


RESEARCH ARTICLE

Open Access



# Prediction of biological age and all-cause mortality by 12-lead electrocardiogram in patients without structural heart disease

Naomi Hirota<sup>\*</sup> , Shinya Suzuki, Takuto Arita, Naoharu Yagi, Takayuki Otsuka and Takeshi Yamashita

## Abstract

**Background:** There is a well-established relationship between 12-lead electrocardiogram (ECG) and age and mortality. Furthermore, there is increasing evidence that ECG can be used to predict biological age. However, the utility of biological age from ECG for predicting mortality remains unclear.

**Methods:** This was a single-center cohort study from a cardiology specialized hospital. A total of 19,170 patients registered in this study from February 2010 to March 2018. ECG was analyzed in a final 12,837 patients after excluding those with structural heart disease or with pacing beats, atrial or ventricular tachyarrhythmia, or an indeterminate axis (R axis > 180°) on index ECG. The models for biological age were developed by principal component analysis (BA) and the Klemera and Doubal's method (not adjusted for age [BA<sub>E</sub>] and adjusted for age [BA<sub>EC</sub>]) using 438 ECG parameters. The predictive capability for all-cause death and cardiovascular death by chronological age (CA) and biological age using the three algorithms were evaluated by receiver operating characteristic analysis.

**Results:** During the mean follow-up period of 320.4 days, there were 55 all-cause deaths and 23 cardiovascular deaths. The predictive capabilities for all-cause death by BA, BA<sub>E</sub>, and BA<sub>EC</sub> using area under the curves were 0.731, 0.657, and 0.685, respectively, which were comparable to 0.725 for CA ( $p = 0.760, 0.141, \text{ and } 0.308$ , respectively). The predictive capabilities for cardiovascular death by BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.682, 0.685, and 0.692, respectively, which were also comparable to 0.674 for CA ( $p = 0.775, 0.839, \text{ and } 0.706$ , respectively). In patients aged 60–74 years old, the area under the curves for all-cause death by BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.619, 0.702, and 0.697, respectively, which tended to be or were significantly higher than 0.482 for CA ( $p = 0.064, 0.006, \text{ and } 0.005$ , respectively).

**Conclusion:** Biological age by 12-lead ECG showed a similar predictive capability for mortality compared to CA among total patients, but partially showed a significant increase in predictive capability among patients aged 60–74 years old.

**Keywords:** Electrocardiogram, Biological age, All-cause death, Cardiovascular death

\* Correspondence: [n-hirota@cvi.or.jp](mailto:n-hirota@cvi.or.jp)

Department of Cardiovascular Medicine, The Cardiovascular Institute, 3-2-19 Nishiazabu, Minato-Ku, Tokyo 106-0031, Japan



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Age is a strong risk factor for mortality. However, chronological age (CA) itself may not be a reliable indicator of functional deterioration because aging can be heterogeneous, with a balance between exposure to damaging properties and resiliency [1, 2]. The concept of biological age was developed to represent the actual status of individual aging. Biological age is estimated as a single variable using complex equations based on multiple biomarkers, which include physical, physiological, or biochemical indicators of individual health status [2–4]. However, a simple, non-invasive, and cost-effective method for estimating biological age is required for its practical use.

Electrocardiography (ECG) is widely used to detect or evaluate the risk of cardiac diseases. ECG parameters can be affected by age, gender, and individual physical conditions [5, 6], especially those involving the circulatory and respiratory systems. The potential mechanisms underlying these effects include changing topography of the heart in relation to the thorax and diaphragm, modification of the various components of the volume conductor (skin, subcutaneous fat, and lung parenchyma), or alterations in cardiac configuration and intracardiac conduction [5, 7].

Aging is also a key factor underlying these electrophysiological and electroanatomical changes [5, 6, 8]. Several studies have utilized ECG to predict biological age or 'heart age' [9–11]. A discrepancy between biological age estimated by ECG and the actual CA was also reported [9], which may relate to differences in the physical conditions of the individuals and the presence of cardiovascular diseases (CVDs). Importantly, this concept may be utilized to provide a simple method for screening patients' health status. Nevertheless, for actual clinical use it would be important to examine the effects of ethnicity on performance of the biological age prediction models [11]. Furthermore, the majority of reported models only examine several representative ECG parameters and with a linear regression model [10, 12], with only one report utilizing artificial intelligence modeling [9].

In the present study, we developed a prediction model for biological age that incorporated hundreds of automatically-measured ECG parameters assessed using the principal component analysis (PCA) algorithm [13] and the Klemra and Doubal's method (KDM) [14] from a single-center cohort in a Japanese cardiovascular hospital. The aim of this study was to evaluate the hypothesis that the predictive capability of biological age for mortality was higher than that of CA.

## Methods

### Study population

The Shinken Database [15] includes all new patients visiting the Cardiovascular Institute, a cardiology

specialized hospital in an urban area of Tokyo, Japan. This single hospital-based database was established in June 2004 to investigate the prevalence and prognosis of various types of CVDs. To investigate the new appearance of CVDs, patients who visited our hospital but who were not diagnosed with CVDs at baseline were also included in the cohort. Patients have been continually registered into the database annually and the registration is ongoing (32,570 patients have been registered up to March 2018). Foreign travelers and patients with active cancer were excluded because of the difficulty in evaluating long-term follow-up. The patients seen included both local residents and patients referred from other clinics for treatment of CVDs. The attending physicians were all cardiologists or cardiothoracic surgeons.

