

Heart age estimated using explainable advanced electrocardiography

Thomas Lindow, MD, PhD^{1,2,3*}, Israel Palencia-Lamela,^{1,4*} Todd T Schlegel, MD^{5,6},
Martin Ugander, MD, PhD^{1,5}

1. Kolling Institute, Royal North Shore Hospital, and University of Sydney, Sydney, Australia
2. Department of Clinical Physiology, Research and Development, Växjö Central Hospital, Region Kronoberg, Sweden
3. Clinical Physiology, Clinical Sciences, Lund University, Sweden
4. Davidson College, Davidson, North Carolina, USA
5. Department of Clinical Physiology, Karolinska Institutet, and Karolinska University Hospital, Stockholm, Sweden
6. Nicollier-Schlegel SARL, Trélex, Switzerland

**Shared first-authorship by equal contribution*

Fundings: TL is currently under the support of postdoctoral research grants from The Swedish Heart-Lung Foundation (grant no 20200553), the Swedish Cardiac Society, the Royal Swedish Academy of Sciences (grant no LM2019-0013), Women and Health Foundation, Region Kronoberg (grant no 8301), The Swedish Heart and Lung Association (grant no LKH1387), Swedish Association of Clinical Physiology, and the Scandinavian Society of Clinical Physiology&Nuclear Medicine. The study was funded in part by grants (PI Ugander) from New South Wales Health, Heart Research Australia, and the University of Sydney.

Author contributions: TL and IPL contributed to study design, performed the data analysis, wrote the first manuscript drafts, prepared the figures, and are shared first authors by equal contributions. TTS curated the data, contributed to study design, performed the prediction modeling, and critically revised the manuscript. MU designed the study, provided supervision, and critically revised the manuscript. All authors read and approved the final version of the manuscript.

Competing interests:

TL, IPL: None

TTS is owner and founder of Nicollier-Schlegel SARL, which performs ECG interpretation consultancy using software that can quantify the advanced ECG measures used in the current study. TTS and MU are owners and founders of Advanced ECG Systems, a company that is developing commercial applications of advanced ECG technology used in the current study.

Corresponding author: Martin Ugander, martin.ugander@ki.se, Professor of Cardiac Imaging, University of Sydney, Sydney, Australia, and Karolinska Institute, Sweden.

39 **Background:** Electrocardiographic (ECG) Heart Age conveying cardiovascular risk has been
40 estimated by both Bayesian and artificial intelligence approaches. We hypothesised that
41 explainable measures from the 10-second 12-lead ECG could successfully predict Bayesian 5-
42 min ECG Heart Age.

43 **Methods:** Advanced analysis was performed on ECGs from healthy subjects and patients
44 with cardiovascular risk or proven heart disease. Regression models were used to predict
45 patients' Bayesian 5-minute ECG Heart Ages from their standard, resting 10-second 12-lead
46 ECGs. The difference between 5-min and 10-second ECG Heart Ages were analyzed, as were
47 the differences between 10-second ECG Heart Age and the chronological age (the Heart Age
48 Gap).

49 **Results:** In total, 2,771 subjects were included (n=1682 healthy volunteers, n=305 with
50 cardiovascular risk factors, n=784 with cardiovascular disease). Overall, 10-second Heart Age
51 showed strong agreement with the 5-minute Heart Age ($R^2=0.94$, $p<0.001$, mean \pm SD bias
52 0.0 ± 5.1 years). The Heart Age Gap was 0.0 ± 5.7 years in healthy individuals, 7.4 ± 7.3 years in
53 subjects with cardiovascular risk factors ($p<0.001$), and 14.3 ± 9.2 years in patients with
54 cardiovascular disease ($p<0.001$).

55 **Conclusions:** Heart Age can be accurately estimated from a 10-second 12-lead ECG in a
56 transparent and explainable fashion based on known ECG measures, without artificial
57 intelligence techniques. The Heart Age Gap increases markedly with cardiovascular risk and
58 disease.

59 **Keywords:** vascular age; electrocardiography cardiovascular risk prediction; advanced
60 electrocardiography; vectorcardiography

61

62

63 **Background**

64 Cardiovascular disease is a significant contributor to mortality, with related pathological
65 processes often beginning early and progressing silently for many years¹⁻³. Many risk factors
66 for developing cardiovascular disease are lifestyle-related⁴. Fortunately, several are
67 modifiable, and risk can therefore be reduced, for example by smoking cessation, dietary
68 changes and increased physical activity^{5,6}. To accomplish this, an individual must fully
69 understand their risk and become motivated to reduce it. One way of communicating risk to
70 the patient is to present it as a “Heart Age”, which can be contrasted to the patient’s
71 chronological age, i.e. a “Heart Age Gap”. A Heart Age can either be determined by
72 translating risk factor scores to what age a resultant score would represent in an individual
73 with no risk factors, or it can be based on electrocardiographic changes⁷⁻¹². Describing the risk
74 to the patient using Heart Age has been reported to reduce metabolic risk factors and may
75 have the advantage of being easily understood by the patient^{11,13}. A similar approach has been
76 applied when conveying risk to smokers by describing how ‘old’ their lungs are, and such an
77 approach increased the chance of smoking cessation¹⁴.

78 Moving beyond but not excluding basic ECG measurements such as heart rate and waveform
79 amplitudes and durations, the diagnostic output from the ECG can be further substantially
80 improved by also utilizing combinations of advanced ECG measures, for example those from
81 12-lead-ECG-derived vectorcardiography, waveform complexity, and measures of beat-to-
82 beat waveform variability¹⁵⁻¹⁷. In 2014, such advanced measures were used along with 5-
83 minute, high-fidelity 12-lead ECGs and a Bayesian statistical approach, to develop an
84 accurate estimate of ECG-based Heart Age. The estimation of Heart Age using this method
85 has been described in detail elsewhere⁹. In brief, the Bayesian approach was used to infer the
86 deviation of the Heart Age from the subject’s known chronological age by assessing and
87 quantifying an individual’s expected versus actual ECG findings. In addition to sex and

88 chronological age itself, the ECG parameters used in the method as modifiers of the Bayesian-
89 predicted Heart Age were the T-wave axis in the frontal plane, P-wave duration, frontal plane
90 vectorcardiographic QRS axis, spatial JT interval, spatial mean QRS-T angle, high-frequency
91 QRS root mean squared voltages across signal-averaged leads after band-pass filtering, beat-
92 to-beat QT- and RR-interval variability, and measures of T-wave complexity based on
93 singular value decomposition and signal averaging of the T wave. After being trained on a set
94 of healthy individuals, this approach was clinically validated by describing an increased
95 difference between Heart Age and chronological age both for subjects at risk of
96 cardiovascular disease, and those with established cardiovascular disease⁹.

