

1 Title page

2 **Genetic liability for cardiovascular disease, physical activity, and mortality**
3 **– findings from The Finnish Twin Cohort**

4 Laura Joensuu, Ph.D., Katja Waller, Ph.D., Anna Kankaanpää, M.Sc., Teemu Palviainen,
5 M.Sc., Jaakko Kaprio, M.D., Ph.D, Elina Sillanpää, Ph.D.

6

7 Author Affiliations

8 From the Faculty of Sport and Health Sciences (L.J., K.W., A.K., E.S.) and the Gerontology
9 Research Center (L.J., A.K., E.S.), University of Jyväskylä, Jyväskylä; the Institute for
10 Molecular Medicine Finland, Helsinki Institute of Life Science (T.P., J.K.), University of
11 Helsinki, Helsinki; and the Wellbeing services county of Central Finland (E.S.), Jyväskylä –
12 all in Finland.

13

14 Corresponding Author

15 Ph.D. Postdoctoral Researcher Laura Joensuu, PO Box 35, FI-40014 University of Jyväskylä,
16 Jyväskylä, Finland; laura.p.joensuu@jyu.fi, @laurajoensuu, Orcid ID: 0000-0002-9544-6552

17 **Abstract**

18 **Background**

19 An overall healthy lifestyle may mitigate the risk of all-cause and cardiovascular disease
20 (CVD) mortality despite the inherited risk for CVD. We assessed if similar phenomenon is
21 observed independently with physical activity (PA).

22

23 **Methods**

24 In a prospective older Finnish Twin Cohort, genetic liability for coronary heart disease,
25 systolic and diastolic blood pressure was estimated with polygenic risk scores (PRSs) derived
26 from the Pan UK Biobank ($N \approx 400,000$) and $>1,000,000$ genetic variants, and leisure-time PA
27 longitudinally with validated and structured questionnaires three times during 1975–1990
28 among 4,897 participants aged 33–60-years (54.3% women). Interactions of PRSs and PA
29 with mortality and the independent main effects of different PA metrics with mortality were
30 evaluated with Cox proportional hazards models. A co-twin control design with 180
31 monozygotic twin pairs discordant for physical activity was used for causal inference.

32

33 **Results**

34 During the 17.4-y (mean) follow-up (85,136 person-years), 1,195 participants died with 389
35 CVD deaths. Although the favorable associations of an overall healthy lifestyle with
36 mortality risk reduction replicated in the cohort, no interactions or significant independent
37 main effects were observed with any of the assessed PA metrics. Adherence to World Health
38 Organizations PA guidelines or vigorous PA during the 15-year observational period did not
39 decrease mortality in more active twins compared to their less active identical twin brother or
40 sister. The findings did not vary with genetic CVD risk.

41

42 **Conclusions**

43 No evidence was found that PA mitigates inherited CVD risk or is causally associated with
44 mortality challenging some of our current understanding.

45

46 (Funded by Academy of Finland; Juho Vainio, Päivikki and Sakari Sohlberg, and Sigrid

47 Juselius Foundations to E.S. Phenotype and genotype data collection by the Wellcome Trust

48 Sanger Institute, the Broad Institute, ENGAGE – European Network for Genetic and

49 Genomic Epidemiology, FP7-HEALTH-F4-2007, grant agreement number 201413, and the

50 Academy of Finland (grants 100499, 205585, 118555, 141054, 264146, 308248, 312073,

51 336823, and 352792 to J.K). Academy of Finland (grants 265240, 263278) and the Sigrid

52 Juselius Foundation to J.K)

53 **Introduction**

54 Cardiovascular diseases (CVD) remain persistently as leading causes of premature death
55 worldwide, posing a considerable burden to both public and personal health.¹ While
56 modifiable lifestyle factors contribute to the onset of the disease,^{2,3} the role of genetic factors
57 is not negligible.⁴ Family history has long been established as an independent risk factor for
58 common CVDs,⁵ with heritability estimates of 40–60% based on twin studies.⁶