We used computerized ECG records, which have been available in our database since February 2010. From a total of 32,570 patients in the Shinken Database, we extracted 19,170 patients registered between February 2010 and March 2018. After excluding patients with structural heart diseases ( $n = 4915$ ), patients aged  $< 20$  years old or  $> 90$  years old ( $n = 168$ ), and patients with an index ECG showing an indeterminate axis (R axis  $> 180^\circ$ ;  $n = 76$ ), pacing beats ( $n = 102$ ), or atrial or ventricular tachyarrhythmia ( $n = 1763$ ), a total of 12,837 patients were included in the present study.

### Data collection at initial visit

After ECG and chest X-ray were performed, cardiovascular status was evaluated using data from an echocardiogram, exercise test, 24-h Holter recording, and blood laboratory tests at the discretion of the attending physician. In addition to gender, age, height, and weight, we collected data on CVDs, including heart failure (New York Heart Association class  $\geq 2$ ), valvular heart disease (moderate or severe stenosis or regurgitation on echocardiogram), coronary artery disease (diagnosed on angiogram or scintigram), hypertrophic and dilated cardiomyopathy (diagnosed on echocardiography or magnetic resonance imaging), congenital heart disease (diagnosed on echocardiography), and history of disabling cerebral infarction or transient ischemic attack (diagnosed on computed tomography or magnetic resonance imaging). Cardiovascular risk factors were defined as hypertension, use of antihypertensive agents, systolic blood pressure  $\geq 140$  mmHg, or diastolic blood pressure  $\geq 90$  mmHg; diabetes mellitus, use of oral hypoglycemic agents or insulin, or glycosylated hemoglobin  $\geq 6.5\%$ ; dyslipidemia, use of statin or drugs for lowering triglyceride, low-density lipoprotein  $\geq 140$  mg/dL, high-density lipoprotein  $< 40$  mg/dL, or triglyceride  $\geq 150$  mg/dL; and chronic kidney disease or estimated glomerular filtration rate (eGFR)  $< 60$  mL/min/1.73 m<sup>2</sup>. The eGFR was calculated using the Japanese coefficient

for the modified isotope dilution mass spectrometry-traceable 4-variable Modification of Diet in Renal Disease study equation ( $eGFR = 194 \times SCr^{-1.004} \times Age^{-0.287} \times 0.739$  [if female]). Body mass index was calculated as weight in kilograms divided by height in meters squared.

**Patient follow-up**

The health status and the incidence of cardiovascular events and mortality were maintained in the database via linking to hospital medical records and by prognosis study documents sent yearly to patients who stopped hospital visits or who were referred to other hospitals. In the present study, we included follow-up data until March 2019 and excluded follow-up data of >3 years after the initial visit to avoid an imbalance of the follow-up period due to the different registration years (between 2010 and 2018).

**Parameters obtained from ECG**

The 12-lead ECG was recorded for 10 s in the supine position using an ECG machine (GE CardioSoft v6.71 and MAC 5500 HD; GE Healthcare, Chicago, IL, USA) at a sampling rate of 500 Hz. Data were stored using the MUSE data management system. Automatic analysis of 639 parameters from the computerized raw ECG data was performed by the GE system. Of these parameters, 201 (nine not lead-specific and 192 [16 × 12 leads] lead-specific) were temporally stored datasets that included the relative coordinate points (i.e., the start point of the P-wave) and calculated values similar to the original parameters (i.e., of the corrected QT [QTc] parameters, the QTc calculation [QTc Bazett] was used while the QTc Framingham and QTc Fridericia were excluded). The remaining 438 parameters (six not lead-specific and 432 [36 × 12 leads] lead-specific) were used in the final analysis (Table 1).

**Evaluation and statistical analysis**

Statistical analyses were performed using SPSS v26.0 (IBM, Chicago, IL, USA) and R v3.5.2 (The R Foundation for Statistical Computing). In all analyses,  $p < 0.05$  was considered statistically significant. Categorical and consecutive data are presented as number (%) and mean ± standard deviation.

**Parameter selection considering collinearity**

First, all ECG parameters were translated into standardized values. We then selected from the 438 ECG parameters using two steps, considering the correlation with CA and the collinearity between the ECG parameters. For step 1, the coefficients of correlation between CA and the 438 ECG parameters were evaluated. The ECG parameters