97 However, importantly, reliable measurement of beat-to-beat RR and QT interval variability
98 requires ECG recordings of at least several minutes. In addition, accurate quantification of
99 high frequency QRS root mean squared voltages requires higher-fidelity recordings with more
100 specialised equipment and sampling rates exceeding the 250-500 samples per second used
101 today in most ECG machines used for standard, 10-second ECG recordings. Therefore, the
102 main aim of this study was to ascertain whether 5-minute ECG Heart Age could be accurately
103 predicted by only those measures available from a standard 10-second 12-lead ECGs.

104 Furthermore, a subsequent secondary aim was to also compare 10-second ECG Heart Ages to
105 chronological ages in healthy subjects, subjects with cardiovascular risk factors, and patients
106 with established cardiovascular disease. We hypothesised that standard-fidelity, 10-second
107 12-lead ECG recordings could accurately predict Bayesian ECG Heart Ages derived from
108 higher-fidelity, 5-minute 12-lead ECG recordings.

109

110 **Methods**

111 A pre-existing database of de-identified subjects with both 5-minute and 10-second 12-lead
112 ECG recordings was utilised for the study^{9,16}. Within that database, healthy individuals,

113 patients at cardiovascular risk, and patients with established cardiovascular disease were
114 included. The healthy subjects, who had volunteered as asymptomatic controls, were recruited
115 at Johnson Space Center (USA), the Universidad de los Andes (Venezuela), the University of
116 Ljubljana hospitals and clinics (Slovenia), or Lund University Hospital (Sweden), as
117 previously described in detail¹⁶. Patients with cardiovascular risk factors or established
118 cardiovascular disease were recruited from cardiology clinics at either Texas Heart Institute
119 (Houston, USA); the University of Texas Medical Branch (Galveston, USA), the University
120 of Texas Health Sciences Center (San Antonio, USA), Brooke Army Medical Center (San
121 Antonio, USA); St. Francis Hospital (Charleston, USA), the Universidad de los Andes
122 (Mérida, Venezuela); and Lund University Hospital (Lund, Sweden).

123

124 All healthy subjects were low risk, asymptomatic volunteers with absence of any
125 cardiovascular or systemic disease, based on clinical history and physical examination.
126 Exclusion criteria for the healthy subjects included increased blood pressure at physical
127 examination ($\geq 140/90$ mm Hg), treatment for hypertension or diabetes, or active smoking.
128 Patients in the established cardiovascular disease group were included based on the presence
129 of either coronary heart disease (determined by coronary angiography with at least one
130 obstructed vessel ($\geq 50\%$) in at least one major native coronary vessel or coronary graft, or, if
131 coronary angiography was either unavailable or clinically not indicated, one or more
132 reversible perfusion defects on ^{99m}Tc-tetrofosmin single-photon emission computed
133 tomography (SPECT)¹⁸⁻²⁰), left ventricular hypertrophy (LVH) based on imaging evidence of
134 at least moderate, concentric wall thickening according to guidelines of the American Society
135 of Echocardiography²¹, left ventricular systolic dysfunction (left ventricular ejection fraction
136 $\leq 50\%$) at echocardiography, cardiac magnetic resonance imaging (CMR) or SPECT, or
137 findings suggestive of dilated/hypertrophic/ischemic cardiomyopathy at echocardiography or

138 CMR¹⁶. ECGs were acquired within 30 days of the cardiac imaging examination. Finally,
139 subjects in the cardiovascular risk group were included based on the presence of
140 cardiovascular risk factors such as hypertension or diabetes but no confirmed, established
141 cardiovascular disease⁹.

142 Based on the above, three groups of study participants were identified: healthy subjects,
143 subjects at cardiovascular risk, and patients with established cardiovascular disease. By
144 methodological design, only healthy subjects were initially included when considering
145 optimal measures of 12-lead ECG available from 10-second recordings for predicting the 5-
146 minute ECG Heart Age. The 10-second ECG measures considered for the prediction model
147 included: (1) From the conventional ECG: heart rate, R-to-R, P-wave, PR, QRS, QT, QTc,
148 and TQ interval durations, as well as the conventional ECG amplitudes and axes; (2) From the
149 transformation of the 12-lead ECG to the Frank X, Y and Z lead vectorcardiogram (VCG) via
150 Kors' transform²²⁻²⁶: the spatial means and peaks QRS-T angles, the spatial ventricular
151 gradient and its individual QRS and T components, the spatial QRS- and T-wave
152 axes (azimuths and elevations), waveform amplitudes and areas, including those in the three
153 individual vectorcardiographic planes, and spatial QRS- and T-wave velocities; and (3)
154 Multiple measures of QRS- and T-wave waveform complexity based on singular value
155 decomposition after signal averaging²⁷⁻²⁹.

156 For all study participants, the Heart Age Gap was calculated as the 10-second ECG Heart Age
157 minus chronological age. All results were compared between the three groups, and for
158 subgroups based on sex and age (above or below 60 years).

159 Each participant gave written informed consent. All recordings were obtained under
160 Institutional Review Board (IRB) approvals from NASA's Johnson Space Center and partner

161 hospitals that fall under IRB exemptions for previously collected and de-identified data. The
162 study was performed in accordance with the Declaration of Helsinki.

163 *Statistical analysis*

164 Continuous variables were described using mean and standard deviation (SD). The chi-
165 squared test was used to test for proportional differences between groups. Student's *t* test was
166 used to compare group means.

167 The 10-second ECG Heart Age was initially derived only in the healthy subject group. First,
168 standard least squares linear regression was used to identify the most promising univariable
169 measures available from 10-second, standard-fidelity ECG for predicting the 5-minute ECG
170 Heart Age. More than 20 such measures were identified, along with chronological age and
171 sex, that had individual $P < 0.0001$. After the univariable linear regression identification
172 procedure, stepwise, standard least squares multiple linear regression was then performed
173 only on the selected measures to best predict, via multivariable model, the 5-minute ECG
174 Heart Age in the healthy subject group. The optimal multivariable model was defined as the
175 most parsimonious one that maintained the highest possible R-squared value while at the
176 same time maintaining all $p < 0.0001$ for the individual measures within the model in the
177 healthy subject group, and at < 0.01 when the model was subsequently applied forward to the
178 non-healthy (cardiovascular risk and disease) groups. Comparisons with 5-minute ECG Heart
179 Age are presented as scatter plots and Bland-Altman plots. A two-sided *p*-value of 0.05 was
180 used as to define statistical significance. Statistical analyses were performed using SAS JMP
181 version 11.0, SAS Institute Inc, Cary, NC, USA, and R version 3.5.3R, R Foundation for
182 Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.

183

184

185 **Results**

186 In total, 2,771 patients were included (n=1682 healthy volunteers, n=305 subjects with
187 cardiovascular risk factors, and n=784 with cardiovascular disease). Baseline characteristics
188 are presented in Table 1, and average values for the ultimately included ECG measures are
189 presented in Table 2. The ECG measures included in the final prediction models and the
190 intercept and coefficients for the regression equations are also presented in Table 3 for males
191 and Table 4 for females.