59

60 Previous landmark studies have promisingly shown that by adhering to an overall healthy
61 lifestyle (regular physical activity, no obesity, no smoking, and healthy diet) the risk of CVD
62 incidence could be mitigated despite possessing a risky genotype.^{2,3} Similar findings have
63 emerged with CVD mortality.⁷ In this study, we evaluated the independent role of physical
64 activity (PA) in risk mitigation while genetic liability for CVD is assessed directly from the
65 genome. PA is commonly promoted by public health authorities to increase life expectancy,⁸
66 with consistent supporting epidemiological evidence among different demographics and with
67 different PA metrics.^{9–14} We investigated if PA during adult years mitigates the risk of
68 mortality despite the genetic liability for CVD, and evaluated evidence for causality of this
69 association.

70 **Methods**

71 *Study oversight*

72 The older Finnish Twin Cohort (FTC) study is a prospective cohort launched in 1975 and
73 described in detail elsewhere.¹⁵ In brief, all same-sex twins born before 1958 and living in
74 Finland were invited to the study with a 84.4% response rate. The study provides a unique
75 45-year follow-up resource for studying the role of lifestyle factors in health. Despite the twin
76 structure, mortality in the FTC does not differ from the Finnish population.¹⁶ The study was
77 funded with grants from Academy of Finland, and several foundations (Juho Vainio

78 foundation, Päivikki and Sakari Sohlberg foundation, and Sigrid Juselius foundation), and the
79 Wellcome Trust Sanger Institute, the Broad Institute, ENGAGE – European Network for
80 Genetic and Genomic Epidemiology); the funders did not influence or involved in any phases
81 of the study. The FTC study was evaluated by ethics committees of the University of
82 Helsinki (113/E3/01 and 346/E0/05), Helsinki University Central Hospital (136/E3/01,
83 01/2011, 270/ 13/03/01/2008, and 154/13/03/00/2011). The authors assume responsibility for
84 the accuracy and completeness of the protocol, data, and analyses, fidelity for their reporting,
85 and complying with the Declaration of Helsinki and guidelines of the Finnish Advisory
86 Board on Research Integrity.

87 ***Study design***

88 In this study, we utilised a genotyped sub-sample of the FTC (See Figures S2 and S3 for flow
89 charts). The participants answered to structured and validated questionnaires in 1975, 1981
90 and 1990 (see Figure S4 for study design). Later, participants were invited to DNA-sampling,
91 taking place at individual time points between 1993 and 2017 (Figure S5). Their vital status
92 was assessed in 2020. The follow-up began from DNA-sampling and ended at death,
93 migration, or end of the follow-up on December 31, 2020.

94 ***Study procedures***

95 In the first phase, we aimed to replicate the previous studies and assessed if an overall healthy
96 lifestyle mitigates the risk of mortality despite genetic liability for CVD in the FTC cohort. In
97 the second phase, we assessed the interactions between PA and genetic liability for CVDs,
98 and independent main effects of PA with mortality utilising different PA metrics. In the third
99 phase, we utilised a co-twin control design with monozygotic (MZ) twin pairs for causal
100 inference. In exploratory analyses we assessed if the genetic predisposition to CVD
101 associates with PA levels.

102 ***Outcomes***

103 All-cause mortality and CVD mortality were assessed. Dates and causes of death by
104 December 31st 2020 were retrieved from the Population Register Centre of Finland, and
105 Statistics Finland. All-cause mortality included all events leading to death. CVD mortality
106 was based on Statistics Finland national time series classification (54 categories). Categories
107 27–30 represent CVD mortality, including ischemic heart diseases (corresponding ICD-10
108 codes (I20–I25), other forms of heart diseases except rheumatic and alcoholism-based (I30–
109 I425, I427–I52), cerebrovascular diseases (I60–I69), hypertensive, pulmonary heart disease,
110 disease of pulmonary circulation and disease of arteries, arterioles and capillaries, veins,
111 lymphatic vessels, lymph nodes, and other unspecified disorders of the circulatory system
112 (I100–I15, I26–I28, I70–I99).^{17,18}