**Table 1** List of ECG parameters used in this study

Parameters available in MUSE database system: 639 parameters
Parameters used for analysis: 438 parameters
(1) Not lead-specific parameters: 6 parameters P-R Interval, P axis, QRS Duration, QTc Calculation (QTc Bazett), R axis, T axis
(2) Lead-specific parameters: 432 [36 × 12 leads] parameters ST at J Point, P Area, P' Area, P Area (Full), P Peak Time, P' Peak Time, P Peak Amplitude, P' Peak Amplitude, P Duration, P' Duration, QRS Area, Q Area, Q Peak Amplitude, Q Duration, R Area, R' Area, R Peak Time, R Duration, R' Duration, S Area, S' Area, S Peak Time, S Duration, S' Duration, T Area, T' Area, T Area (Full), T Peak Time, T Peak Amplitude, T' Peak Amplitude, T Duration, T' Duration, Minimum ST level, Max R Amplitude, Maximum ST level, Max S Amplitude
Parameters excluded: 201 parameters
(1) Not lead-specific parameters: 9 parameters P Onset, P Offset, QRS Count, QTc Framingham, QTc Fridericia, Q-T Interval, Q Onset, Q Offset, T Offset
(2) Lead-specific parameters: 192 [16 × 12 leads] parameters P Onset Amplitude, QRS Balance, QRS Deflection, QRS Intrinsicoid, Q Peak Time, R' Peak Time, R Peak Amplitude, R' Peak Amplitude, S' Peak Time, S Peak Amplitude, S' Peak Amplitude, T' Peak Time, T End, ST at End ST, ST at Mid ST, Special T

P', R', S', and T' indicated the second components of P, R, S and T wave, respectively, which could be positive or negative polarity

with a correlation coefficient ≥ 0.2 were selected. For step 2, from the parameters selected in step 1, the coefficient of correlation for any pairs of the parameter combinations (if the number was X then X × [X - 1] combinations) were evaluated, excluding pairs of each parameter with itself. The parameter pairs with a correlation coefficient ≥ 0.9 (defined as a 'strong correlation') were determined and the parameters that demonstrated the highest coefficient of correlation for CA in step 1 compared with any counterparts were selected for further analysis. Furthermore, the parameters not included in any pairs with a 'strong correlation' were selected for further analysis.

**Modeling of biological age using ECG parameters**

**PCA** Biological age by PCA (BA) was modeled using ECG parameters by two steps, as previously reported [13].

$$pre-BA = \left( \sum_{i=1}^m \sum_{j=1}^n \beta_{ij} \left( \frac{x_j - \bar{x}_j}{sd(x_j)} \right) p_i \right), \quad (1)$$

where m indicates the number of principal components, i indicates their individual orders, n indicates the number of ECG parameters, j indicates their individual orders, β indicates the coefficient in the PCA, x indicates each ECG parameter, and  $\bar{x}_j$  and  $sd(x)$  indicate the average value and the standard deviation of each ECG parameter, respectively. The  $p_i$  was calculated using the following formula:

$$p_i = \frac{R^2 \text{ in a univariate linear regression model with each principal component for CA}}{\text{(sum of } R^2 \text{ in the univariate linear regression models with each principal component for CA)}} \quad (2)$$

For step 2, BA was calculated using the following formula:

$$BA = \text{pre-BA} \times sd(CA) + \overline{CA} + (CA - \overline{CA}) \times (1-B), \quad (3)$$

where  $sd(CA)$  and  $\overline{CA}$  indicate the standard deviation and the average value of CA, respectively, and  $B$  indicates the standardized coefficient in the univariate linear regression analysis in which pre-BA and CA are the dependent and independent variables, respectively.

**KDM Biological age by KDM algorithm (BA<sub>E</sub> and BA<sub>EC</sub>)** was modeled using ECG parameters with the following equations [14].

$$BA_E = \frac{\sum_{j=1}^m (x_j - q_j) \left(\frac{k_j}{s_j^2}\right)}{\sum_{j=1}^m \left(\frac{k_j}{s_j}\right)^2}, \quad (4)$$

$$BA_{EC} = \frac{\sum_{j=1}^m (x_j - q_j) \frac{k_j}{s_j^2} + \frac{CA}{s_{BA}^2}}{\sum_{j=1}^m \left(\frac{k_j}{s_j}\right)^2 + \frac{1}{s_{BA}^2}}, \quad (5)$$

$$r_{char} = \frac{\sum_{j=1}^m \frac{r_j^2}{\sqrt{1-r_j^2}}}{\sum_{j=1}^m \frac{r_j}{\sqrt{1-r_j^2}}}, \quad (6)$$

$$s_{BA}^2 = \left( \frac{\sum_{j=1}^n (|BA_{Ei} - CA_i| - \sum_{i=1}^n |BA_{Ei} - CA_i|/n)^2}{n} \right) - \left( \frac{1-r_{char}^2}{r_{char}^2} \right) \times \left( \frac{(CA_{max} - CA_{min})^2}{12m} \right), \quad (7)$$

where  $k_j$  indicates the beta of an ECG parameter regressed on BA<sub>E</sub>,  $q_j$  indicates the beta of a constant regressed on BA<sub>E</sub>, and  $s_j$  indicates the root mean squared error of an ECG parameter regressed on BA<sub>E</sub>. However, given that BA<sub>E</sub> was not measurable, the root mean squared errors from the regressions between each ECG parameter and CA (rather than BA<sub>E</sub>) were used [16]. The value  $r_j^2$  indicates the variance explained by the regression of CA on  $m$  parameters.