192 Agreement between the 10-second Heart Age and the 5-minute Heart Age was excellent
193 ($R^2=0.94$, $p<0.001$, mean \pm SD bias 0.0 \pm 5.2 years), Figure 1. Agreement was strong for both
194 males and females ($R^2=0.91$, $p<0.001$, and $R^2=0.92$, $p<0.001$ respectively). In healthy
195 subjects, on average, the Heart Age and chronological age were no different (0.0 \pm 5.7 years),
196 i.e. the Heart Age Gap was zero. In subjects with cardiovascular risk factors, the Heart Age
197 Gap was larger (7.4 \pm 7.3 years, $p<0.001$). Patients with cardiovascular disease showed the
198 largest Heart Age Gap (14.3 \pm 9.2 years, $p<0.001$ vs. subjects at cardiovascular risk), Figure 2.
199 This was evident both for females (Heart Age Gap: healthy subjects: 0.4 \pm 5.8 years;
200 cardiovascular risk: 7.8 \pm 8.0 years; cardiovascular disease: 14.8 \pm 9.0 years, $p<0.001$) and for
201 males (Heart Age Gap: healthy subjects: -0.3 \pm 5.6 years; cardiovascular risk: 7.1 \pm 6.6 years;
202 cardiovascular disease: 14.0 \pm 9.3 years, $p<0.001$). A similar pattern was observed for younger
203 (age<60 years) individuals (n=2,142; Heart Age Gap: healthy subjects: 0.1 \pm 5.7 years;
204 cardiovascular risk: 7.8 \pm 6.4 years; cardiovascular disease: 15.4 \pm 9.8 years, $p<0.001$) and for
205 older individuals (n=629; Heart Age Gap: healthy subjects: -1.7 \pm 6.8 years; cardiovascular
206 risk: 7.8 \pm 6.4 years; cardiovascular disease: 15.4 \pm 9.8 years).

207 **Discussion**

208 We show that Heart Age can be estimated from standard-fidelity, conventional 10-second
209 ECG recordings without requiring more specialised, higher-fidelity, 5-minute long recordings.

210 Consequently, ECG-based Heart Age as estimated using Bayesian techniques can now be
211 more readily implemented in clinical practice.

212 ECG-based estimation of Heart Age, as a marker of increased cardiovascular risk, has the
213 advantage of being easily understood by the patient. It might also provide strong incentives
214 for lifestyle changes or compliance to medication, similar to the reporting of “lung age” in the
215 context of smoking cessation¹⁴. For example, in an asymptomatic patient with newly
216 diagnosed hypertension, a value of 160/85 mm Hg may be difficult for the patient to
217 understand in relation to future cardiovascular risk, and the communication of that risk may
218 be difficult for the physician to translate. Instead, if this patient had a chronological age of 50
219 years but his Heart Age was 10 years higher, it may be more intuitive for the patient to
220 understand that actions are needed. Additionally, an elevated blood pressure is merely a risk
221 factor, and we do not know if the individual patient will suffer from the condition associated
222 with the risk or not. By comparison, changes in the pattern of the ECG and advanced ECG
223 often reflect cardiovascular end-organ changes, some potentially reversible, moving one step
224 beyond the reporting of traditional risk factors.

225 In addition, we found Heart Age to be similar to chronological age in healthy individuals,
226 while the Heart Age Gap was increasingly larger with increasing cardiovascular disease
227 status. This suggests that Heart Age might also provide useful information for cardiovascular
228 risk prediction, although validation in other datasets is necessary. Moreover, more tailored
229 ECG approaches directly geared toward predicting outcomes, rather than indirectly geared
230 toward predicting outcomes through a Heart Age per se, would also likely outperform strictly
231 Heart Age-based attempts at risk stratification.

232 In the dataset used, healthy subjects were younger than subjects with cardiovascular risk
233 factors or disease, and there was also a male preponderance in all three groups. The Heart Age

234 Gap showed a similar pattern in both males and females, and in both younger (<60 years) and
235 older subjects.

236 Estimation of both the 5-minute and 10-second Heart Age currently requires sinus rhythm,
237 since P-wave duration is included in the related calculation. Thus, the assignment of a Heart
238 Age to patients in atrial fibrillation is currently precluded. However, the risk associated with
239 atrial fibrillation is well-established, and patients are often symptomatic. Consequently, the
240 use of Heart Age to incrementally incentivise lifestyle interventions or improve compliance to
241 medications may not be as necessary. Nonetheless, future studies to derive similar methods in
242 patients with atrial fibrillation may also be of value.

243 *Explainability and transparency of variables that contribute to ECG Heart Age*

244 It is important to note that the patient's chronological age was purposely included in the
245 estimation of Heart Age, just as in the original (5-min ECG) Bayesian model. The original
246 model was not developed to predict a patient's actual chronological age, arguably a task of
247 negligible clinical importance. Rather, the model was developed to estimate the "age" of the
248 heart, and specifically how much that estimate differed from chronological age – quantified as
249 the Heart Age Gap. A Bayesian statistical approach is predicated on knowing the patient's
250 chronological age, with Heart Age being an adjustment of the chronological age based on
251 electrocardiographic characteristics and sex. For an estimation of Heart Age to be accurate in
252 predicting a Heart Age that is similar to the chronological age when the heart is healthy, and
253 increased when the heart is diseased, it is desirable that the included ECG measures change
254 with age, and that the change is augmented with increasing cardiovascular risk or disease
255 severity.

256 The two ECG measures that had the strongest influence (highest t ratio) on the model were P-
257 wave duration and spatial QT duration. This is not surprising, given that P-wave duration
258 increases with age³⁰, and increased P-wave duration can also be seen in advanced

259 cardiovascular pathologies, e.g. heart failure and cardiac amyloidosis³¹. Similarly, QT
260 duration increases with age³². Further, QT prolongation is associated with increased of
261 cardiovascular risk, even beyond the rare long QT syndromes³³, and with incrementally
262 increased risk in advanced ages³⁴. These general characteristics are also true for increased
263 heart rate and for leftward shifting of the frontal plane QRS axis³⁵⁻³⁸. The other measures
264 included in the score track changes in the vectorcardiographic QRS and T, and in T-wave
265 complexity by singular value decomposition. Such changes are also known to occur in
266 conditions associated with increased cardiovascular risk, such as hypertension and diabetes³⁹,
267 and in established cardiovascular disease, in which they often provide strong diagnostic and
268 prognostic information^{22-25,27-29}. Notably, these changes are not easily detectable by visual
269 interpretation of a standard 12-lead ECG. How these ECG measures can affect the Heart Age
270 is exemplified in Figure 3. Taken together, the described ECG measures that contribute in a
271 multivariable fashion to the Heart Age all have physiologically reasonable associations with
272 age and disease in a way that is transparent to the assessing clinician, thus providing
273 important explainability to the model.

