

24 **Abstract**

25 Biological aging is marked by a decline in resilience at the cellular and systemic levels, driving an
26 exponential increase in mortality risk. Here, we evaluate several clinical and epigenetic clocks for their
27 ability to predict mortality, demonstrating that clocks trained on survival and functional aging
28 outperform those trained on chronological age. We present an enhanced clinical clock that predicts
29 mortality more accurately and provides actionable insights for guiding personalized interventions.
30 These findings highlight the potential of mortality-predicting clocks to inform clinical decision-making
31 and promote strategies for healthy longevity.

32 **Main**

33 Biological aging is characterized by the progressive decline in intrinsic biological resilience that is
34 associated with an exponential increase in mortality, expressed in the demographic “Gompertz
35 mortality law”¹. Not all humans age at the same rate since genetics, lifestyle and stochastic factors
36 can affect future mortality and morbidity trajectories. Consequently, individual true biological age (BA)
37 is not identical to calendar or chronological age (CA). The true BA of an individual can be uniquely
38 defined as the age at which subjects of a reference cohort have the same risk of age-dependent
39 disease and all-cause mortality as the subject in question. Tools to accurately track changes in true
40 BA are essential for the development and validation of novel life- and healthspan-optimizing diet,
41 lifestyle, supplement and drug interventions.

42

43 Biological aging “clocks” are computational tools that estimate individual true BA based on
44 demographic, clinical, and/or molecular data. CA itself is widely used for both clinical prognostication
45 and decision-making, and can be viewed as a first order approximation of true BA. The ideal BA clock
46 should predict individual Gompertz mortality risk with higher accuracy than CA. Some aging clocks,
47 including most clinical clocks, explicitly include CA as a covariate, using biological features to estimate
48 a correction factor aimed at providing a better estimate of true BA. CA in this case is used as a proxy
49 for effects and mechanisms, such as entropic damage, not captured by the clock itself. Of course, the
50 ideal clock would include all relevant processes, wherein the model would assign zero or negligible
51 weight to CA.

