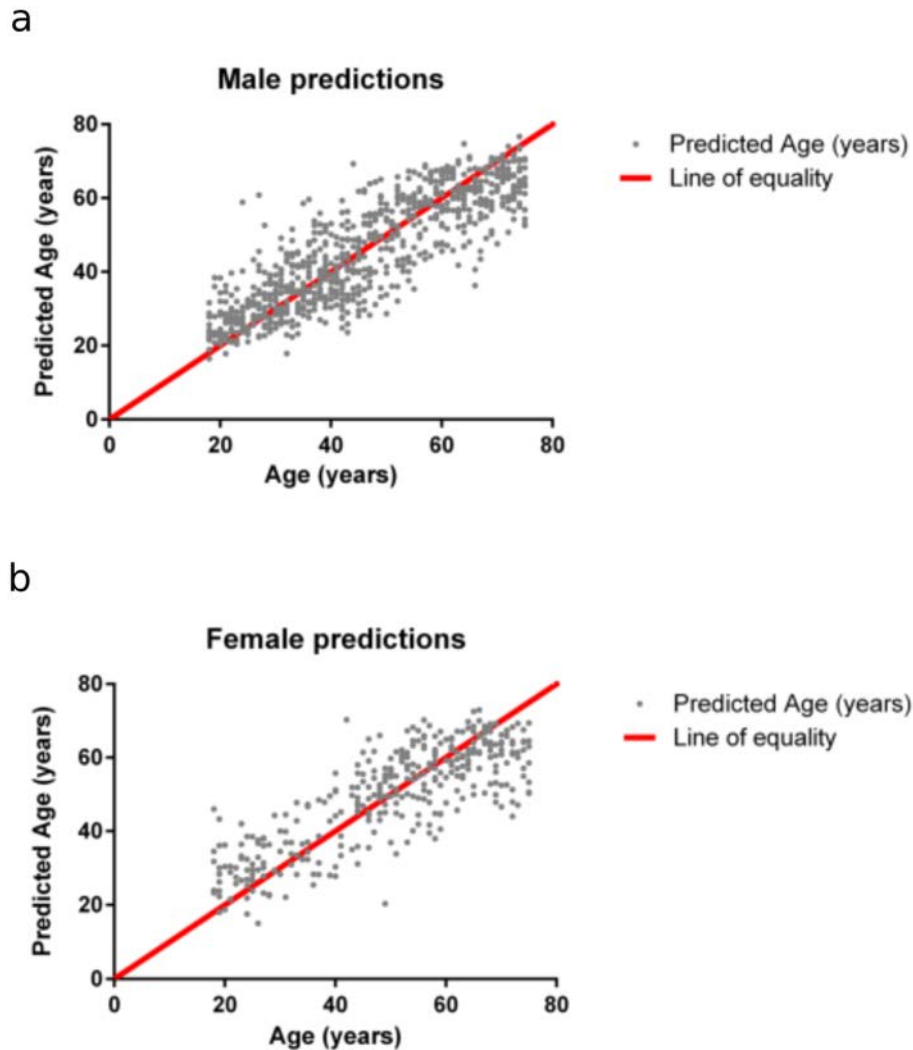


Fig. 3. Predictions of 819 male subjects (a, top) and 361 female subjects (b, bottom).

their study subjects were assumed to have no structural heart diseases, it was stated by the authors that it will be necessary to validate the results of their study in populations of healthy subjects. In our study, we only studied healthy individuals, giving the advantage of being a much needed benchmark study, which enables the validation of future studies in patients versus our data.

#### Performance of the model

The relation between chronological and predicted physiologic age was associated with an  $R^2$  of 0.72. Although with a smaller dataset than used by Attia et al., our predictions have a similar performance, probably because of the healthy population in our study, which we expect reduces the variability of the association. Given the large number of influencing factors that can affect ECG parameters the  $R^2$  of 0.72 of our models seems sufficient to detect a pharmacodynamic effect in a cohort of subjects. Use of the entire dataset with a larger number of subjects may improve future performance of the model.