274

275 *Differences compared to Bayesian 5-minute ECG Heart Age*

276 The original, Bayesian 5-minute ECG Heart Age requires information from measures of beat-
277 to-beat heart rate and QT variability,⁹ and of the root-mean square voltage or other aspects of
278 high-frequency (high fidelity) components of the QRS complex⁴⁰⁻⁴². However, 10-second-
279 duration recordings of standard fidelity do not allow for such measures, and therefore they
280 were not included in the 10-second ECG Heart Age. However, unlike the original Bayesian 5-
281 minute ECG Heart Age, the 10-second ECG Heart Age should be derivable from any standard
282 12-lead ECG machine, as long as it is sufficiently equipped with software that can quantify
283 the included measures and calculate the 10-second ECG Heart Age. The presented 10-second

284 ECG Heart Age might therefore be anticipated to contribute to more widespread clinical
285 penetration and use.

286

287 *Comparison with other Heart (or Vascular) Ages*

288 Different means of expressing a heart or vascular age have recently been published¹³,
289 including the use of information in the 10-second 12-lead ECG to estimate ECG Heart Age
290 based on artificial intelligence (AI)^{7,8,10}. For example, Attia, et al, showed that by using a deep
291 neural network (DNN) AI technique, a patient's chronological age could be predicted, and
292 that if the difference between the predicted and actual age was small, prognosis was good⁸.
293 When Heart Age by Attia et al's technique was notably older than the chronological age, the
294 risk of future death was increased⁷. Such results at least superficially correspond to the
295 findings in our study that the Heart Age Gap increased with increasing burden of
296 cardiovascular risk. Another AI method similar to that of Attia et al also more recently
297 reported similarly encouraging results¹⁰.

298

299 Although the results of such AI studies are promising, DNN-based AI techniques are
300 nonetheless inherently problematic in several respects, especially in relation to their lack of
301 transparency and explainability, i.e., the 'black box' of AI^{43,44}. Without the ability to know the
302 exact features of the 12-lead ECG that are most important in a given DNN model's output,
303 both interpretability and ethical accountability are compromised⁴⁵. Moreover, it is effectively
304 impossible for a clinician to identify, when critically evaluating the diagnostic output of a
305 DNN-based AI model, the contribution to the result from methodological artifact or bias
306 merely related to noise or to differing technical specifications between different ECG
307 machines⁴⁶. Or the extent to which unanticipated results might merely relate to excess
308 dependency on the particular characteristics of a given DNN AI model's training set⁴⁷. In

309 addition, a major flaw in all DNN-based AI ECG age models of which we are aware is that
310 their age predictions were made using training datasets that also included individuals with
311 both cardiovascular risk factors and established disease^{8,10}. For ECG Heart Age to be used as
312 a marker of potentially reversible cardiovascular disease and risk, it is imperative that ECG
313 Heart Age agrees with chronological age in healthy populations, since it is the deviations from
314 the line of identity in this relationship that form the basis for disease-related assessment and
315 risk.

316

317 Hence, we believe that the pursuit of an ECG Heart Age developed from heart-healthy
318 subjects of varying ages, but without a black-box DNN or related AI methodology is valuable,
319 and that the present results provide sufficient confirmation of accuracy to encourage further
320 development. Moreover, that the use of more transparent regression models will also
321 increase the ability of clinicians to better understand the origin of any unexpected result, and
322 to thereafter relay it to the patient with a more convincing sense of trust and accountability⁴⁵.
323 Finally, Heart Age can also now be retrospectively determined for clinical purposes or for
324 retrospective scientific studies whenever raw digital data of acceptable quality are available
325 from stored, standard 10-sec ECG recordings.

326

327 *Limitations*

328 The dataset was not strictly divided into a training and test set. However, the fact that the
329 selected ECG measures were derived only in the Healthy group, then applied forward to the
330 other groups wherein the Heart Age Gap was incrementally larger with increasing
331 cardiovascular risk and overt disease, does serve as a form of clinical validation. The
332 prognostic value of the Heart Age Gap was not investigated and requires further evaluation in
333 future studies. The use of three different levels of cardiovascular risk or disease status

334 (healthy, cardiovascular risk without established disease, and established cardiovascular
335 disease) might also be construed as somewhat arbitrary, although they do constitute a logical
336 grouping with incrementally increasing risk from one group to the next. Future studies are
337 nonetheless also required to validate any association between Heart Age and outcomes
338 associated with different cardiovascular pathologies. Finally, although Heart Age was highly
339 accurate, its precision could not be reliably defined in this study. These aspects therefore need
340 to be addressed in future studies.

341

342 *Conclusion*

343 We show that Heart Age can be accurately, transparently, and explainably estimated from a
344 standard 10-second, resting 12-lead ECG utilizing multiple, discrete conventional and
345 advanced ECG measures. The Heart Age Gap increases with increasing cardiovascular risk
346 and disease. However, further prospective evaluation in future studies is also required.

347

348

349 **References**

- 350 1 Bergström, G. *et al.* Prevalence of subclinical coronary artery atherosclerosis in the
351 general population. *Circulation* **144**, 916-929, doi:10.1161/circulationaha.121.055340
352 (2021).
- 353 2 Townsend, N. *et al.* Cardiovascular disease in Europe: epidemiological update 2016.
354 *Eur Heart J* **37**, 3232-3245, doi:10.1093/eurheartj/ehw334 (2016).
- 355 3 Badimon, L. & Vilahur, G. Thrombosis formation on atherosclerotic lesions and
356 plaque rupture. *J Intern Med* **276**, 618-632, doi:10.1111/joim.12296 (2014).
- 357 4 Andersson, C. & Vasan, R. S. Epidemiology of cardiovascular disease in young
358 individuals. *Nat Rev Cardiol* **15**, 230-240, doi:10.1038/nrcardio.2017.154 (2018).
- 359 5 Elliot, C. A. & Hamlin, M. J. Combined diet and physical activity is better than diet or
360 physical activity alone at improving health outcomes for patients in New Zealand's
361 primary care intervention. *BMC Public Health* **18**, 230, doi:10.1186/s12889-018-
362 5152-z (2018).
- 363 6 Lim, S. S. *et al.* A comparative risk assessment of burden of disease and injury
364 attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a
365 systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2224-
366 2260, doi:10.1016/s0140-6736(12)61766-8 (2012).
- 367 7 Ladejobi, A. O. *et al.* The 12-lead electrocardiogram as a biomarker of biological age.
368 *Eur Heart J - Dig Health* **2**, 379-389, doi:10.1093/ehjdh/ztab043 (2021).
- 369 8 Attia, Z. I. *et al.* Age and Sex Estimation Using Artificial Intelligence From Standard
370 12-Lead ECGs. *Circulation. Arrhythmia and electrophysiology* **12**, e007284,
371 doi:10.1161/circep.119.007284 (2019).