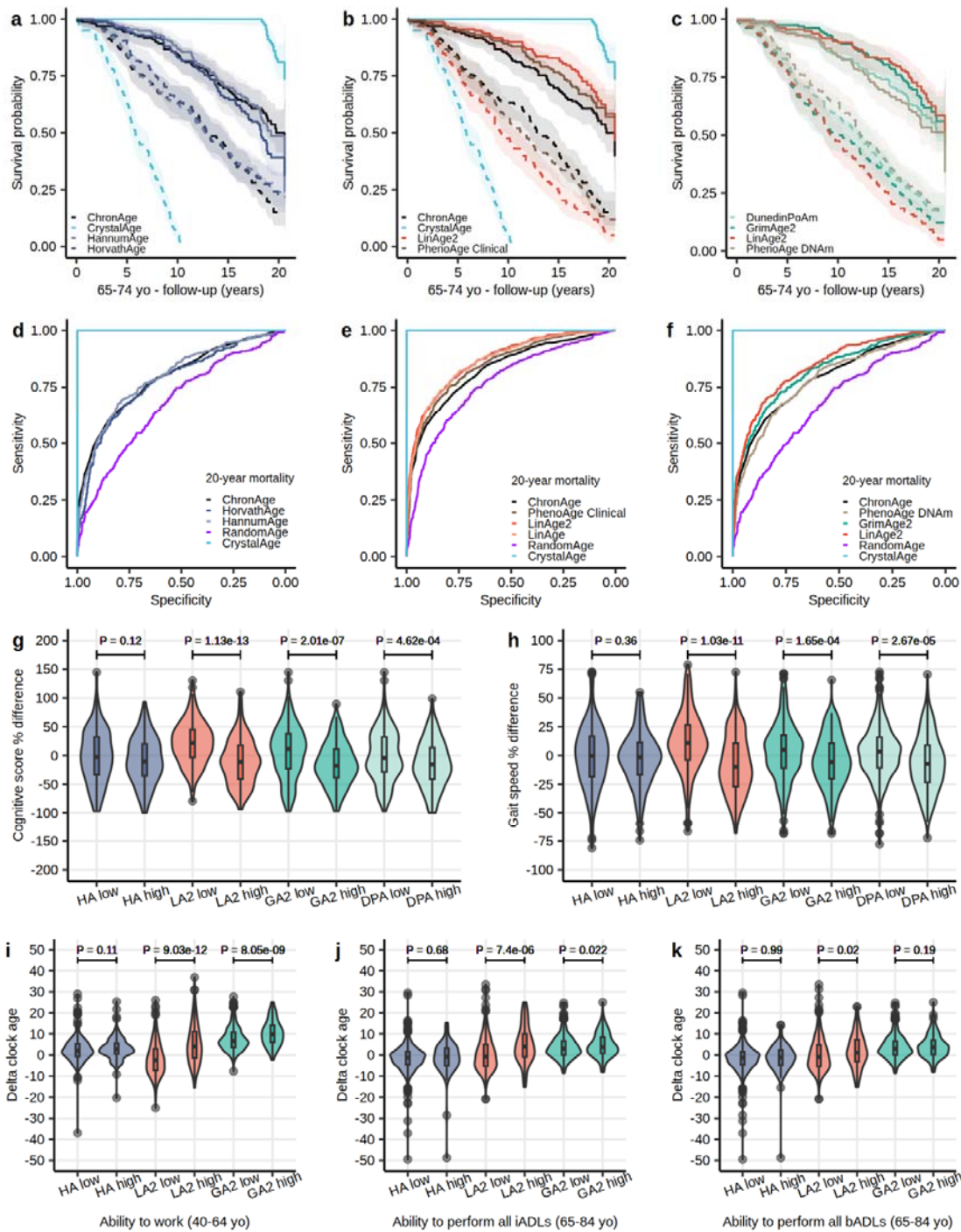
52

53 Aging clocks are generalizations of current clinical risk markers that predict disease-specific morbidity
54 and, in some cases, mortality. Aging clocks should similarly enable early detection of hidden or
55 subclinical diseases, surpassing the capabilities of diagnostics by identifying disease processes years
56 or decades before overt disease is present. Secondly, to inform risk-to-benefit estimates (clinical
57 equipoise), aging clocks should capture all-cause mortality holistically, providing value beyond organ
58 or disease-specific risks. Thirdly, aging clocks must be sensitive to individual variations in biological
59 resilience. Finally, aging clocks should provide tools for mechanistic interpretation and provide
60 actionable insights, facilitating targeted interventions. To date, none of the existing clocks meet all
61 these criteria.

62

63 To evaluate the performance of aging clocks, we can compare them to a hypothetical “ideal” clock,
64 which we term “CrystalAge”. This optimal clock would predict disease-specific and all-cause mortality
65 at the individual level with near-perfect accuracy, essentially forecasting an individual’s date of death
66 (Fig. 1a and d). While practically impossible, in retrospective studies, we can determine the
67 theoretically optimal performance of CrystalAge and use it as a benchmark to evaluate the
68 performance of existing aging clocks.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



69
70
71
72
73
74
75
76
77
78
79
80

Fig. 1: LinAge2 predicts 20-year all-cause mortality and tracks with healthspan markers. a-c, Kaplan-Meier survival curves showing 20-year survival in the 65-74 CA bin ($n=631$). For each clock, subjects were stratified by selecting the lowest (best, solid line) and highest (worst, dotted line) 25% quartiles for BA. Clocks within the same quartile were compared using log-rank tests with Benjamini-Hochberg correction. Areas shaded indicate 95% error bands for lines of the same color. **b,** Compared to ChronAge, use of LinAge2 BA results in a significant survival difference for the lowest 25% BA quartile ($P=6.16E-04$), but not for the highest 25% quartile ($P=0.07$). PhenoAge Clinical did not significantly outperform ChronAge in predicting survival in this age bin. **c,** LinAge2 significantly outperformed DunedinPoAm ($P=1.09E-02$) and PhenoAge DNAm ($P=1.37E-03$) in the lowest 25% BA quartile, but not GrimAge2 ($P=0.22$). In the highest 25% quartile, while LinAge2 significantly outperformed PhenoAge DNAm ($P=0.03$), the differences between LinAge2 and DunedinPoAm

81 ($P=0.11$) and GrimAge2 ($P=0.58$) did not reach statistical significance. **e**, ROC analysis revealed that
82 LinAge2 (area under the curve (AUC)=0.8684) was significantly more informative than PhenoAge
83 Clinical (AUC=0.8479, $P=6.35E-05$) and ChronAge (AUC=0.8288, $P=3.16E-10$) in predicting future
84 mortality ($n=2,036$). LinAge2 performed similarly to LinAge (AUC=0.8647). **f**, LinAge2 also
85 outperformed PhenoAge DNAm (AUC=0.7859, $P=4.44E-07$) and GrimAge2 (AUC=0.8233, $P=0.02$) in
86 predicting 20-year mortality ($n=1,065$). Although GrimAge2 outperformed ChronAge (AUC=0.7933,
87 $P=2.74E-03$) in predicting 20-year mortality, PhenoAge DNAm did not ($P=0.47$). **a,d**, HorvathAge,
88 HannumAge, and ChronAge did not significantly differ in predicting mortality risk (AUCs=0.7776,
89 0.7978 and 0.7933, respectively, $n=1,065$). ROC curves were compared using DeLong's test.
90 **a,b,d,e,f**, CrystalAge, a theoretical perfect clock shown for reference, accurately identifies individuals
91 at risk of dying (AUC=1), whereas RandomAge adds random gaussian noise of ± 10 years to CA. **g-k**,
92 Violin plots for each clock categorized into low (biologically younger/best 25% quartile) and high
93 (biologically older/worst 25% quartile) groups plotted against healthspan markers: cognitive scores
94 (digit symbol substitution test), gait speed, ability to work, and ability to perform all instrumental and
95 basic activities of daily living (iADLs and bADLs). Groups (BA high versus low) were compared using
96 two-sided t-tests. Median value, lower (25th) and upper (75th) percentiles are indicated. Lines extend
97 to ± 1.5 times interquartile range, with points outside this range drawn individually. The violin shape
98 indicates the probability density function. yo, years old. HA, HorvathAge. LA2, LinAge2. GA2,
99 GrimAge2. DPA, DunedinPoAm.