113 ***Genetic liability for cardiovascular diseases, and physical activity***

114 *Polygenic risk scores*

115 Polygenic risk scores (PRSs) summarise the genome-wide information between single-
116 nucleotide polymorphisms (SNPs) and a certain phenotype into a single variable that
117 quantifies person's genetic liability to a disease or a trait.¹⁹ PRSs were calculated for three
118 CVDs; coronary heart disease (CHD), systolic blood pressure (SBP) and diastolic blood
119 pressure (DBP). After evaluating a few candidates for PRSs (Tables S2-S4), PRSs based on
120 genome-wide association data from Pan-UK Biobank (restricted to participants with
121 European ancestry) and the SBayesR method were selected. Pan-UK Biobank data included
122 396,663–419,724 participants (Table S1). We used SBayesR pipeline to construct the PRSs
123 based on 1,005,933–1,006,472 genetic variants. More information related to construction of
124 the PRSs is provided in the Supplemental Appendix, Chapter 1.1.

125

126 *Physical activity*

127 PA during leisure-time was assessed with structured and validated questionnaires in 1975,
128 1981 and 1990.^{15,20,21} For utilized questions, see Supplemental Appendix Chapter 2.3. Several
129 PA metrics were formed. A metabolic equivalent of task (MET) in hours per week ($\text{h} \cdot \text{week}^{-1}$)
130 ¹) was calculated as a product of reported intensity, duration and frequency of activity.²⁰
131 Meeting the World Health Organisations PA guidelines for aerobic activity was defined as a
132 minimum of $7.5 \text{ MET h} \cdot \text{week}^{-1}$,¹¹ and calculated based on the mean $\text{MET h} \cdot \text{week}^{-1}$ in
133 1975–1990. Additional metrics were the continuous mean $\text{MET h} \cdot \text{week}^{-1}$ in 1975–1990,
134 meeting the PA guidelines in all three measurement points (1975, 1981 and 1990), and
135 adhering to vigorous PA in all three measurement points (1975, 1981 and 1990). Vigorous
136 physical activity was considered as performing at least intermittent jogging.

137 *Covariates*

138 Age, sex, educational attainment, body mass index, smoking, fruit and vegetable
139 consumption, and alcohol consumption were assessed by questionnaires. The last known
140 and/or most comprehensive data was used (1981 or 1990). Additionally, health status, i.e.,
141 self-reported symptoms, physician-diagnosed diseases and utilized medications were assessed
142 by questionnaires and from national health register data between 1971 and 1983. Ten
143 principal components of ancestry were used to adjust for any genetic stratification which may
144 occur in the Finnish population. For further details related to covariates, see Supplementary
145 Appendix, Chapter 2.3.

146 *Statistical analysis*

147 The description and rationale of the statistical procedures, along with utilized codes, are
148 provided in the Statistical Analysis Plan (SAP), which was written after accessing the data
149 and prior to final analyses. A convenience sample was used, with all assumed and eligible

150 twins in Finland being contacted to participate in 1975. In post hoc analysis, the main effects
151 analyses indicate sufficient 80% power to detect the expected 29% decrease in mortality,
152 while the power in the co-twin control design may be limited (SAP Chapter 3.2). A complete
153 case approach was selected despite potential bias to firmly retain the twin structure in the
154 data. Decision was supported by analyses where the complete case sample did not differ from
155 the full sample by PA at baseline (Table S5).

156

157 After examining proportional hazard assumptions with Shoenfeld and Martingale residuals
158 (Supplemental Appendix, Chapter 2.2), outcomes were analysed with Cox Proportional
159 Hazards Model. In the first phase, the main effects of different combinations of different
160 lifestyles (described in detail in Table S7) and genetic disease risk categories (sample-
161 specific low (<33.3%), intermediate (33.3–66.6%), or high genetic risk for CVD (> 66.6%))
162 were compared against mortality, with participants with high genetic risk and unhealthy
163 lifestyles utilized as reference category. In the second phase, the interactions between PRSs
164 and PA were examined (PRS × PA) to evaluate the potential moderator effect. Additionally,
165 the independent main effects of PA and PRSs with mortality were assessed. Models were
166 adjusted for covariates.