- 372 9 Ball, R. L., Feiveson, A. H., Schlegel, T. T., Starc, V. & Dabney, A. R. Predicting
373 "heart age" using electrocardiography. *J Pers Med* **4**, 65-78, doi:10.3390/jpm4010065
374 (2014).
- 375 10 Lima, E. M. *et al.* Deep neural network-estimated electrocardiographic age as a
376 mortality predictor. *Nature communications* **12**, 5117, doi:10.1038/s41467-021-
377 25351-7 (2021).
- 378 11 Lopez-Gonzalez, A. A. *et al.* Effectiveness of the Heart Age tool for improving
379 modifiable cardiovascular risk factors in a Southern European population: a
380 randomized trial. *Eur J Prev Cardiol* **22**, 389-396, doi:10.1177/2047487313518479
381 (2015).
- 382 12 Soureti, A., Hurling, R., Murray, P., van Mechelen, W. & Cobain, M. Evaluation of a
383 cardiovascular disease risk assessment tool for the promotion of healthier lifestyles.
384 *Eur J Cardiovasc Prev Rehab* **17**, 519-523, doi:10.1097/HJR.0b013e328337ccd3
385 (2010).
- 386 13 Groenewegen, K. A. *et al.* Vascular age to determine cardiovascular disease risk: A
387 systematic review of its concepts, definitions, and clinical applications. *Eur J Prev*
388 *Cardiol* **23**, 264-274, doi:10.1177/2047487314566999 (2016).
- 389 14 Parkes, G., Greenhalgh, T., Griffin, M. & Dent, R. Effect on smoking quit rate of
390 telling patients their lung age: the Step2quit randomised controlled trial. *BMJ* **336**,
391 598-600, doi:10.1136/bmj.39503.582396.25 (2008).
- 392 15 Maanja, M. *et al.* The electrical determinants of increased wall thickness and mass in
393 left ventricular hypertrophy. *J Electrocardiol* **58**, 80-86,
394 doi:10.1016/j.jelectrocard.2019.09.024 (2020).
- 395 16 Schlegel, T. T. *et al.* Accuracy of advanced versus strictly conventional 12-lead ECG
396 for detection and screening of coronary artery disease, left ventricular hypertrophy and

- 397 left ventricular systolic dysfunction. *BMC cardiovasc disord* **10**, 28,
398 doi:10.1186/1471-2261-10-28 (2010).
- 399 17 Johnson, K. *et al.* Advanced Electrocardiography Identifies Left Ventricular Systolic
400 Dysfunction in Non-Ischemic Cardiomyopathy and Tracks Serial Change over Time. *J*
401 *Cardiovasc Dev Dis* **2**, 93-107, doi:10.3390/jcdd2020093 (2015).
- 402 18 Vrtovec, B., Sinkovec, M., Starc, V., Radovancevic, B. & Schlegel, T. T. Coronary
403 artery disease alters ventricular repolarization dynamics in type 2 diabetes. *Pacing*
404 *Clin Electrophysiol* **28 Suppl 1**, S178-181, doi:10.1111/j.1540-8159.2005.00076.x
405 (2005).
- 406 19 Persson, E., Carlsson, M., Palmer, J., Pahlm, O. & Arheden, H. Evaluation of left
407 ventricular volumes and ejection fraction by automated gated myocardial SPECT
408 versus cardiovascular magnetic resonance. *Clin Physiol Funct Imaging* **25**, 135-141,
409 doi:10.1111/j.1475-097X.2005.00599.x (2005).
- 410 20 Trägårdh, E. *et al.* High-frequency electrocardiogram analysis in the ability to predict
411 reversible perfusion defects during adenosine myocardial perfusion imaging. *J*
412 *Electrocardiol* **40**, 510-514, doi:10.1016/j.jelectrocard.2007.03.242 (2007).
- 413 21 Lang, R. M. *et al.* Recommendations for cardiac chamber quantification by
414 echocardiography in adults: an update from the American Society of
415 Echocardiography and the European Association of Cardiovascular Imaging. *Eur*
416 *Heart J Cardiovasc Imaging* **16**, 233-270, doi:10.1093/ehjci/jev014 (2015).
- 417 22 Yamazaki, T., Froelicher, V. F., Myers, J., Chun, S. & Wang, P. Spatial QRS-T angle
418 predicts cardiac death in a clinical population. *Heart rhythm* **2**, 73-78,
419 doi:10.1016/j.hrthm.2004.10.040 (2005).

- 420 23 Fayn, J., Rubel, P., Pahlm, O. & Wagner, G. S. Improvement of the detection of
421 myocardial ischemia thanks to information technologies. *Int J Cardiol.* **120**, 172-180,
422 doi:10.1016/j.ijcard.2006.09.025 (2006).
- 423 24 Kardys, I. *et al.* Spatial QRS-T angle predicts cardiac death in a general population.
424 *European heart journal* **24**, 1357-1364, doi:10.1016/S0195-668X(03)00203-3 (2003).
- 425 25 Horinaka, S. *et al.* Ventricular gradient variability. New ECG method for detection of
426 ischemic heart disease. *J Electrocardiol* **28**, 177-183, doi:10.1016/S0022-
427 0736(05)80255-9 (1995).
- 428 26 Kors, J. A., van Herpen, G., Sittig, A. C. & van Bommel, J. H. Reconstruction of the
429 Frank vectorcardiogram from standard electrocardiographic leads: diagnostic
430 comparison of different methods. *Eur Heart J* **11**, 1083-1092 (1990).
- 431 27 Zabel, M. *et al.* Analysis of T-wave morphology from the 12-lead electrocardiogram
432 for prediction of long-term prognosis in male US veterans. *Circulation (New York,*
433 *N.Y.)* **105**, 1066-1070, doi:10.1161/hc0902.104598 (2002).
- 434 28 Okin, P. M. *et al.* Repolarization Abnormality for Prediction of All-Cause and
435 Cardiovascular Mortality in American Indians: The Strong Heart Study. *J Cardiovasc*
436 *Electrophysiol* **16**, 945-951, doi:10.1111/j.1540-8167.2005.40808.x (2005).
- 437 29 Priori, S. G. *et al.* Evaluation of the spatial aspects of T-wave complexity in the long-
438 QT syndrome. *Circulation (New York, N.Y.)* **96**, 3006-3012,
439 doi:10.1161/01.CIR.96.9.3006 (1997).
- 440 30 Chhabra, L., Devadoss, R., Chaubey, V. K. & Spodick, D. H. Interatrial Block in the
441 Modern Era. *Curr Cardiol Rev* **10**, 181-189,
442 doi:10.2174/1573403X10666140514101748 (2014).