100

101 Taking inspiration from Levine's PhenoAge clinical clock², we recently developed and validated clinical
102 aging clocks (PCAge, LinAge) based on linear dimensionality reduction by matrix factorization
103 (singular value decomposition) and demonstrated them to be highly predictive in terms of future
104 disease-specific and all-cause mortality³. These clocks have since been applied in a range of clinical
105 settings and, taking advantage of user feedback, we have implemented several improvements,
106 creating an updated version of these clocks (LinAge2). Like LinAge, we trained LinAge2 in the
107 National Health and Nutrition Examination Survey (NHANES) IV 1999-2000 wave before testing it in
108 the 2001-2002 wave. LinAge2 further reduces the number of rarely measured parameters and
109 emphasizes interpretability. For a detailed description of LinAge2's features and construction, refer to
110 Methods.

111

112 Many aging clocks have been developed, with epigenetic or DNA methylation (DNAm) clocks most
113 widely recognized and well-established. Several epigenetic clocks have been commercially licensed
114 for applications, including estimating CA (HorvathAge⁴, HannumAge⁵), optimizing life insurance
115 policies (PhenoAge DNAm⁶, GrimAge⁷), and monitoring the rate of aging (DunedinPoAm⁸)⁹. Recently,
116 a dataset of pre-calculated epigenetic clock ages has been published for the NHANES 1999-2002
117 waves, permitting direct comparison of the predictive power of CA, the original LinAge, LinAge2 and
118 PhenoAge clinical clocks, and the HorvathAge, HannumAge, PhenoAge DNAm, GrimAge2¹⁰ and
119 DunedinPoAm epigenetic clocks.

120

121 To compare efficacy in predicting mortality, we performed survival and receiver operating
122 characteristic (ROC) analyses on 20- and 10-year mortality in the NHANES 2001-2002 test cohort.
123 Compared to CA, LinAge2 demonstrated significant survival differences across all age bins, whereas
124 PhenoAge Clinical did not (Fig. 1b, Extended Data Fig. 1b and e). LinAge2 performed similarly to
125 LinAge and demonstrated superior predictive power for future mortality compared to PhenoAge
126 Clinical and CA (Fig. 1e and Extended Data Fig. 2b). Surprisingly, LinAge2 also outperformed
127 PhenoAge DNAm and DunedinPoAm in predicting age-specific survival differences (Fig. 1c, Extended
128 Data Fig. 1c and f) and future mortality (Fig. 1f and Extended Data Fig. 2c). In contrast, PhenoAge
129 DNAm, HorvathAge, and HannumAge did not significantly differ from CA in predicting future mortality
130 (Fig. 1a, d and f, Extended Data Fig. 1a and d, and Extended Data Fig. 2a and c). LinAge2 and
131 GrimAge2 performed similarly in predicting future mortality (Fig. 1f and Extended Data Fig. 2c) and
132 survival across all age bins (Fig. 1c, Extended Data Fig. 1c and f).