**Evaluation of the predictability of biological age for mortality**

The predictive capabilities of CA, BA (by PCA), and BA<sub>E</sub> and BA<sub>EC</sub> (by KDM) for all-cause death and cardiovascular death were evaluated by the area under the curve (AUC) with the receiver operating curve. Patients were also divided into four CA categories of 20–39 years old, 40–59 years old, 60–74 years old, and ≥ 75 years old, and a similar evaluation was performed for each separate CA category. The comparison among CA and the biological age algorithms (BA, BA<sub>E</sub>, and BA<sub>EC</sub>) or the age categories were tested by the paired or the unpaired Delong’s test for two ROC curves [17], respectively.

**Results**

**Patient characteristics**

The study patients included 6897 men (53.7%) and the mean age was 55.5 ± 15.0 years. In men, the mean ages of alive and deceased patients were 54.1 ± 14.4 years and 70.9 ± 12.1 years, respectively. In women, the mean ages of alive and deceased patients were 56.9 ± 15.6 years and 70.1 ± 14.1 years, respectively. The patients’ characteristics are shown in Table S1 (see Additional file 1).

During the mean follow-up period of 320.4 days, all-cause death occurred in 55 patients (0.5 per 100 patient-years). Among the 55 all-cause death patients, 23 and 32 were cardiovascular deaths and non-cardiovascular deaths, respectively. The distributions of the deceased and alive patients are shown in Table 2.

**Parameter selection**

For step 1, among the 438 ECG parameters, the correlation coefficient with CA was ≥ 0.2 for 71 parameters in men and for 99 parameters in women; these parameters were selected for the next step. For step 2, the coefficients of correlation were evaluated for all pairs of the parameters selected from step 1 (men: 71 × 70 = 4970 combinations; women: 99 × 98 = 9702 combinations). For both men and women, all of the parameters selected from step 1 (i.e., 71 for men and 99 for women) had combinations with a correlation coefficient ≥ 0.9. From these ECG parameters, we selected 61 parameters for

**Table 2** The number of deceased subjects with respect to individual CA

	Alive	All-cause death	cardiovascular death	Non-cardiovascular death
Total	12,782	55	23	32
CA				
≤ 39 years	2138	–	–	–
40–59 years	5223	11	7	4
60–74 years	4097	19	7	12
≥ 75 years	1379	25	9	16

Abbreviation: CA chronological age



men and 80 parameters for women that had a higher coefficient of correlation in step 1 compared with any counterparts. As there were no parameters not included in any pairs with a ‘strong correlation’, a total of 61 parameters for men (61 + 0) and 80 parameters for women (80 + 0) were selected for biological age modeling (Table S2; see Additional file 2).

**Construction of the biological age models**

**PCA**

The PCA model was constructed using the 26 ECG parameters. The model consisted of eight unrotated principal components with corresponding eigenvalues  $\geq 1.0$ . The factor loadings of the 26 ECG parameters of the PCA model are presented in Table S3 (see Additional file 3). BA by PCA ranged from - 5.31 to 132.97 (Table 3).

**KDM**

BA<sub>E</sub> and BA<sub>EC</sub> were calculated using the 26 ECG parameters. The values of  $r_{char}$  and  $S^2_{BA}$  were 8.83 and 765.39, respectively. BA<sub>E</sub> ranged from - 66.21 to 202.17, whereas BA<sub>EC</sub> ranged from - 44.94 to 178.58 (Table 3).

**Predictive capability of biological age for mortality**

The predictive capabilities by AUC for all-cause death and cardiovascular death are shown in Table 4 and the *p*-values by the paired and unpaired Delong’s test for comparing AUCs are shown in Table S4; see Additional file 4. The AUCs for all-cause death for CA, BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.725, 0.731, 0.657, and 0.685 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs, except for 0.038 for BA vs BA<sub>E</sub> and 0.002 for BA<sub>E</sub> vs BA<sub>EC</sub>), respectively, while those for cardiovascular death were 0.674, 0.682, 0.685, and 0.692 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs), respectively.

The AUCs for all-cause death and cardiovascular death according to CA categories are also shown in Table 4. Because there were no deaths in patients with a CA of 20–39 years, receiver operating curve analysis was not performed in this CA category. In patients with a

CA of 40–59 years, the AUCs for all-cause death for CA, BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.573, 0.592, 0.583, and 0.586 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs), respectively, while those for cardiovascular death were 0.605, 0.521, 0.473, and 0.481, respectively (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs). In patients with a CA of 60–74 years, the AUCs for all-cause death for CA, BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.482, 0.619, 0.702, and 0.697 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs, except for 0.006 for CA vs BA<sub>E</sub> and 0.005 for CA vs BA<sub>EC</sub>), respectively, while those for cardiovascular death were 0.549, 0.673, 0.811, and 0.803 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs), respectively. In patients with a CA  $\geq 75$  years, the AUCs for all-cause death for CA, BA, BA<sub>E</sub>, and BA<sub>EC</sub> were 0.680, 0.575, 0.534, and 0.539 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs, except for 0.010 for CA vs BA<sub>EC</sub>), respectively, while those for cardiovascular death were 0.599, 0.615, 0.686, and 0.689 (*p*-values by the paired Delong’s test were  $> 0.05$  for all pairs), respectively. When the difference of the predictive capability among age categories was compared in a same biological age algorithm, the AUC for all-cause death by BA<sub>EC</sub> was higher in patients with a CA of 60–74 years than in patients with a CA  $\geq 75$  years (unpaired Delong’s test; *p* = 0.007). The AUCs for cardiovascular death by BA<sub>E</sub> and BA<sub>EC</sub> were higher in patients with a CA of 60–74 years than in patients with a CA of 40–59 years (unpaired Delong’s test; *p* = 0.011 and 0.015, respectively).