167

168 In the third phase, the within-pair differences in mortality were assessed among MZ twin
169 pairs discordant for their PA levels. In a co-twin control design the other twin is exposed to
170 an environmental factor (e.g., meeting selected PA criteria) while the other twin is not. This
171 difference in their exposure is evaluated against difference in their survival. The co-twin
172 control design in MZ twins is considered to be an elegant study design for causal inference in
173 epidemiological studies, as many potential confounding factors are shared among identical
174 twins (genotype and childhood environments, with many characteristics in adulthood).^{22,23}

175 Models are by design controlled for genetic factors and early life environmental factors.

176 Models with highest number of twin pairs were additionally adjusted for main lifestyle risk
177 factors; BMI, and smoking status.

178

179 In exploratory analyses, the association between genetic liability for CVD and different
180 lifestyle factors were evaluated with linear and multinomial logistic regression. Throughout
181 the study, all analyses were conducted separately for the three different genetic liabilities,
182 with results related to genetic liability for CHD presented in the main text. As diseases are a
183 potential source for reverse causality, sensitivity analyses were conducted with only
184 apparently healthy participants. Across all analyses the twin structure of the data was
185 acknowledged. Data was assumed to be missing at random. Statistical significance was set at
186 $P < 0.01$ after Bonferroni correction for multiple testing (for three different CVDs) and 99%
187 Confidence Interval reported. Analyses were performed with Stata/IC 16.0.

188 **Results**

189 *Characteristics of the study sample*

190 The descriptives of the eligible 4,897 participants are presented in Table 1 by their survival
191 status at the end of follow-up. During the 17.4-y (mean) follow-up (85,136 person-years)
192 1,195 participants died with 389 CV deaths.

193

194 *Validation and replication*

195 PRSs indicating genetic liability to CVDs indicated valid mortality risk in the FTC. In crude
196 models, one standard deviation increment in PRSs was associated with 6% to 7% and 13% to
197 17% increased risk in all-cause and CV mortality, respectively (Table S2). These associations
198 persisted after adjustments in all-cause mortality for PRS for DPB, and in CVD mortality for
199 all PRSs (Tables 2 and 3 for CHD, Tables S9–S12 for others).

200

201 The favourable associations of an overall healthy lifestyle replicated in the FTC and was
202 associated with a 43% to 61% lower risk of all-cause and 57% to 76% CVD mortality, except
203 in participants with high genetic liability for CHD (Figures S6 and S7).

204

205 *Interactions and independent main effects of PA with mortality*

206 No interactions between different PA metrics and PRSs were observed with all-cause or CVD
207 mortality (Table S8). The independent favorable associations of meeting the PA guidelines
208 with mortality gradually attenuated when accounting for other healthy lifestyles from Hazard
209 Ratio 0.85 (99% Confidence interval 0.72 to 0.99) to 0.93 (0.78 to 1.11) in all-cause mortality
210 and from 0.78 (0.59 to 1.04) to 0.92 (0.67 to 1.25) in CVD mortality (Tables 2 and 3 for
211 CHD, and S9–S12 for others). Of covariates, smoking was consistently associated with
212 higher risk of mortality, BMI only with CVD mortality and high alcohol consumption only
213 with all-cause mortality. No statistically significant associations were observed with any of
214 the additionally evaluated PA metrics (Tables S13–S18), or in sensitivity analyses with only
215 healthy participants (Tables S19–S21).