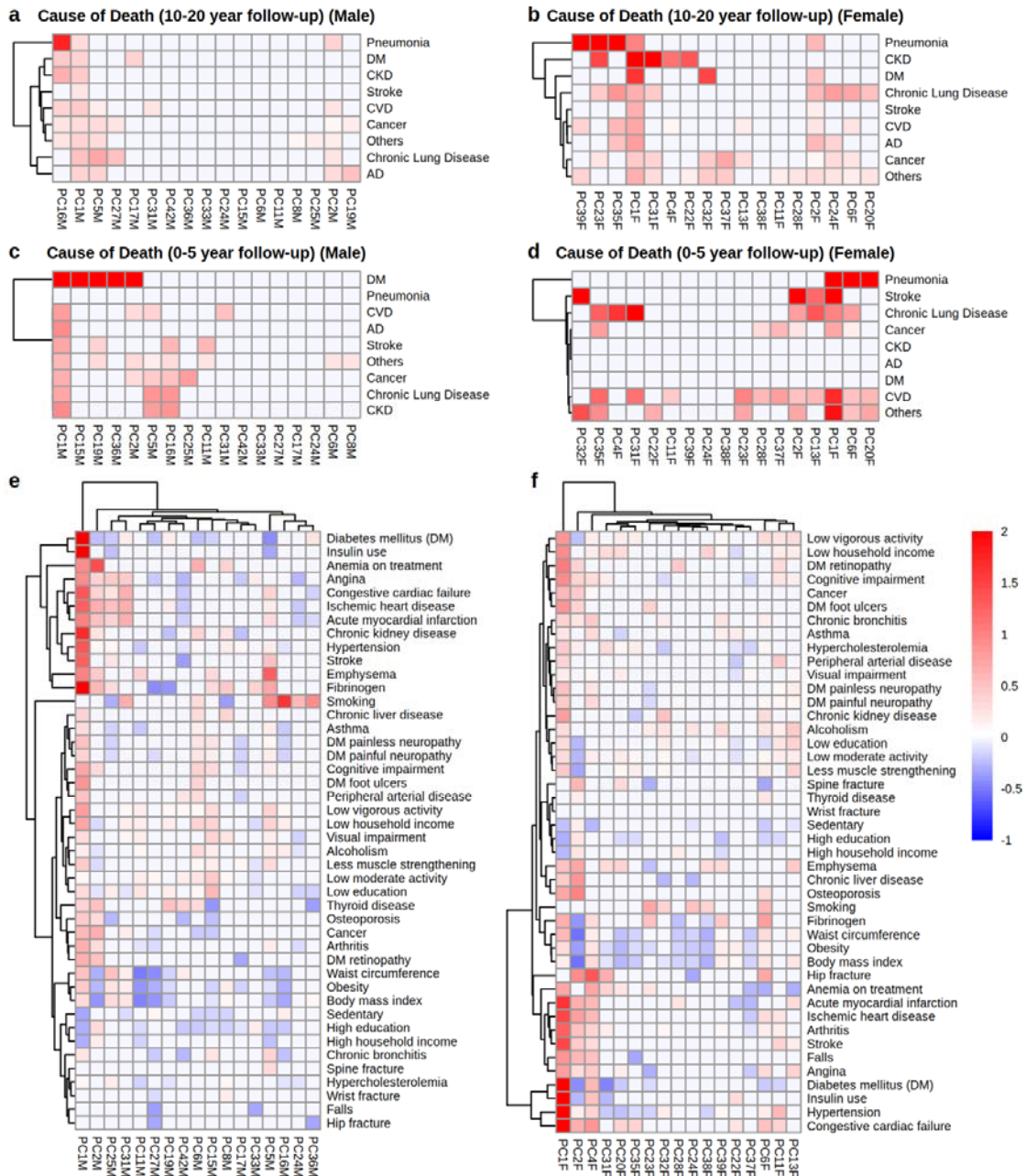
133

134 While mortality prediction is an important function of aging clocks, it is important to evaluate if clock
135 ages are similarly predictive of functional status and healthspan. We tested this for the same clinical
136 and epigenetic clocks by comparing markers of functional and health status in individuals selected by
137 each clock to be in the lowest 25% BA quartile (biologically younger, low) with those in the highest
138 25% quartile (biologically older, high). Our analysis revealed that LinAge2 low was associated with
139 superior healthspan markers, including higher cognitive scores, faster gait speed, ability to work, and
140 performance of all instrumental and basic activities of daily living (iADLs and bADLs) (Fig. 1g-k).
141 Conversely, individuals in the LinAge2 high group had poorer healthspan, with statistically significant
142 differences between the two groups across all markers (Fig. 1g-k). Similar trends were observed for
143 GrimAge2 and DunedinPoAm, with statistically significant differences between the low and high
144 groups across most healthspan markers, except for the ability to perform all bADLs (Fig. 1g-k and
145 Extended Data Fig. 3). In contrast, no statistically significant differences were found between
146 HorvathAge low and HorvathAge high across healthspan markers (Fig. 1g-k). Our findings on
147 healthspan markers and mortality, for HorvathAge, HannumAge, PhenoAge DNAm and GrimAge2,
148 corroborate similar findings for 10-year survival in 490 subjects of the Irish Longitudinal Study on
149 Aging¹¹.

150

151 A significant drawback of many existing aging clocks is that they lack interpretability and actionable
152 insights, making it challenging to develop targeted interventions. However, one key benefit of clinical
153 clocks is that they are built from parameters directly related to the underlying disease mechanisms,
154 enabling easier interpretation of clock residuals and providing actionable insights into disease
155 pathophysiology. Individual age-associated principal components (PCs) identify clusters of features
156 that change in a coordinated manner during aging. Analyzing individual PCs can provide valuable
157 insights into the underlying patterns and trajectories of aging-related changes. Using heatmaps, we
158 visualized the predictive power of individual PCs relative to clinical outcomes, including sex-specific
159 causes of death and chronic diseases (Fig. 2). Supplementary Table 5 provides detailed insights into
160 the interpretation of each PC, including associations with causes of death, chronic diseases, lifestyle
161 factors, and potential aging mechanisms, as well as suggested interventions to optimize each PC and
162 lower BA.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



163
164
165
166
167
168
169
170
171
172
173
174

Fig. 2: Heatmaps illustrating the associations between clinical outcomes and PCs analyzed using multivariate logistic regression. Associations of PCs with specific causes of death at **a-b**, 10-20 year and **c-d**, 0-5 year follow-up for male and females, respectively. Strength of association (odds ratio) is represented using a red color scale. PCs that are strongly positively associated with a specific cause of death are bright red. For cause of death, negative associations were truncated by setting their values to zero (i.e. no harm). **e-f**, Association of PCs with specific chronic diseases, and measures of lifestyle and socioeconomic status. Strength of association is represented using a blue-red scale ranging from -1 (blue, negative association) to 2 (red, positive association). PCs that are strongly positively associated (positive risk ratios) with a specific disease are bright red, whereas PCs that are strongly negatively associated (negative risk ratios) with a specific disease are bright blue.

175 Accurately predicting patient outcomes and allocating healthcare resources is a significant challenge
176 in clinical practice¹². Currently, clinicians rely heavily on CA to make these decisions. However, here

177 we show that mortality-predicting clocks, such as LinAge2 and GrimAge2, outperform CA in predicting
178 mortality risk across timeframes, ranging from 2-20 years (Fig. 1, Extended Data Fig. 4). Moreover,
179 clinical clocks can also predict specific causes of death within a 5-year window (Fig. 2c and d). This
180 illustrates that clock-based BAs are more accurate and informative estimates of true BA than CA itself.
181 By providing a more precise metric of biological status than CA alone, BA can enable clinicians to
182 better support patients and their caregivers in navigating healthcare choices, including end-of-life
183 care.