**Discussion**

A number of studies have examined the utility of medical records, vital signs, laboratory data [18], and epigenetic changes [19] for prediction of biological age. Differences between biological age and CA are thought to reflect acceleration of epigenetic age because of the associations with a higher risk of all-cause mortality [20, 21], CVD [19, 22], and cross-sectionally with obesity [23], earlier menopause [24], and frailty [25].

ECG can be performed easily and repeatedly, and can be analyzed instantly. ECG may be a candidate tool for estimating biological age because ECG parameters can

**Table 3** CA and biological age estimated by three algorithms

	CA				BA				BA <sub>E</sub>				BA <sub>EC</sub>			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Total	55.50	15.04	20.00	90.00	53.76	21.15	-5.31	132.97	44.39	26.40	-66.21	202.17	48.03	21.84	-44.94	178.58
CA of 20–39 years	32.38	5.14	20.00	39.00	24.38	9.36	-5.31	76.49	26.39	19.92	-64.43	145.85	28.24	15.75	-44.94	124.70
CA of 40–59 years	50.00	5.63	40.00	59.00	46.31	10.78	11.64	96.54	41.09	23.84	-66.21	165.13	44.08	18.40	-43.60	143.33
CA of 60–74 years	66.46	4.16	60.00	74.00	68.00	10.32	37.62	118.56	52.78	25.72	-40.22	202.09	57.23	19.34	-20.19	174.76
CA $\geq 75$ years	79.55	3.76	75.00	90.00	85.20	11.55	47.63	132.97	59.91	27.60	-3.65	202.17	66.31	20.25	22.44	178.58

Abbreviations: CA chronological age, BA biological age by principal component analysis method, BA<sub>E</sub> biological age by the Klemera and Doubal’s method without adjustment for chronological age, BA<sub>EC</sub> biological age by the Klemera and Doubal’s method with adjustment for chronological age, SD standard deviation

**Table 4** The predictive capability for all-cause death and cardiovascular death using CA and biological age by three algorithms

	CA			BA			BA <sub>E</sub>			BA <sub>EC</sub>		
	AUC	95% CI	P-value	AUC	95% CI	P-value	AUC	95% CI	P-value	AUC	95% CI	P-value
All-cause death												
Total	0.725	0.654–0.796	< 0.001	0.731	0.662–0.799	< 0.001	0.657	0.584–0.730	< 0.001	0.685	0.616–0.754	< 0.001
CA of 40–59 years	0.573	0.455–0.691	0.404	0.592	0.437–0.746	0.293	0.583	0.410–0.756	0.342	0.586	0.416–0.756	0.325
CA of 60–74 years	0.482	0.349–0.615	0.783	0.619	0.489–0.750	0.071	0.702	0.579–0.826	0.002	0.697	0.574–0.820	0.002
CA ≥75 years	0.680	0.563–0.797	0.001	0.575	0.477–0.673	0.197	0.534	0.414–0.655	0.556	0.539	0.421–0.658	0.498
CV death												
Total	0.674	0.556–0.791	0.003	0.682	0.556–0.807	0.002	0.685	0.568–0.802	0.002	0.692	0.573–0.811	0.001
CA of 40–59 years	0.605	0.451–0.758	0.338	0.521	0.319–0.724	0.845	0.473	0.258–0.689	0.808	0.481	0.270–0.692	0.862
CA of 60–74 years	0.549	0.282–0.817	0.653	0.673	0.452–0.894	0.113	0.811	0.704–0.918	0.004	0.803	0.686–0.921	0.005
CA ≥75 years	0.599	0.414–0.783	0.306	0.615	0.461–0.769	0.232	0.686	0.517–0.856	0.053	0.689	0.524–0.854	0.050

Abbreviations: CA chronological age, BA biological age by principal component analysis method, BA<sub>E</sub> biological age by the Klemra and Doubal's method without adjustment for chronological age, BA<sub>EC</sub> biological age by the Klemra and Doubal's method with adjustment for chronological age, AUC area under the curve, CI confidence interval

be affected by age [5, 6]. As ECG reflects the cardiac condition, which is closely associated with the circulatory and respiratory systems, biological age estimated by ECG is suggested to reflect 'heart age' [10, 12]. Accordingly, when biological age estimated by ECG is utilized for predicting mortality, the prediction would primarily involve cardiovascular death. In the present study, while the AUCs for all-cause death and cardiovascular death with biological age by ECG were comparable to that with CA in total patients, the AUCs for all-cause death by biological age in patients with a CA of 60–74 years was partly higher than that with CA, and the AUCs for cardiovascular death by biological age in patients with a CA of 60–74 years showed the trend to be high. These findings suggest that biological age by ECG may provide, at least in part, a prediction of mortality related to non-cardiovascular causes. Of interest, several studies have reported that ECG can be affected by various extracardiac diseases. For example, ventricular repolarization was altered by hemodialysis [26], prolonged QTc was observed in end-stage liver disease [27, 28] and non-alcoholic fatty liver disease [29], ST segment and T waves can be altered in acute cholecystitis [30], ST depression, left ventricular hypertrophy, prolonged QTc, and T wave inversion were observed in patients with intracranial hemorrhage [31, 32], other ECG abnormalities were reported in patients with brain injury and stroke [31], and higher heart rate, prolonged QTc, and low voltage was observed in patients with thyroid dysfunction [33].