216

217 *Causal relationship between PA and mortality*

218 The twin pairs in co-twin control design did not differ from one another by other
219 characteristics than their PA (Figure 1). Meeting the PA guidelines did not statistically
220 significantly decrease the risk of all-cause 0.74 (0.37 to 1.48), or CVD mortality 0.39 (0.10 to
221 1.46) in within-pair comparison between more and the less active twins (Table 5). Findings
222 were similar across genetic liabilities for different CVDs and genetic disease risk categories
223 (Table 5, S22 and S23), in healthy twin pairs (Tables S24 and S25) and among the three
224 evaluated PA discordances (Tables S26–S29).

225

226 *Exploratory analyses*

227 The exploratory analyses did not suggest that higher genetic liability for CVDs predisposes to
228 lower PA levels. Interestingly, higher genetic liability for SBP and DBP, but not for CHD,
229 was associated with higher BMI: regression coefficient 0.24 (99% Confidence Interval 0.09
230 to 0.38), 0.27 (0.12 to 0.41) and 0.12 (-0.02 to 0.26), respectively (Table S30).

231 **Discussion**

232 In this study, we evaluated the role of PA in reducing mortality risk of genetic liability for
233 CVD. Our study supported the previous findings that adherence to an overall healthy lifestyle
234 attenuates the risk of all-cause and CVD mortality despite the inherited risk for CVD.⁷
235 However, after extensive evaluations, we did not find any independent favorable associations
236 between PA and mortality. These findings question the potential to mitigate the adverse
237 outcomes of inherited risk of CVD with PA.

238

239 The main observations are highlighted and critically discussed. Firstly, the phenomenon of
240 positive correlations between an overall healthy lifestyle, favorable health outcomes, and
241 reduced mortality despite genetic liability to CVD is reliably replicated. However, our
242 findings underline that not all lifestyle factors contribute equally, and the role of PA is
243 limited. Interestingly, this interpretation is supported by the seminal studies but
244 overshadowed by emphasis on the overall healthy lifestyle concept.^{2,3} We recommend to
245 carefully consider if overall lifestyle scores provide information that sufficiently supports the
246 development of health promotion strategies.

247

248 Secondly, our study found no evidence of interactions between PA and genetic risk in
249 relation to mortality. Although the existing literature is limited, our novel findings suggest

250 that physical activity may not moderate the association between inherited CVD risk and
251 mortality.

252

253 Thirdly, in our sample, the independent main effects of PA were attenuated in a step-wise
254 manner when adjusting for other healthy lifestyles. Physically active individuals often
255 possess also other healthy lifestyles,²⁴ and this clustering might cause unobserved
256 confounding if not appropriately considered. It is also important to acknowledge that
257 although the main risk factors (smoking, BMI) may be most relevant, the associations may
258 further attenuate with additional adjustments.

259

260 Finally, the twins who were more active during the extensive 15-year observational period
261 did not live longer compared to their less active cotwin in identical twin pairs. Findings were
262 similar with different PA metrics (meeting PA guidelines or not, adhering to vigorous PA or
263 not), and with different genetic disease risk levels for different CVDs. Assessing causality
264 between PA and mortality has been shown to be methodologically challenging,²⁵ with our
265 study being no exception. However, our study is in line with previous findings from
266 randomized controlled trials (RCTs), animal studies, and Mendelian randomization studies
267 where no convincing evidence for causal relationship between PA and mortality has been
268 observed.²⁶⁻²⁸

269

270 We acknowledge that self-reported metrics are prone to biases, and data was collected during
271 an era with potentially limited generalizability to modern lifestyles. Data lost during the long
272 observational period reduced the number of subjects in complete cases analysis, potentially
273 introducing bias. Larger sample sizes or PA discordances might have strengthened some of
274 the observed associations.