184

185 Overall, our analysis reveals that, regardless of feature space (methylation or clinical), aging clocks
186 trained to predict mortality or functional aging outcomes provide more predictive value in terms of
187 clinical decision-making. Surprisingly, clinical aging clocks still outperform several prominent mortality-
188 predicting and functional epigenetic clocks, including PhenoAge DNAm and DunedinPoAm, in
189 predicting future mortality. A key advantage of LinAge2 lies in its interpretability. Because principal
190 component analysis is a linear matrix factorization technique, the resulting model is easier to interpret
191 than nonlinear alternatives¹³. Latent variables based on linear dimensionality reduction (PCs),
192 especially those based on clinical parameters, are comparatively easy to understand and interpret,
193 making them more actionable. This enables clinical aging clocks like LinAge2 to detect hidden or
194 subclinical diseases and inform primordial prevention strategies. By identifying individuals at high risk
195 of developing specific diseases, healthcare providers can implement targeted interventions early and
196 proactively. By casting specific risk in terms of BA acceleration, aging clocks can significantly increase
197 compliance and adherence with specific health recommendations¹⁴. For example, male smokers with
198 high PC5M values in LinAge2 are at increased risk of death from chronic lung disease and should be
199 screened and advised to quit smoking (Fig. 2 and Supplementary Table 5).

200

201 All current aging clocks, regardless of feature space (e.g. clinical, methylation, proteomics, etc.) and
202 target (mortality, functional outcomes, disease, or CA) share significant limitations. Most importantly,
203 many current clocks employ linear techniques (e.g. principal component analysis/singular value
204 decomposition, regression-based predictions), which limit their ability to distinguish between aging
205 signatures and those of age-dependent diseases, and to learn U-shape response patterns. Current
206 clocks therefore inherently conflate intrinsic biological aging with disease-specific signatures (hidden

207 sickness, primordial disease signatures). Nonlinear approaches including generative artificial
208 intelligence and artificial neural networks are being investigated and could offer improved models, but
209 their increased complexity pose a significant challenge for interpretation^{15,16}. Next generation clocks
210 will need to differentiate between disease signatures and intrinsic aging, and quantify intrinsic
211 biological resilience. Further theoretical work will be required to deconvolute these disease-centric
212 signatures from determinants of intrinsic resilience and entropic aging¹⁷⁻¹⁹. Advancing next generation
213 clocks is crucial to equip healthcare providers with the essential tools needed to make informed
214 decisions regarding targeted interventions that support healthy longevity in populations where
215 healthcare needs are increasingly dominated by aging.

216 **Methods**

217 **Motivation for enhancing LinAge2**

218 The original PCAge and LinAge³ both utilized some parameters that are not routinely collected.
219 LinAge has been utilized by several clinics worldwide, and we have received informal feedback
220 regarding its use. Common suggestions for enhancing the clock include: (i) improving handling of
221 outliers and threshold effects, (ii) further refining the clinical parameters, especially removing serum
222 fibrinogen due to the need for a specialized sodium citrate tube, (iii) providing additional tools to
223 improve the interpretability of PCs, and (iv) providing specific strategies to optimize each PC to lower
224 BA. PCs can also be sensitive to outliers, thresholding and batch effects. We developed LinAge2 in
225 response to these concerns.

226

227 We followed the same workflow, as previously described³, to construct LinAge2 but with several
228 modifications. To enhance LinAge2, we refined the clinical parameters by reducing the total number to
229 60, removing serum fibrinogen (Supplementary Table 2). We also addressed outliers and thresholding
230 by capping outliers at six standard deviations and log-transforming additional parameters
231 (Supplementary Table 2). Batch effects were mitigated through z-score normalization by median and
232 median absolute deviation to a younger, generally healthy cohort (age 40-50 years), separately for
233 males and females (Supplementary Table 2), generating sex-specific PCs. The loadings for male and
234 female PCs are provided in Supplementary Table 3, and sex-specific weights of the Cox proportional
235 hazards models are listed in Supplementary Table 4.

236

237 A parametrized version of LinAge2 is provided as previously described³ (Supplementary Table 2). The
238 baseline characteristics of the study participants are listed in Supplementary Table 1.

239

240 **PhenoAge Clinical and epigenetic clocks**

241 PhenoAge Clinical was implemented using the equation from the original publication. The dataset of
242 pre-calculated epigenetic clock ages published for the NHANES 1999-2002 waves were obtained
243 from <https://wwwn.cdc.gov/nchs/nhanes/dnam/> and analyzed.

244

245 **Construction of healthspan markers**

246 The digit symbol substitution test score (NHANES variable 'CFDRIGHT') was used as a cognitive
247 measure. Gait speeds were obtained by taking the total distance walked (20 feet or 6.096 meters)
248 divided by the time taken (NHANES variable 'MSXWTIME'). Differences in cognitive scores and gait
249 speeds were calculated as the percent difference between a control group (middle 50% of all
250 subjects), for younger (best 25% quartile) and older (worst 25% quartile) groups. The ability to work
251 was established using the NHANES variable 'PFQ048'. The ability to perform all instrumental
252 activities of daily living (iADLs) was a combination of the NHANES variables 'PFQ060A', 'PFQ060F',
253 'PFQ060G', 'PFQ060Q', 'PFQ060R' and 'PFQ060S', while the ability to perform all basic activities of
254 daily living (bADLs) was a combination of the NHANES variables 'PFQ060B', 'PFQ060C', 'PFQ060H',
255 'PFQ060I', 'PFQ060J', 'PFQ060K' and 'PFQ060L'. Participants had to have either no difficulty or
256 some difficulty in all the variables to be deemed able to perform all iADLs or all bADLs.