The predictive capability of biological age for various prognoses is generally considered to decline in older people because of the increased biological heterogeneity [34]. Nevertheless, in the present study the predictive capability of biological age for all-cause death and cardiovascular death was higher than that for CA in

patients with a CA of 60–74 years, but was mostly comparable to CA in patients with a CA ≥75 years. However, the range of biological age in patients with a CA ≥75 years was narrower than that for patients with a CA of 60–74 years or 20–59 years, suggesting 'decreased' biological heterogeneity in older patients.

In the present study, we used three types of biological age assessment (BA by PCA, and BA<sub>E</sub> and BA<sub>EC</sub> by KDM). The distribution of BA (by PCA) was generally good, although the minimum value was lower than zero in patients with a CA of 20–39 years. By contrast, BA<sub>E</sub> and BA<sub>EC</sub> (by KDM) showed an extremely wide distribution, ranging from –66.21 to 202.17 and –44.94 to 178.58, respectively. Furthermore, for the distribution of BA<sub>E</sub> and BA<sub>EC</sub> in each CA category, only a BA<sub>EC</sub> of ≥75 years showed a minimum value over zero. Of note, BA<sub>E</sub> and BA<sub>EC</sub> showed a high predictive capability for all-cause and cardiovascular death in patients with a CA of 60–74 years. Thus, despite their wide distribution, BA<sub>E</sub> and BA<sub>EC</sub> may be useful for predicting all-cause and cardiovascular death in patients with a CA 60–74 years.

The cost-effectiveness and non-invasive nature are major advantages of using ECG to assess biological age. We found that biological age by ECG is particularly useful for discriminating high or low risk for mortality in patients aged 60–74 years old and in discriminating the risk for cardiovascular death in patients aged ≥75 years old. Our data confirm that biological age by ECG provides an indicator of 'heart age'. Furthermore, we provide new evidence that the predictive capability of biological age by ECG varies according to age categories.

#### Limitations

There were several limitations of this study. First, all participants were patients who visited a cardiovascular hospital in an urban area. Although we analyzed the

patients without structural heart diseases, they have some reasons to visit a cardiology specialized hospital, including, at least, mild symptoms or minor ECG abnormalities. Therefore, our data should be carefully interpreted and are not easily extrapolated to general populations. Second, we used the ECG parameters provided by a commercial ECG machine (GE Healthcare). Given that the approaches or algorithms used to measure the ECG waves may differ between machines from different manufacturers, revalidation with other ECG machines may be necessary. Finally, patients' characteristics such as cardiac anatomical information, comorbidities, concomitant medications, and frailty were not included in our models.

## Conclusion

We developed a prediction model for biological age using 12-lead ECG parameters in patients without structural heart diseases. This model showed a similar predictive capability to CA for all-cause death and cardiovascular death among total patients, but partially showed a significant increase in the predictive capability among patients aged 60–74 years old.

## Abbreviations

AUC: Area under the curve; CA: Chronological age; CVDs: Cardiovascular diseases; ECG: Electrocardiogram; eGFR: Estimated glomerular filtration rate; KDM: Klemmer and Doubal's method; PCA: Principal component analysis

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12877-021-02391-8>.

- Additional file 1: Table S1.** Patient characteristics.
- Additional file 2: Table S2.** Selected ECG parameters for biological age models.
- Additional file 3: Table S3.** Factor weights of the 26 ECG parameters by the PCA model.
- Additional file 4: Table S4.** Comparison of AUCs.

## Acknowledgments

We thank Mikio Kishi, Hiroaki Semba, Hiroto Kano, Shunsuke Matsuno, Yuko Kato, Tokuhisa Uejima, Yuji Oikawa, Minoru Matsuhama, Tatsuya Inoue, and Junji Yajima for cooperating for the data acquisition. We also thank Shiro Ueda and Nobuko Ueda at Medical Edge Company, Ltd., for assembling the database using the Clinical Study Supporting System, and Yurika Hashiguchi, Hiroaki Arai, Takashi Osada, and Hiroshi Nakai for data management and system administration. We also thank Edanz Group (<https://en-author-services.edanzgroup.com/ac>) for editing a draft of this manuscript.

## Authors' contributions

NH, SS, and TY conceived the study concept and design. NH and SS analyzed the data. NH and SS drafted the manuscript, and TA, NY, TO, and TY revised the manuscript. NH, SS, TA, NY, TO, and TY approved the final version.