In the present study, the impact of physiologic aging on the various ECG features was analyzed using SHAP values. Several changes are clearly visible in the ECG figures. Some of these are already well known in clinical practice, such as prolongation of PR and QT interval and deceleration of heart rate [16]. Other changes, however, could only be recognized by using machine learning, while these may be

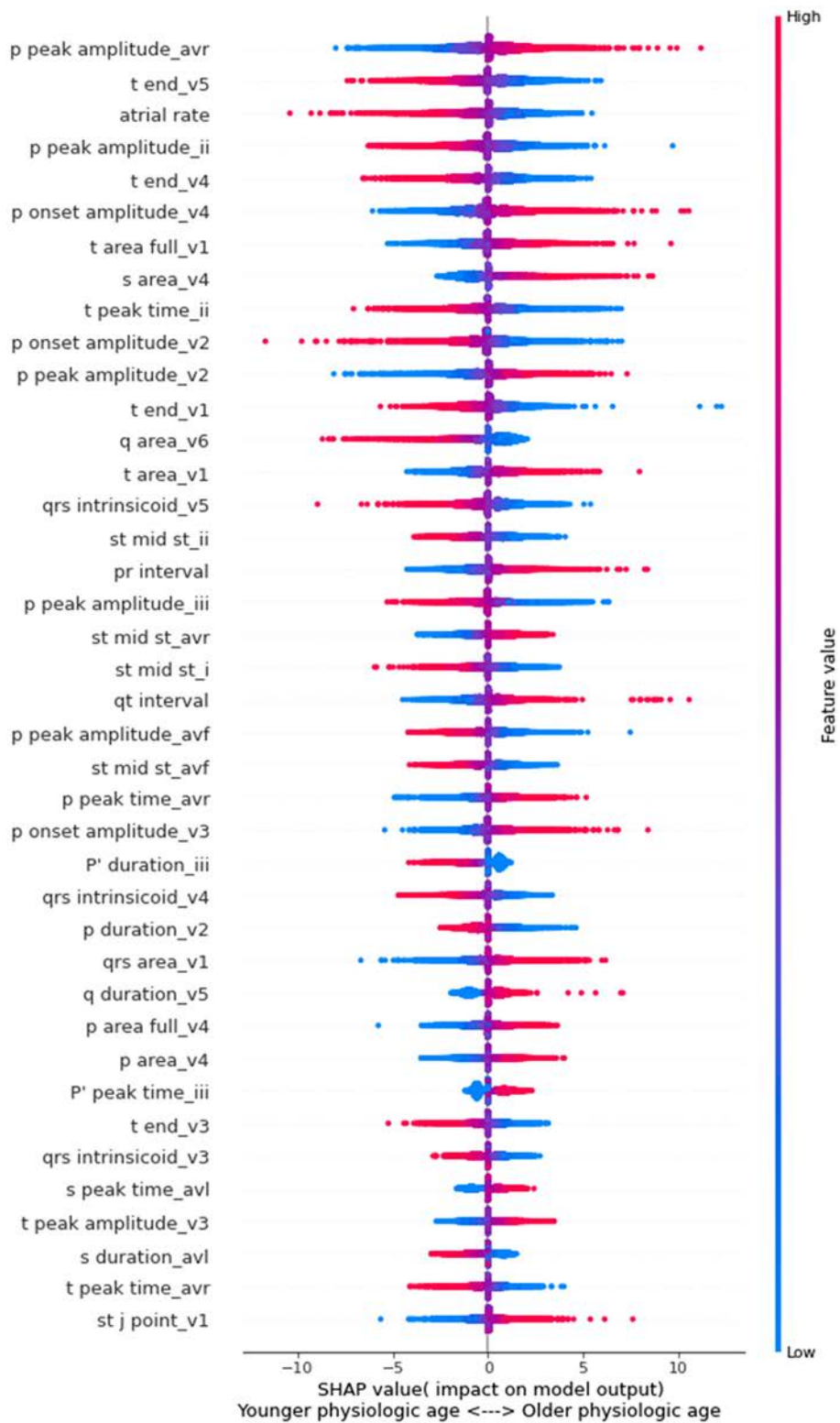
evenly important. Moreover, when multiple features change at the same time, it becomes difficult to judge whether the change in the ECG is good or bad without using machine learning. By means of machine learning techniques a combination of various ECG changes allows a more accurate insight into the physiologic health changes of the heart.