257

258 **Heatmap analysis**

259 Using the 'nnet'²⁰ (version 7.3-19) R package, heatmaps were generated to evaluate the predictive
260 values for PCs included in LinAge2. For each parameter, we attempted to predict status
261 (diseased/compromised or not) using multivariate logistic regression with the clock PCs as covariates.
262 PCs that received a statistically insignificant ($P \geq 0.05$) weight in the logistic regression model were
263 assigned zero weights (white). The remaining PCs ($P < 0.05$) were assigned color values according to
264 their weight in the model (see Fig. 2f legend for color mapping).

265

266 **Statistics and reproducibility**

267 For the NHANES IV 1999-2002 waves, we excluded: participants top-coded at age 85 years, as we
268 could not ascertain the exact CAs of these adults, and participants who died from accidental deaths,
269 as these were deemed to be not age-related.

270

271 Survival analyses were performed using log-rank tests with Benjamini-Hochberg correction. ROC
272 curves were compared using DeLong's test. For healthspan markers, two-sided t-tests were used to
273 compare between the low and high clock groups. All statistical analyses were performed using R
274 version 4.2.0 (<https://www.R-project.org/>).

275

276 Data availability

277 All datasets used are publicly available online at <https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>.

278 There were no restrictions on data availability. This study was reported according to STROBE
279 guidelines for cohort studies.

280

281 References

- 282 1 Gompertz, B. *On the nature of the function expressive of the law of human mortality : and on*
283 *a new mode of determining the value of life contingencies ; In a letter to Francis Baily / by*
284 *Benjamin Gompertz.* (Printed by W. Nicol, 1825).
- 285 2 Liu, Z. *et al.* A new aging measure captures morbidity and mortality risk across diverse
286 subpopulations from NHANES IV: A cohort study. *PLoS Med* **15**, e1002718 (2018).
287 <https://doi.org/10.1371/journal.pmed.1002718>
- 288 3 Fong, S. *et al.* Principal component-based clinical aging clocks identify signatures of healthy
289 aging and targets for clinical intervention. *Nat Aging* **4**, 1137-1152 (2024).
290 <https://doi.org/10.1038/s43587-024-00646-8>
- 291 4 Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* **14**, R115
292 (2013). <https://doi.org/10.1186/gb-2013-14-10-r115>
- 293 5 Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human
294 aging rates. *Mol Cell* **49**, 359-367 (2013). <https://doi.org/10.1016/j.molcel.2012.10.016>
- 295 6 Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging*
296 (*Albany NY*) **10**, 573-591 (2018). <https://doi.org/10.18632/aging.101414>
- 297 7 Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*
298 (*Albany NY*) **11**, 303-327 (2019). <https://doi.org/10.18632/aging.101684>
- 299 8 Belsky, D. W. *et al.* Quantification of the pace of biological aging in humans through a blood
300 test, the DunedinPoAm DNA methylation algorithm. *Elife* **9** (2020).
301 <https://doi.org/10.7554/eLife.54870>
- 302 9 Moqri, M. *et al.* Biomarkers of aging for the identification and evaluation of longevity
303 interventions. *Cell* **186**, 3758-3775 (2023). <https://doi.org/10.1016/j.cell.2023.08.003>
- 304 10 Lu, A. T. *et al.* DNA methylation GrimAge version 2. *Aging (Albany NY)* **14**, 9484-9549 (2022).
305 <https://doi.org/10.18632/aging.204434>
- 306 11 McCrory, C. *et al.* GrimAge Outperforms Other Epigenetic Clocks in the Prediction of Age-
307 Related Clinical Phenotypes and All-Cause Mortality. *J Gerontol A Biol Sci Med Sci* **76**, 741-
308 749 (2021). <https://doi.org/10.1093/gerona/glaa286>
- 309 12 Orlovic, M. *et al.* Accuracy of clinical predictions of prognosis at the end-of-life: evidence from
310 routinely collected data in urgent care records. *BMC Palliat Care* **22**, 51 (2023).
311 <https://doi.org/10.1186/s12904-023-01155-y>
- 312 13 Sayed, N. *et al.* An inflammatory aging clock (iAge) based on deep learning tracks
313 multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nat Aging* **1**, 598-615
314 (2021). <https://doi.org/10.1038/s43587-021-00082-y>
- 315 14 Bonner, C. *et al.* Interventions Using Heart Age for Cardiovascular Disease Risk
316 Communication: Systematic Review of Psychological, Behavioral, and Clinical Effects. *JMIR*
317 *Cardio* **5**, e31056 (2021). <https://doi.org/10.2196/31056>
- 318 15 Qiu, W., Chen, H., Kaeberlein, M. & Lee, S. I. ExplaiNABLE BioLogical Age (ENABL Age): an
319 artificial intelligence framework for interpretable biological age. *Lancet Healthy Longev* **4**,
320 e711-e723 (2023). [https://doi.org/10.1016/S2666-7568\(23\)00189-7](https://doi.org/10.1016/S2666-7568(23)00189-7)
- 321 16 Nusinovi, S. *et al.* Application of a deep-learning marker for morbidity and mortality
322 prediction derived from retinal photographs: a cohort development and validation study.
323 *Lancet Healthy Longev* **5**, 100593 (2024). [https://doi.org/10.1016/S2666-7568\(24\)00089-8](https://doi.org/10.1016/S2666-7568(24)00089-8)
- 324 17 Denisov, K. A., Gruber, J. & Fedichev, P. O. Discovery of Thermodynamic Control Variables
325 that Independently Regulate Healthspan and Maximum Lifespan. *bioRxiv*,
326 2024.2012.2001.626230 (2024). <https://doi.org/10.1101/2024.12.01.626230>