- 443 31 Lindow, T. & Lindqvist, P. The prevalence of advanced interatrial block and its
444 relationship to left atrial function in patients with transthyretin cardiac amyloidosis. *J*
445 *Clin Med* **10**, doi:10.3390/jcm10132764 (2021).
- 446 32 Rautaharju, P. M., Mason, J. W. & Akiyama, T. New age- and sex-specific criteria for
447 QT prolongation based on rate correction formulas that minimize bias at the upper
448 normal limits. *Int J Cardiol* **174**, 535-540, doi:10.1016/j.ijcard.2014.04.133 (2014).
- 449 33 Shah, S. R., Park, K. & Alweis, R. Long QT Syndrome: A comprehensive review of
450 the literature and current evidence. *Curr Prob Cardiol* **44**, 92-106,
451 doi:10.1016/j.cpcardiol.2018.04.002 (2019).
- 452 34 Nielsen, J. B. *et al.* Risk prediction of cardiovascular death based on the QTc interval:
453 evaluating age and gender differences in a large primary care population. *Eur Heart J*
454 **35**, 1335-1344, doi:10.1093/eurheartj/ehu081 (2014).
- 455 35 Goorakani, Y. *et al.* Correlation of resting heart rate with anthropometric factors and
456 serum biomarkers in a population-based study: Fasa PERSIAN cohort study. *BMC*
457 *cardiovasc disord* **20**, 319-319, doi:10.1186/s12872-020-01594-y (2020).
- 458 36 Ogliari, G. *et al.* Resting heart rate, heart rate variability and functional decline in old
459 age. *Can Med Assoc J* **187**, E442-e449, doi:10.1503/cmaj.150462 (2015).
- 460 37 Morcet, J. F., Safar, M., Thomas, F., Guize, L. & Benetos, A. Associations between
461 heart rate and other risk factors in a large French population. *J Hypertens* **17**, 1671-
462 1676, doi:10.1097/00004872-199917120-00003 (1999).
- 463 38 Evans, J. G., Prior, I. A. & Tunbridge, W. M. Age-associated change in QRS axis:
464 intrinsic or extrinsic ageing? *Gerontology* **28**, 132-137, doi:10.1159/000212523
465 (1982).

- 466 39 Bergfeldt, L. *et al.* Spatial peak and mean QRS-T angles: A comparison of similar but
467 different emerging risk factors for cardiac death. *J Electrocardiol* **61**, 112-120,
468 doi:10.1016/j.jelectrocard.2020.05.013 (2020).
- 469 40 Ringborn, M. M. D. *et al.* Comparison of high-frequency QRS components and ST-
470 segment elevation to detect and quantify acute myocardial ischemia. *J Electrocardiol*
471 **43**, 113-120, doi:10.1016/j.jelectrocard.2009.11.009 (2010).
- 472 41 Trägårdh, E., Pahlm, O., Wagner, G. S. & Pettersson, J. Reduced high-frequency QRS
473 components in patients with ischemic heart disease compared to normal subjects. *J*
474 *Electrocardiol* **37**, 157-162, doi:10.1016/j.jelectrocard.2004.02.004 (2004).
- 475 42 Schlegel, T. T. *et al.* Real-Time 12-Lead High-Frequency QRS Electrocardiography
476 for Enhanced Detection of Myocardial Ischemia and Coronary Artery Disease. *Mayo*
477 *Clin Proc* **79**, 339-350, doi:10.4065/79.3.339 (2004).
- 478 43 Gladding, P. A., Hewitt, W. & Schlegel, T. T. Going deep with ecg and aortic
479 stenosis: Touchdown or incomplete pass? *J Am Heart Assoc* **9**, e016193-e016193,
480 doi:10.1161/JAHA.120.016193 (2020).
- 481 44 The Lancet Respiratory Medicine. Opening the black box of machine learning. *Lancet*
482 *Resp Med* **6**, 801-801, doi:10.1016/S2213-2600(18)30425-9 (2018).
- 483 45 Yoon, C. H., Torrance, R. & Scheinerman, N. Machine learning in medicine: should
484 the pursuit of enhanced interpretability be abandoned? *J Med Ethics*,
485 doi:10.1136/medethics-2020-107102 (2021).
- 486 46 Brisk, R. *et al.* The effect of confounding data features on a deep learning algorithm to
487 predict complete coronary occlusion in a retrospective observational setting. *Eur*
488 *Heart J - Dig Health* **2**, 127-134, doi:10.1093/ehjdh/ztab002 (2021).

489 47 Siontis, K. C. *et al.* Use of Artificial Intelligence Tools Across Different Clinical
490 Settings: A Cautionary Tale. *Circulation. Cardiovasc Qual Outcomes* **14**, e008153,
491 doi:10.1161/circoutcomes.121.008153 (2021).

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

Table 1. Baseline characteristics

	All (n=2,771)	Healthy (n=1,682)	Subjects at CV risk (n=305)	Patients with CV disease (n=784)	p
Age, years	46.2 [16.0]	38.6 [13.0]	54.8 [11.2]	59.0 [13.2]	<0.001
Male sex, n (%)	1645 [59.4]	999 [59.4]	163 [53.4]	483 [61.6]	0.048
5-minute ECG Heart					
Age, years	51.1 [21.7]	38.3 [14.1]	61.9 [13.2]	74.2 [15.6]	<0.001
CAD, n (%)	421 [15.2]	-	-	421 [53.7]	-
ICM, n (%)	120 [4.3]	-	-	120 [15.3]	-
HCM, n (%)	92 [3.3]	-	-	92 [11.7]	-
LVH, n (%)	96 [3.5]	-	-	96 [12.2]	-
NICM, n (%)	53 [1.9]	-	-	53 [6.8]	-

Age and 5-minute ECG Heart Age are presented as mean (standard deviation).

Abbreviations: CAD: coronary artery disease; ICM: ischemic cardiomyopathy; HCM: hypertrophic cardiomyopathy; LVH: left ventricular hypertrophy; NICM: non-ischemic cardiomyopathy

509

510

511

512

513

514

515

Table 2. ECG measures in all patients, stratified by health, cardiovascular risk or cardiovascular disease

Measure	All (n=2,771)	Healthy (n=1,682)	Subjects at CV risk (n=305)	Patients with CV disease (n=784)	p
P-wave duration, ms	105 [15]	99 [13]	109 [14]	114 [15]	<0.001
Spatial QT interval, ms	432 [38]	428 [32]	425 [37]	445 [46]	<0.001
Heart rate, min ⁻¹	65 [11]	63 [10]	70 [11]	67 [13]	<0.001
QRS max amplitude in VCG lead Y, μ V	988 [454]	1144 [392]	828 [363]	717 [462]	<0.001
Frontal plane QRS axis, radians	0.64 [0.44]	0.82 [0.23]	0.54 [0.44]	0.29 [0.54]	<0.001
T-wave complexity, Ln $\sum(\text{EV3:8}) / (\text{EV1-EV2})$, unitless	-2.9 [0.7]	-3.2 [0.5]	-2.9 [0.6]	-2.3 [0.9]	<0.001
Spatial ventricular gradient minus Spatial mean QRS, mV*s	0.044 [0.030]	0.059 [0.021]	0.038 [0.017]	0.016 [0.028]	<0.001
QRS RMS in VCG vector magnitude lead, mV	0.52 [0.15]	0.53 [0.13]	0.47 [0.11]	0.50 [0.21]	<0.001
QRS average spatial velocity in VCG vector magnitude lead, mV/s	45 [12]	46 [10]	43 [16]	43 [10]	<0.001
Portion of QRS loop in posterior superior quadrant of left sagittal plane by VCG, %	11 [15]	6 [8]	11 [15]	17 [22]	<0.001

All values are presented as mean (standard deviation).