#### Gender differences

The accuracy of predicting physiologic age was found to be higher in males than in the female subjects. This may be due to the somewhat smaller female study population, but it may also reflect the atypical ECG repolarization patterns which are known to occur frequently in women [26]. The SHAP values show that impact of gender on physiologic age prediction was only of minor importance. Future studies, analyzing sex- and age- interaction could clarify this.

#### Pharmaceutical drug testing and potential implications

The prediction of the physiologic age for one single person is less relevant in this model. However for larger groups or cohorts of multiple subjects, the prediction could be more accurate. For example, for a group of 30 test subjects, the average deviation is only less than two years from average physiologic age. Therefore, our models could be particularly suitable as benchmark for testing new pharmaceutical drugs or



**Fig. 4.** SHAP values of the 40 most important features for predicting physiologic age. High values of the features are represented in red. Low values are represented in blue. On the x-axis, the predicted physiologic age. Shorter bars mean less impact on physiologic aging.

other interventions which may have impact on cardiac health in the near future. Differences between physiologic ECG age and chronological age have been shown to predict all-cause and cardiovascular mortality and reflect physiologic age, cardiovascular health and long term outcomes [27]. It has also been found that a difference in predicted

(cardiac) age and chronological age (higher cardiac age) was greater in patients with peripheral microvascular endothelial function [28]. Additionally, patients with an ECG-age more than 8 years greater than chronological age had a higher mortality rate [29]. Our models, trained with healthy subjects, would therefore be a good benchmark and

could be used to predict the mean cardiac age of a cohort before (baseline) and after an intervention to determine its effect on the heart.

The proper use of a model - trained on the entire dataset - in early drug development can provide important information that can be used to make a go/no-go decision regarding further development of new drugs. Similarly, this can be used to guide the decision-making process regarding the dosage range to be used in phase II studies, determining a therapeutic window, and even identifying the target study population [30]. This way novel pharmacological drugs could be tested for effect on cardiac physiologic aging in the early phase of development.

### Limitations

Our population consisted of only 29% female subjects. This may have influenced the accuracy of the model, but SHAP value analysis showed that gender only had a minimal impact on the predictions of physiologic age.

No data from children were available for the present study. At the moment, legal age determination of children is a common probable and depends on imaging techniques that use radiation. Future studies, including age determination in children using ECG could be very useful.

The models have been trained on a limited range of ages. Therefore, the models are limited to predict inside this range, which means that wrong predictions among the youngest participants are always higher and among the oldest participants always lower. This is clearly visible in Fig. 2. Future studies, including data with a bigger age range, might reduce these limitations.

ECG changes do not need to have a purely cardiac cause, but they may also be caused by effects of age on the position of the heart in the thorax, the presence of fat layers around the heart, and the shape of the thorax shape. Therefore, the found relationship does not necessarily mean older heart per se, but can also mean an older body.

### Conclusion

The application of machine learning to the ECG using a neural network regression model, allows estimation of physiologic cardiac age. This technique could be used to pick up subtle age-related cardiac changes, but also estimate the reversing of these age-associated effects by administered treatments.

Disclosure: No disclosures for all authors.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jelectrocard.2022.03.001>.