- 327 18 Perevoshchikova, K. & Fedichev, P. O. Differential Responses of Dynamic and Entropic Aging
328 Factors to Longevity Interventions. *bioRxiv*, 2024.2002.2025.581928 (2024).
329 <https://doi.org/10.1101/2024.02.25.581928>
330 19 Tarkhov, A. E., Denisov, K. A. & Fedichev, P. O. Aging clocks, entropy, and the challenge of
331 age reversal. *AgingBio* 2, e20240031 (2024). <https://doi.org/10.59368/agingbio.20240031>
332 20 Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. Fourth edn, (Springer,
333 2002).
334

335 **Acknowledgements**

336 We thank the National Health and Nutrition Examination Survey participants and staff who made this
337 study possible. We thank C. Chen for her careful reading of this manuscript. This research was
338 funded by the Ministry of Education in Singapore, grant numbers IG21-SG007 and A-0007215-00-00,
339 to J.G. S.F. is supported by the Research Training Fellowship (MOH-001294-00) from the National
340 Medical Research Council Singapore. This work was supported by the Lien Foundation.

341

342 **Author contributions**

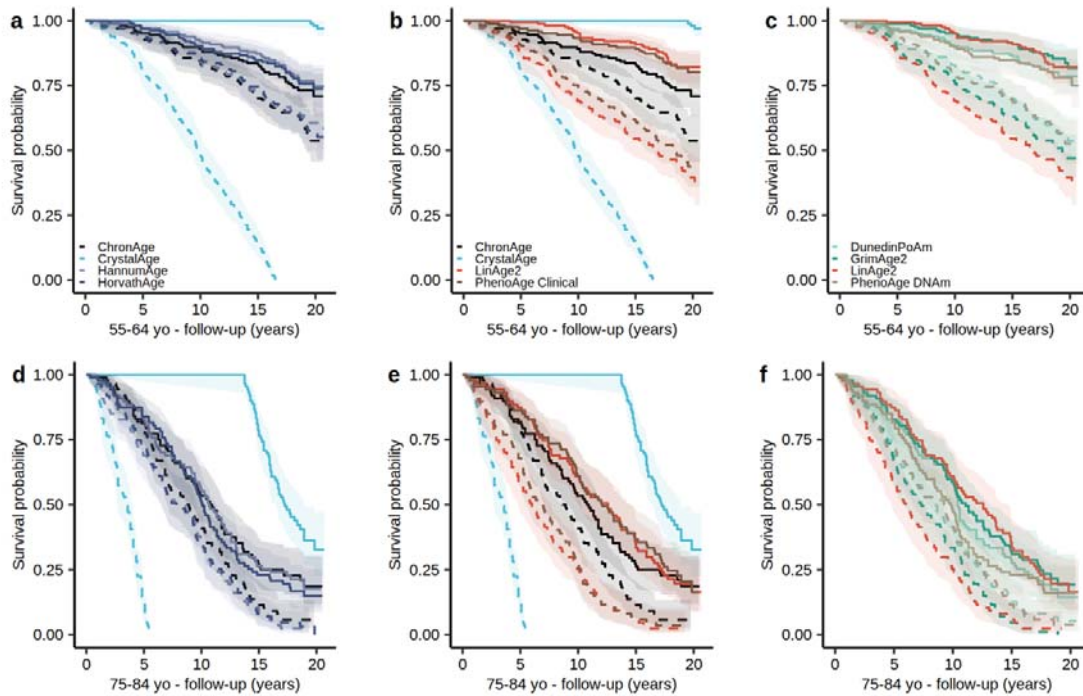
343 S.F., B.K.K. and J.G. conceived, conceptualized and designed the study. S.F., K.A.D. and J.G.
344 analyzed and interpreted the data. S.F., B.K.K. and J.G. wrote the first draft of the paper.

345

346 **Competing interests**

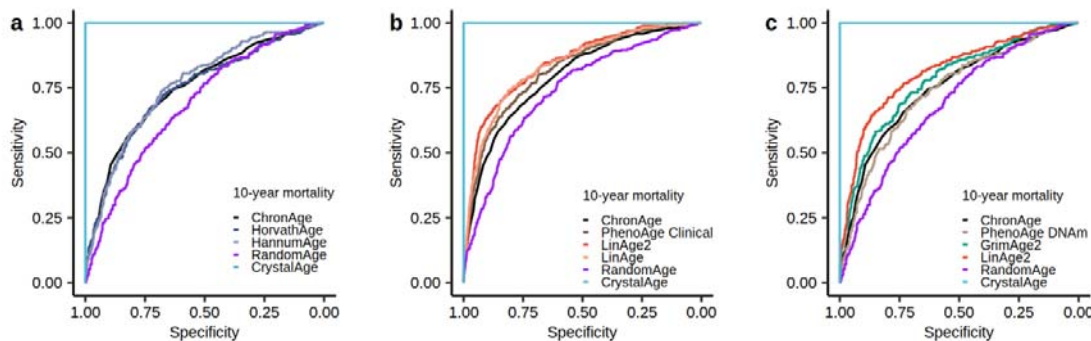
347 The authors declare no competing interests.