275 *“Physical activity might not bring more years to life, but more life to years”*

276 The above mentioned quote from professor Urho Kujala summarizes the public health
277 message of this study. This study does not refute well-established findings from RCTs and
278 the biological rationale that support the role of exercise and PA in primary and secondary
279 CVD risk modification.^{29,30} Notably, previous findings from the FTC show that more active
280 MZ twins manifest more favorable tissue and metabolic health alongside with better physical
281 fitness compared to their less active co-twins, supporting the findings from RCTs.²²
282 However, the discrepancy between the systematic observational findings and persistent
283 inability to confirm causality between PA and mortality requires further thought. Plausible
284 explanations are multifactorial; potential bias in observational studies may be induced by
285 healthy exerciser bias, unobserved and residual confounding, reverse causation, and genetic
286 predisposition for both lesser physical activity and increased disease risk.^{26,31–33} Hence,
287 physical activity may predominantly reflect one’s health and fitness status by conveying the
288 underlying biological milieu rather than causally reduce mortality. The limited independent
289 role of PA is evident in identical twin pairs, where higher engagement in PA does not
290 associate with longer life expectancy.

291

292 In conclusion, no evidence was found that PA mitigates the inherited CVD disease risk or is
293 causally associated with mortality. The benefits of an overall lifestyle were explained mainly
294 by lesser smoking and healthy BMI. These critical findings provide novel information for
295 future health policies.

296 **References**

- 297 1. Kyu HH, Abate D, Abate KH, et al. Global, regional, and national disability-adjusted life-
298 years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195
299 countries and territories, 1990–2017: a systematic analysis for the Global Burden of
300 Disease Study 2017. *The Lancet*. 2018;392(10159):1859-1922. doi:10.1016/S0140-
301 6736(18)32335-3
- 302 2. Khera AV, Emdin CA, Drake I, et al. Genetic Risk, Adherence to a Healthy Lifestyle, and
303 Coronary Disease. *N Engl J Med*. 2016;375(24):2349-2358. doi:10.1056/NEJMoa1605086
- 304 3. Rutten-Jacobs LC, Larsson SC, Malik R, et al. Genetic risk, incident stroke, and the
305 benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank
306 participants. *BMJ*. Published online October 24, 2018:k4168. doi:10.1136/bmj.k4168
- 307 4. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic Susceptibility to
308 Death from Coronary Heart Disease in a Study of Twins. *N Engl J Med*.
309 1994;330(15):1041-1046. doi:10.1056/NEJM199404143301503
- 310 5. Shea S, Ottman R, Gabrieli C, Stein Z, Nichols A. Family history as an independent risk
311 factor for coronary artery disease. *J Am Coll Cardiol*. 1984;4(4):793-801.
312 doi:10.1016/S0735-1097(84)80408-8
- 313 6. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U.
314 Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish
315 twins. *J Intern Med*. 2002;252(3):247-254. doi:10.1046/j.1365-2796.2002.01029.x
- 316 7. Yun JS, Jung SH, Shivakumar M, et al. *Polygenic Risk, Lifestyle, and Cardiovascular*
317 *Mortality: A Prospective Population-Based UK Biobank Study*. medRxiv; 2021.
318 doi:10.1101/2021.02.15.21251790
- 319 8. OECD, World Health Organization. *Step Up! Tackling the Burden of Insufficient Physical*
320 *Activity in Europe*. OECD; 2023. doi:10.1787/500a9601-en
- 321 9. Garcia L, Pearce M, Abbas A, et al. *Non-Occupational Physical Activity and Risk of 22*
322 *Cardiovascular Disease, Cancer, and Mortality Outcomes: A Dose-Response Meta-*
323 *Analysis of Large Prospective Studies*. medRxiv; 2022. doi:10.1101/2022.03.02.22271753
- 324 10. Jayedi A, Gohari A, Shab-Bidar S. Daily Step Count and All-Cause Mortality: A
325 Dose–Response Meta-analysis of Prospective Cohort Studies. *Sports Med*. 2022;52(1):89-
326 99. doi:10.1007/s40279-021-01536-4
- 327 11. Watts EL, Matthews CE, Freeman JR, et al. Association of Leisure Time Physical
328 Activity Types and Risks of All-Cause, Cardiovascular, and Cancer Mortality Among
329 Older Adults. *JAMA Netw Open*. 2022;5(8):e2228510.
330 doi:10.1001/jamanetworkopen.2022.28510
- 331 12. Tarp J, Fagerland MW, Dalene KE, et al. Device-measured physical activity,
332 adiposity and mortality: a harmonised meta-analysis of eight prospective cohort studies. *Br*
333 *J Sports Med*. 2022;56(13):725-732. doi:10.1136/bjsports-2021-104827