Abbreviations: CV: cardiovascular; Ln: Natural logarithm; EV: Eigenvalues; VCG: vectocardiographic; RMS: root mean square

Table 3. Measures included in the 10-second ECG Heart Age for male subjects.

Measure	Estimate	Standard error	t ratio
<i>(Intercept)</i>	-60.575	2.353	-25.73
Age, years	0.820	0.007	113.73
P-wave duration, ms	0.288	0.007	39.99
Spatial QT interval, ms	0.081	0.004	22.86
Heart rate, min ⁻¹	0.242	0.012	20.02
QRS max amplitude in VCG lead Y, μ V	-0.007	0.000	-18.20
Frontal plane QRS axis, radians	-5.834	0.395	-14.75
T-wave complexity, $\text{Ln} \sum(\text{EV}_{3:8}) / (\text{EV}_1 - \text{EV}_2)$, unitless	2.152	0.164	13.12
Spatial ventricular gradient minus Spatial mean QRS, mV*s	-55.841	4.464	-12.51
Male sex (value=1)	1.996	0.215	9.27
QRS RMS in VCG vector magnitude lead, mV	10.315	1.612	6.40
QRS average spatial velocity in VCG vector magnitude lead, mV/s	0.106	0.017	6.30
Portion of QRS loop in posterior superior quadrant of left sagittal plane by VCG, %	0.033	0.009	3.55

Abbreviations: Ln: Natural logarithm; EV: Eigenvalues; VCG: vectocardiographic; RMS: root mean square

Table 4. Measures included in the 10-second ECG Heart Age for female subjects.

Measure	Estimate	Standard error	t ratio
<i>(Intercept)</i>	-70.477	2.662	-26.47
Age, years	0.850	0.008	104.26
P-wave duration, ms	0.331	0.008	40.73
Spatial QT interval, ms	0.095	0.004	23.58
Heart rate, min ⁻¹	0.281	0.014	20.57
QRS max amplitude in VCG lead Y, μ V	-0.009	0.000	-18.54
Frontal plane QRS axis, radians	-6.354	0.447	-14.21
Spatial ventricular gradient minus Spatial mean QRS, mV*s	-67.642	5.049	-13.40
T-wave complexity, $\text{Ln} \sum(\text{EV}3:8) / (\text{EV}1-\text{EV}2)$, unitless	2.359	0.185	12.72
QRS RMS in VCG vector magnitude lead, mV	11.701	1.822	6.42
QRS average spatial velocity in VCG vector magnitude lead, mV/s	0.116	0.019	6.10
Portion of QRS loop in posterior superior quadrant of left sagittal plane by VCG, %	0.034	0.010	3.22

Abbreviations: Ln: Natural logarithm; EV: Eigenvalues; VCG: vectocardiographic; RMS: root mean square

517

518

519

520

521

522

523

524 **Figure 1.** Left panel: Scatter plot showing the relation between the 10-second ECG Heart Age
525 and the 5-minute ECG Heart Age in all participants. The R^2 value was 0.94 ($p < 0.001$). Right
526 panel: Bland-Altman plot showing the difference between the 10-second and 5-minute ECG
527 Heart Age in relation to the mean of both ECG Heart Ages. The agreement between methods
528 is strong, with minimal deviation from the identity line (dashed) or bias (0.0 ± 5.2 years).

529

530 **Figure 2.**

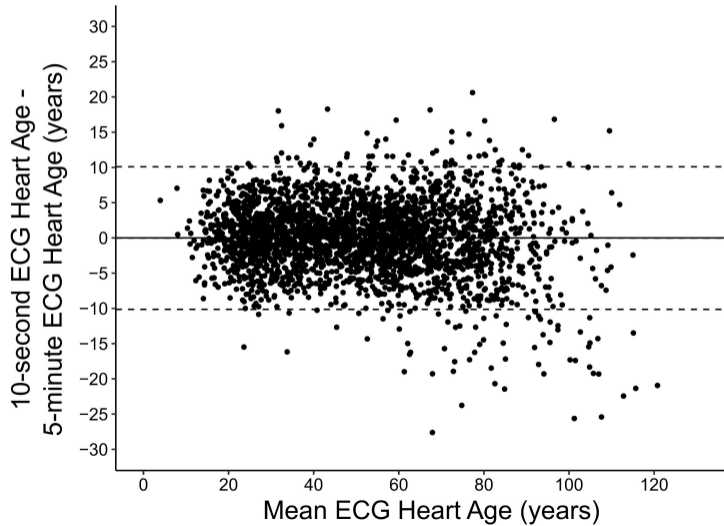
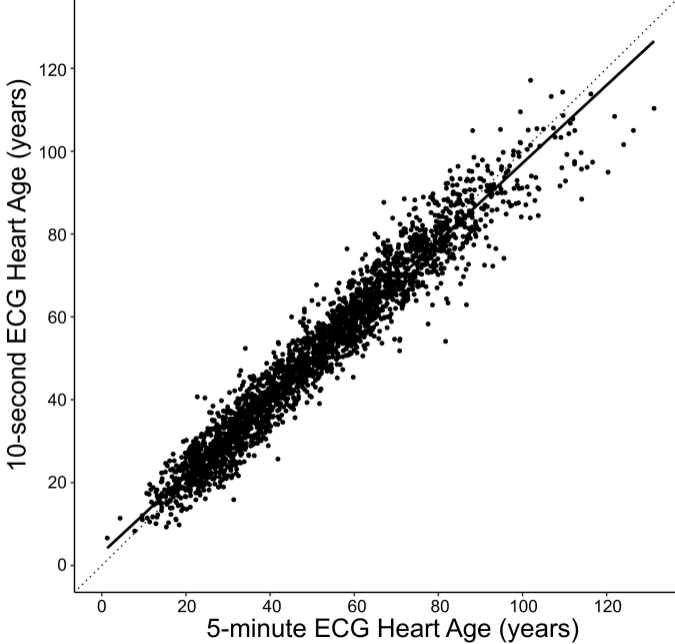
531 *Panel A:* The difference (the Heart Age Gap) between the 10-second ECG Heart Age and
532 chronological age in healthy subjects (left, dark green), subjects at cardiovascular (CV) risk
533 (middle, light blue), and patients with CV disease (right, yellow). On average, there is no
534 difference between Heart Age and chronological age in healthy subjects. Heart Age Gap is
535 higher in subjects at CV risk, and highest for those with overt CV disease.

536 *Panel B:* Scatter plots showing the relationship between the 10-second ECG Heart Age and
537 chronological age in healthy subjects (left, dark green), subjects at CV risk (middle, light
538 blue), and patients with CV disease (right, yellow). The dashed diagonal line is the identity
539 line, i.e. indicating no difference between Heart Age and chronological age.