### References

- [1] van Dam PM, et al. The relation of 12 lead ECG to the cardiac anatomy: the normal CineECG. *J Electrocardiol.* 2021;69:67–74.
- [2] Biernacka A, Frangogiannis NG. Aging and cardiac fibrosis. *Aging Dis.* 2011;2(2):158.
- [3] Hayashi H, et al. Aging-related increase to inducible atrial fibrillation in the rat model. *J Cardiovasc Electrophysiol.* 2002;13(8):801–8.
- [4] Siontis KC, et al. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol.* 2021;18(7):465–78.
- [5] Wang F, Syeda-Mahmood T, Beymer D. Information extraction from multimodal ECG documents. 2009 10th international conference on document analysis and recognition. IEEE; 2009.
- [6] Roetker NS, et al. Prospective study of epigenetic age acceleration and incidence of cardiovascular disease outcomes in the ARIC study (Atherosclerosis Risk in Communities). *Circulation: Genomic Precision Med.* 2018;11(3):e001937.
- [7] Kistler PM, et al. Electrophysiologic and electroanatomic changes in the human atrium associated with age. *J Am Coll Cardiol.* 2004;44(1):109–16.
- [8] Vicent L, Martínez-Sellés M. Electrocardiogeriatrics: ECG in advanced age. *J Electrocardiol.* 2017;50(5):698–700.
- [9] Lu T-P, et al. Develop and apply electrocardiography-based risk score to identify community-based elderly individuals at high-risk of mortality. *Front Cardiovasc Med.* 2021;8.
- [10] Molander U, et al. ECG abnormalities in the elderly: prevalence, time and generation trends and association with mortality. *Aging Clin Exp Res.* 2003;15(6):488–93.
- [11] Breitling LP, et al. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. *Clin Epigenetics.* 2016;8(1):21.
- [12] Perna L, et al. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics.* 2016;8(1):64.
- [13] Wang Z, et al. Predicting age by mining electronic medical records with deep learning characterizes differences between chronological and physiological age. *J Biomed Inform.* 2017;76:59–68.
- [14] Horvath S, et al. Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci.* 2014;111(43):15538–43.
- [15] Levine ME, et al. Menopause accelerates biological aging. *Proc Natl Acad Sci.* 2016;113(33):9327–32.
- [16] Rijnbeek PR, et al. Normal values of the electrocardiogram for ages 16–90 years. *J Electrocardiol.* 2014;47(6):914–21.
- [17] Macfarlane P, et al. Effects of age, sex, and race on ECG interval measurements. *J Electrocardiol.* 1994;27:14–9.
- [18] Mason JW, Hancock EW, Gettes LS. Recommendations for the standardization and interpretation of the electrocardiogram: part II: electrocardiography diagnostic statement list: a scientific statement from the American Heart Association electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology. *Circulation.* 2007;115(10):1325–32.
- [19] Kligfield P, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American Heart Association electrocardiography and arrhythmias committee, council on clinical cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol.* 2007;49(10):1109–27.
- [20] Khane RS, Surdi AD, Bhatkar RS. Changes in ECG pattern with advancing age; 2011..
- [21] Attia ZI, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ Arrhythm Electrophysiol.* 2019;12(9):e007284.
- [22] Chawla NV, et al. SMOTE: synthetic minority over-sampling technique. *J Artificial Intell Res.* 2002;16:321–57.
- [23] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems;* 2017.
- [24] Altmann A, et al. Permutation importance: a corrected feature importance measure. *Bioinformatics.* 2010;26(10):1340–7.
- [25] Hirota N, et al. Prediction of biological age and all-cause mortality by 12-lead electrocardiogram in patients without structural heart disease. *BMC Geriatr.* 2020;21(460).
- [26] Okin PM. Electrocardiography in women: taking the initiative. *Am Heart Assoc.* 2006;113(4):464–6.
- [27] Ladejobi A, et al. ECG-derived age and survival: validating the concept of physiologic age detected by ECG using artificial intelligence. *J Am Coll Cardiol.* 2020;75(11 Supplement 1):3469.
- [28] Toya T, et al. Vascular aging detected by peripheral endothelial dysfunction is associated with ECG-derived physiological aging. *J Am Heart Assoc.* 2021;10(3):e018656.
- [29] Lima EM, et al. *Deep neural network estimated electrocardiographic-age as a mortality predictor.* medRxiv; 2021..
- [30] Groeneveld GJ, Hay JL, Van Gerven JM. Measuring blood-brain barrier penetration using the NeuroCart, a CNS test battery. *Drug Discov Today Technol.* 2016;20:27–34.