- 334 13. Paluch AE, Bajpai S, Bassett DR, et al. Daily steps and all-cause mortality: a meta-
335 analysis of 15 international cohorts. *Lancet Public Health*. 2022;7(3):e219-e228.
336 doi:10.1016/S2468-2667(21)00302-9
- 337 14. Shaked O, Cohen G, Goshen A, Shimony T, Shohat T, Gerber Y. Physical Activity
338 and Long-Term Mortality Risk in Older Adults with and without Cardiovascular Disease:
339 A Nationwide Cohort Study. *Gerontology*. 2022;68(5):529-537. doi:10.1159/000518169
- 340 15. Kaprio J, Bollepalli S, Buchwald J, et al. The Older Finnish Twin Cohort - 45 Years
341 of Follow-up. *Twin Res Hum Genet*. 2019;22(4):240-254. doi:10.1017/thg.2019.54
- 342 16. Kaprio J. The Finnish Twin Cohort Study: An Update. *Twin Res Hum Genet*.
343 2013;16(1):157-162. doi:10.1017/thg.2012.142
- 344 17. World Health Organization. International Statistical Classification of Diseases and
345 Related Health Problems 10th Revision. Accessed October 30, 2022.
346 <https://icd.who.int/browse10/2019/en#/>
- 347 18. Statistics Finland. Causes of death. Data resources catalogue. Accessed October 30,
348 2022. [https://aineistokatalogi.fi/catalog/studies/778c33bf-aceb-423f-89d9-
349 e5abb5a0585c/datasets/fcc7f03d-3f83-49df-973d-f2d94b3742a6](https://aineistokatalogi.fi/catalog/studies/778c33bf-aceb-423f-89d9-e5abb5a0585c/datasets/fcc7f03d-3f83-49df-973d-f2d94b3742a6)
- 350 19. Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk
351 score analyses. *Nat Protoc*. 2020;15(9):2759-2772. doi:10.1038/s41596-020-0353-1
- 352 20. Kujala UM, Kaprio J, Sarna S, Koskenvuo M. Relationship of Leisure-Time Physical
353 Activity and Mortality: The Finnish Twin Cohort. *JAMA*. 1998;279(6):440.
354 doi:10.1001/jama.279.6.440
- 355 21. Waller K, Kujala UM, Kaprio J, Koskenvuo M, Rantanen T. Effect of Physical
356 Activity on Health in Twins: A 30-yr Longitudinal Study. *Med Sci Sports Exerc*.
357 2010;42(4):658-664. doi:10.1249/MSS.0b013e3181bdeea3
- 358 22. Kujala UM, Leskinen T, Rottensteiner M, et al. Physical activity and health: Findings
359 from Finnish monozygotic twin pairs discordant for physical activity. *Scand J Med Sci
360 Sports*. 2022;32(9):1316-1323. doi:10.1111/sms.14205
- 361 23. Goldberg J, Fischer M. Co-twin Control Methods. In: Everitt BS, Howell DC, eds.
362 *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Ltd; 2005:bsa143.
363 doi:10.1002/0470013192.bsa143
- 364 24. Kankaanpää A, Tolvanen A, Heikkinen A, Kaprio J, Ollikainen M, Sillanpää E. The
365 role of adolescent lifestyle habits in biological aging: A prospective twin study. *eLife*.
366 2022;11:e80729. doi:10.7554/eLife.80729
- 367 25. Bahls M, Baurecht H, Hanssen H, Van Craenenbroeck EM. How to establish
368 causality between physical inactivity and mortality? *Eur J Prev Cardiol*. 2022;29(8):e266-
369 e267. doi:10.1093/eurjpc/zwac021
- 370 26. Ballin M, Nordström P. Does exercise prevent major non-communicable diseases and
371 premature mortality? A critical review based on results from randomized controlled trials.
372 *J Intern Med*. 2021;290(6):1112-1129. doi:10.1111/joim.13353