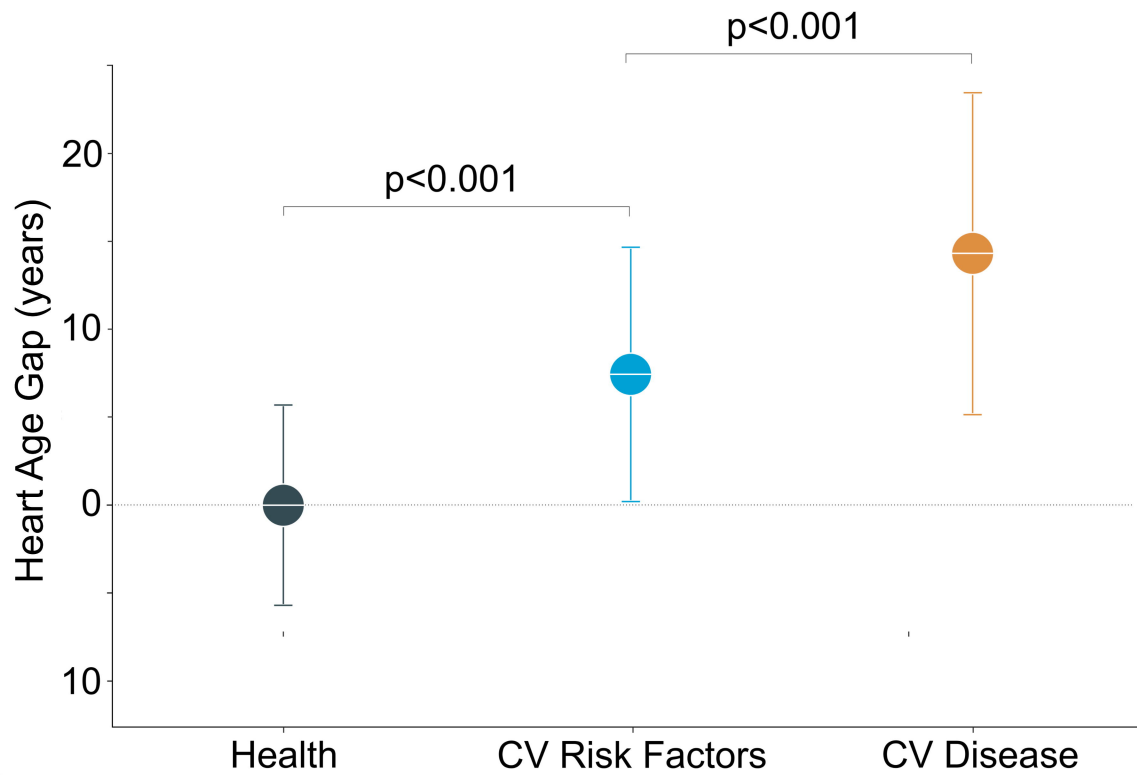
540

541 **Figure 3.** Example of the transparency and explainability of the Heart Age from two subjects
542 with equal chronological age but different Heart Ages, illustrated by a commensurately aged
543 female heart, but a disproportionately aged male heart. The ECG measures for each of the two
544 patients are shown in the table in the middle of the figure, presented in the order of the
545 relative strength (strongest first, based on t ratio [not shown]) of contribution to the Heart
546 Age. Notably, P-wave duration is markedly different between these two patients, and heart
547 rate is higher for the male than the female, helping drive the Heart Age higher in the male.
548 The R-wave amplitude in lead Y is also much larger in the female, and the difference between

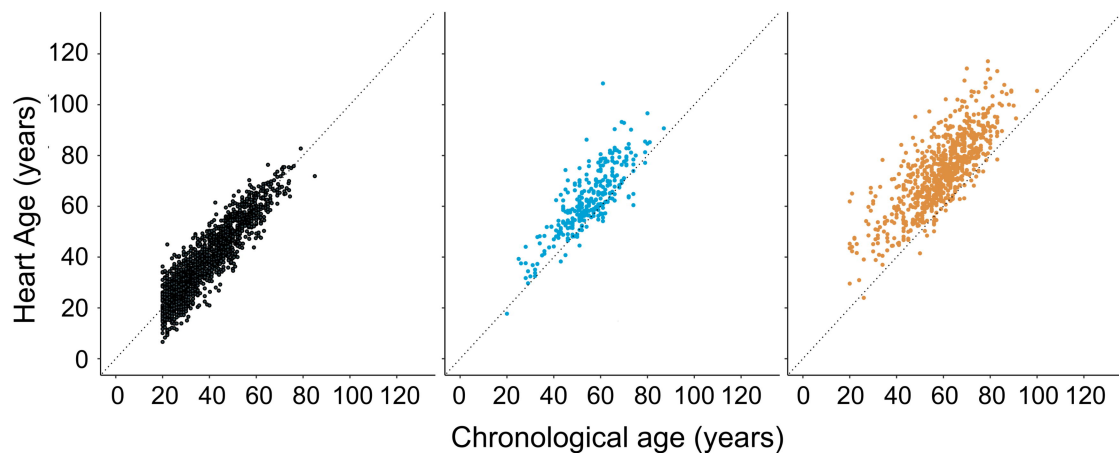
549 the spatial ventricular gradient and the spatial mean QRS in the female is also larger, likely
550 due to preserved T-wave amplitudes, contributing to her relatively younger Heart Age.
551 Furthermore, possibly due to ischemic myocardial injuries, T-wave complexity is increased in
552 the male, suggesting that increased myocardial repolarisation heterogeneity also contributes to
553 driving Heart Age higher in the male, in spite of his shorter QT interval.

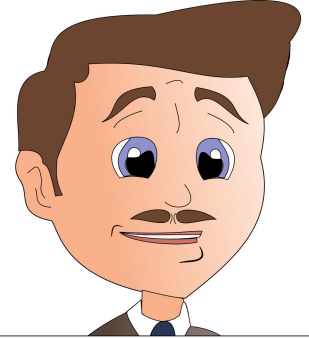


A



B





60-year old, healthy female
without cardiovascular
risk factors or medications

60-year old male with
heart failure following
myocardial infarction

ECG measures

P-wave duration	90 ms	148 ms
Spatial QT interval	447 ms	380 ms
Heart rate	55/min	70/min
QRS max amplitude in VCG lead Y	835 μ V	299 μ V
Frontal plane QRS axis	0.58 radians	0.64 radians
Spatial ventricular gradient - spatial mean QRS	0.038 mV*s	0.003 mV*s
T-wave complexity, Ln	-1.92	1.21
QRS RMS in VCG vector magnitude lead	0.43 mV	0.31 mV
QRS average spatial velocity in vector magnitude lead	39.1 mV/s	23.5 mV/s
Portion of QRS loop in posterior superior quadrant of left sagittal plane by VCG	12.7%	25.2%

ECG HEART AGE

60 years

79 years