348 **Additional information**



349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365

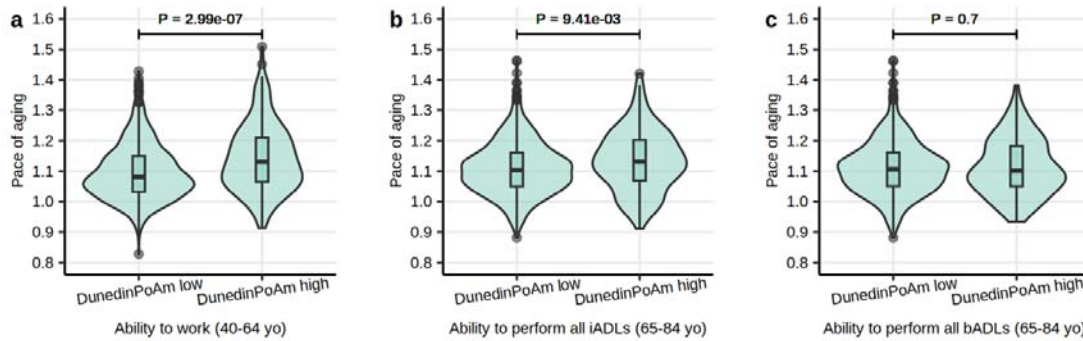
Extended Data Fig. 1: LinAge2 predicts survival in chronologically 55-64 and 75-84 year old individuals. Kaplan-Meier survival curves showing 20-year survival in the **a-c**, 55-64 CA bin ($n=657$) and **d-f**, 75-84 CA bin ($n=348$) in the test cohort. **a,d**, HannumAge, HorvathAge, and ChronAge showed no statistically significant differences in survival. **b**, LinAge2 demonstrated significant survival differences compared to ChronAge in both the best 25% ($P=2.01E-03$) and worst 25% ($P=0.03$) quartiles. In contrast, PhenoAge Clinical showed a significant difference only in the best 25% quartile ($P=3.80E-02$). **c,f**, In the best 25% quartile, LinAge2 outperformed DunedinPoAm ($P=6.45E-03$ and $P=1.39E-02$ in the 55-64 and 75-84 CA bins, respectively) and PhenoAge DNAm ($P=8.90E-03$ and $P=1.73E-02$ in the 55-64 and 75-84 CA bins, respectively). LinAge2 and GrimAge2 performed similarly with no significant differences between them. In the worst 25% quartile, no significant differences in survival were found between LinAge2, DunedinPoAm, PhenoAge DNAm, and GrimAge2. **e**, Compared to ChronAge, both LinAge2 ($P=4.56E-03$) and PhenoAge Clinical ($P=3.16E-02$) showed significant survival differences in the best 25% quartile, but not in the worst 25% quartile. Clocks were compared using log-rank tests with Benjamini-Hochberg correction. Areas shaded indicate 95% error bands for lines of the same color. yo, years old.



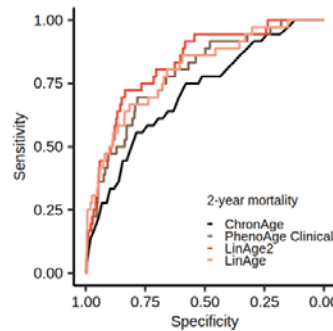
366
367
368
369
370
371
372

Extended Data Fig. 2: ROC curves for 10-year all-cause mortality in the test cohort. **a**, There were no significant differences in the AUCs between HorvathAge (AUC=0.7425), HannumAge (AUC=0.7612), and ChronAge (AUC=0.7501) ($n=1,065$). **b**, LinAge2 (AUC=0.8468) was significantly more informative than PhenoAge Clinical (AUC=0.8203, $P=6.91E-05$) and ChronAge (AUC=0.7946, $P=2.65E-09$) in predicting future mortality ($n=2,036$). LinAge2 performed similarly to LinAge (AUC=0.8383). **c**, Compared to LinAge2 (AUC=0.8144), PhenoAge DNAm (AUC=0.7390, $P=2.84E-$

373 06) and GrimAge2 (AUC=0.7801, $P=1.81E-03$) were significantly less predictive of 10-year follow-up
374 ($n=1,065$). Although GrimAge2 outperformed ChronAge (AUC=0.7501, $P=0.01$) in predicting 10-year
375 mortality, PhenoAge DNAm did not ($P=0.39$). ROC curves were compared using DeLong's test.
376



377 **Extended Data Fig. 3: DunedinPoAm tracks with healthspan markers.** Differences between
378 DunedinPoAm low (slow aging) and DunedinPoAm high (fast aging) were significant for: **a**, ability to
379 work; **b**, ability to perform instrumental activities of daily living (iADLs); but not for **c**, ability to perform
380 basic activities of daily living (bADLs). Median value, lower (25th) and upper (75th) percentiles are
381 indicated. Lines extend to ± 1.5 times interquartile range, with points outside this range drawn
382 individually. The violin shape indicates the probability density function.
383
384



385 **Extended Data Fig. 4: Clinical clocks are better predictors of 2-year mortality than CA.** ROC
386 curves for 2-year all-cause mortality in the test cohort. LinAge2 (AUC=0.8294, $P=4.02E-06$), LinAge
387 (AUC=0.8003, $P=3.91E-04$) and PhenoAge Clinical (AUC=0.7870, $P=3.24E-03$) were significantly
388 more predictive of 2-year follow-up than ChronAge (AUC=0.7120) ($n=2,036$). ROC curves were
389 compared using DeLong's test.
390