## Funding

This study was partially supported by the Practical Research Project for Life-Style related Diseases including Cardiovascular Diseases and Diabetes Mellitus from the Japan Agency for Medical Research and Development (AMED; JP17ek0210082). However, they were not involved with the design of the

study and collection, analysis, and interpretation of data and writing the manuscript.

Dr. Suzuki received research funding from Mitsubishi Tanabe Pharm. Daiichi Sankyo and Dr. Yamashita has received research funds and/or lecture fees from Daiichi Sankyo, Bayer Yakuhin, Bristol-Myers Squibb, Pfizer, Nippon Boehringer Ingelheim, Eisai, Mitsubishi Tanabe Pharm, Ono Pharmaceutical, and Toa Eiyo.

## Availability of data and materials

The data cannot be shared publicly because there was no such approval in the study protocol or informed consent. Data are available from the Ethics Review Committee at the Cardiovascular Institute for researchers who meet the criteria for access to confidential data (contact via the corresponding author).

## Declarations

### Ethics approval and consent to participate

This study was performed in accordance with the ethical norms based on the Declaration of Helsinki (revised in 2013) and the Ethical Guidelines for Medical and Health Research Involving Human Subjects (Public Notice of the Ministry of Education, Culture, Sports, Science and Technology, and the Ministry of Health, Labour and Welfare, Japan, issued in 2017). Written informed consent was obtained from all participants. The study protocol was reviewed by the Institutional Review Board of the Cardiovascular Institute.

### Consent for publication

Not applicable.

### Competing interests

None.

Received: 7 May 2020 Accepted: 13 July 2021

Published online: 11 August 2021

## References

- Arbeev KG, Ukraintseva SV, Yashin AI. Dynamics of biomarkers in relation to aging and mortality. *Mech Ageing Dev.* 2016;156:42–54. <https://doi.org/10.1016/j.mad.2016.04.010>.
- Levine ME. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol A Biol Sci Med Sci.* 2013;68(6):667–74. <https://doi.org/10.1093/geron/gls233>.
- Park J, Cho B, Kwon H, Lee C. Developing a biological age assessment equation using principal component analysis and clinical biomarkers of aging in Korean men. *Arch Gerontol Geriatr.* 2009;49(1):7–12. <https://doi.org/10.1016/j.archger.2008.04.003>.
- Yoo J, Kim Y, Cho ER, Jee SH. Biological age as a useful index to predict seventeen-year survival and mortality in Koreans. *BMC Geriatr.* 2017;17(1):7. <https://doi.org/10.1186/s12877-016-0407-y>.
- Rijnbeek PR, van Herpen G, Bots ML, Man S, Verweij N, Hofman A, et al. Normal values of the electrocardiogram for ages 16–90 years. *J Electrocardiol.* 2014;47(6):914–21. <https://doi.org/10.1016/j.jelectrocard.2014.07.022>.
- Mason JW, Ramseth DJ, Chanter DO, Moon TE, Goodman DB, Mendzelevski B. Electrocardiographic reference ranges derived from 79,743 ambulatory subjects. *J Electrocardiol.* 2007;40(3):228–34. <https://doi.org/10.1016/j.jelectrocard.2006.09.003>.
- Vicent L, Martinez-Selles M. Electrocardiogeriatrics: ECG in advanced age. *J Electrocardiol.* 2017;50(5):698–700. <https://doi.org/10.1016/j.jelectrocard.2017.06.003>.
- Kistler PM, Sanders P, Fynn SP, Stevenson IH, Spence SJ, Vohra JK, et al. Electrophysiologic and electroanatomic changes in the human atrium associated with age. *J Am Coll Cardiol.* 2004;44(1):109–16. <https://doi.org/10.1016/j.jacc.2004.03.044>.
- Attia ZI, Friedman PA, Noseworthy PA, Lopez-Jimenez F, Ladewig DJ, Satam G, et al. Age and sex estimation using artificial intelligence from standard 12-Lead ECGs. *Circ Arrhythm Electrophysiol.* 2019;12(9):e007284. <https://doi.org/10.1161/CIRCEP.119.007284>.
- Ball RL, Feiveson AH, Schlegel TT, Starc V, Dabney AR. Predicting "heart age" using electrocardiography. *J Pers Med.* 2014;4(1):65–78. <https://doi.org/10.3390/jpm4010065>.

11. Hirsch JR, Waits G, Li Y, Soliman EZ. Racial differences in heart age and impact on mortality. *J Natl Med Assoc.* 2018;110(2):169–75. <https://doi.org/10.1016/j.jnma.2017.08.003>.
12. Khane RS, Surdi AD, Bhatkar RS. Changes in ECG pattern with advancing age. *J Basic Clin Physiol Pharmacol.* 2011;22(4):97–101. <https://doi.org/10.1515/JBCPP.2011.017>.
13. Nakamura E, Miyao K, Ozeki T. Assessment of biological age by principal component analysis. *Mech Ageing Dev.* 1988;46(1–3):1–18. [https://doi.org/10.1016/0047-6374\(88\)90109-1](https://doi.org/10.1016/0047-6374(88)90109-1).
14. Klemra P, Doubal S. A new approach to the concept and computation of biological age. *Mech Ageing Dev.* 2006;127(3):240–8. <https://doi.org/10.1016/j.mad.2005.10.004>.
15. Suzuki S, Otsuka T, Sagara K, Semba H, Kano H, Matsuno S, et al. Nine-year trend of anticoagulation use, thromboembolic events, and major bleeding in patients with non-Valvular atrial fibrillation- Shinken database analysis. *Circ J.* 2016;80(3):639–49. <https://doi.org/10.1253/circ.CJ-15-1237>.
16. Cho IH, Park KS, Lim CJ. An empirical comparative study on biological age estimation algorithms with an application of work ability index (WAI). *Mech Ageing Dev.* 2010;131(2):69–78. <https://doi.org/10.1016/j.mad.2009.12.001>.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837–45. <https://doi.org/10.2307/2531595>.
18. Wang Z, Li L, Glicksberg BS, Israel A, Dudley JT, Ma'ayan A. Predicting age by mining electronic medical records with deep learning characterizes differences between chronological and physiological age. *J Biomed Inform.* 2017;76:59–68. <https://doi.org/10.1016/j.jbi.2017.11.003>.
19. Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. Prospective study of epigenetic age acceleration and incidence of cardiovascular disease outcomes in the ARIC study (atherosclerosis risk in communities). *Circ Genom Precis Med.* 2018;11(3):e001937. <https://doi.org/10.1161/CIRCGEN.117.001937>.
20. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY).* 2016;8(9):1844–65. <https://doi.org/10.18632/aging.101020>.
21. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* 2015;16(1):25. <https://doi.org/10.1186/s13059-015-0584-6>.
22. Perna L, Zhang Y, Mons U, Holleczeck B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics.* 2016;8(1):64. <https://doi.org/10.1186/s13148-016-0228-z>.
23. Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schonfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci U S A.* 2014;111(43):15538–43. <https://doi.org/10.1073/pnas.1412759111>.
24. Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, et al. Menopause accelerates biological aging. *Proc Natl Acad Sci U S A.* 2016; 113(33):9327–32. <https://doi.org/10.1073/pnas.1604558113>.
25. Breitling LP, Saum KU, Perna L, Schottker B, Holleczeck B, Brenner H. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. *Clin Epigenetics.* 2016;8(1):21. <https://doi.org/10.1186/s13148-016-0186-5>.
26. Sivri S, Celik M. Evaluation of index of cardiac-electrophysiological balance before and after hemodialysis in patients with end-stage renal disease. *J Electrocardiol.* 2019;54:72–5. <https://doi.org/10.1016/j.jelectrocard.2019.03.011>.
27. Glowczynska R, Galas M, Oldakowska-Jedynak U, Peller M, Tomaniak M, Raszeja-Wyszomirska J, et al. Pretransplant QT interval: the relationship with severity and etiology of liver disease and prognostic value after liver transplantation. *Ann Transplant.* 2018;23:622–30. <https://doi.org/10.12659/AOT.908769>.
28. Moller S, Hove JD, Dixen U, Bendtsen F. New insights into cirrhotic cardiomyopathy. *Int J Cardiol.* 2013;167(4):1101–8. <https://doi.org/10.1016/j.ijcard.2012.09.089>.
29. Ismaiel A, Colosi HA, Rusu F, Dumitrascu DL. Cardiac arrhythmias and electrocardiogram modifications in non-alcoholic fatty liver disease. A systematic review. *J Gastrointest Liver Dis.* 2019;28(4):483–93. <https://doi.org/10.15403/jgld-344>.
30. Ozeki M, Takeda Y, Morita H, Miyamura M, Sohmiya K, Hoshiga M, et al. Acute cholecystitis mimicking or accompanying cardiovascular disease among Japanese patients hospitalized in a cardiology department. *BMC Res Notes.* 2015;8(1):805. <https://doi.org/10.1186/s13104-015-1790-8>.
31. Katsanos AH, Korantzopoulos P, Tsvigoulis G, Kyritsis AP, Kosmidou M, Giannopoulos S. Electrocardiographic abnormalities and cardiac arrhythmias in structural brain lesions. *Int J Cardiol.* 2013;167(2):328–34. <https://doi.org/10.1016/j.ijcard.2012.06.107>.
32. Takeuchi S, Nagatani K, Otani N, Wada K, Mori K. Electrocardiograph abnormalities in intracerebral hemorrhage. *J Clin Neurosci.* 2015;22(12):1959–62. <https://doi.org/10.1016/j.jocn.2015.04.028>.
33. Tayal B, Graff C, Selmer C, Kragholm KH, Kihlstrom M, Nielsen JB, et al. Thyroid dysfunction and electrocardiographic changes in subjects without arrhythmias: a cross-sectional study of primary healthcare subjects from Copenhagen. *BMJ Open.* 2019;9(6):e023854. <https://doi.org/10.1136/bmjopen-2018-023854>.
34. Jylhava J, Pedersen NL, Hagg S. Biological age predictors. *EBioMedicine.* 2017;21:29–36. <https://doi.org/10.1016/j.ebiom.2017.03.046>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