- 373 27. Karvinen S, Waller K, Silvennoinen M, et al. Physical activity in adulthood: genes
374 and mortality. *Sci Rep*. 2015;5(1):18259. doi:10.1038/srep18259
- 375 28. Wang Z, Emmerich A, Pillon NJ, et al. Genome-wide association analyses of physical
376 activity and sedentary behavior provide insights into underlying mechanisms and roles in
377 disease prevention. *Nat Genet*. 2022;54(9):1332-1344. doi:10.1038/s41588-022-01165-1
- 378 29. Lin X, Zhang X, Guo J, et al. Effects of Exercise Training on Cardiorespiratory
379 Fitness and Biomarkers of Cardiometabolic Health: A Systematic Review and Meta-
380 Analysis of Randomized Controlled Trials. *J Am Heart Assoc*. 2015;4(7):e002014.
381 doi:10.1161/JAHA.115.002014
- 382 30. Valenzuela PL, Ruilope LM, Santos-Lozano A, et al. Exercise benefits in
383 cardiovascular diseases: from mechanisms to clinical implementation. *Eur Heart J*.
384 Published online April 3, 2023:ehad170. doi:10.1093/eurheartj/ehad170
- 385 31. Kujala UM. Is physical activity a cause of longevity? It is not as straightforward as
386 some would believe. A critical analysis. *Br J Sports Med*. 2018;52(14):914.
387 doi:10.1136/bjsports-2017-098639
- 388 32. Sillanpää E, Palviainen T, Ripatti S, Kujala UM, Kaprio J. Polygenic Score for
389 Physical Activity Is Associated with Multiple Common Diseases. *Med Sci Sports Exerc*.
390 2022;54(2):280-287. doi:10.1249/MSS.0000000000002788
- 391 33. Tynkkynen NP, Törmäkangas T, Palviainen T, et al. *Associations of Polygenic*
392 *Inheritance of Physical Activity with Aerobic Fitness, Cardiometabolic Risk Factors and*
393 *Diseases: The HUNT Study*. medRxiv; 2023. doi:10.1101/2023.03.24.23287686
- 394

395 **Tables and Figures**

396 Table 1. Subject characteristics according to the latest known information by survival status at the end
397 of follow-up.

N=4,897	Alive		Deceased	
	Mean (SD)		Mean (SD)	
	N	or %	N	or %
Characteristic				
Age (years)	3,702	44.8 (7.2)	1,195	51.3 (6.9)
Male (%)	1,540	41.6	699	58.5
Female (%)	2,162	58.4	496	41.5
Height (cm)	3,702	168.3 (8.8)	1,195	169.3 (8.8)
Weight (kg)	3,702	70.1 (13.1)	1,195	74.4 (13.4)
BMI (kg · m ⁻²)	3,702	24.6 (3.7)	1,195	25.9 (3.9)
Educational level^a				
Low (%)	1,223	33.0	577	48.3
Middle (%)	1,827	49.4	502	42.0
High (%)	652	17.6	116	9.7
Health status^a				
Not healthy (%)	484	13.1	268	22.4
Healthy (%)	3,218	86.9	927	77.6
Smoking				
Never (%)	1,830	49.4	475	39.8
Occasional (%)	114	3.1	35	2.9
Former (%)	925	25.0	316	26.4
Light (%)	212	5.7	54	4.5
Medium (%)	362	9.8	135	11.3
Heavy (%)	259	7.0	180	15.1
Fruit and vegetable consumption per day^a				
Not once (%)	237	6.4	120	10.0

Once or twice (%)	3,071	83.0	993	83.1
Three times or more (%)	394	10.6	82	6.9
Alcohol consumption				
Lifetime abstainer (%)	275	7.4	106	8.9
Former (%)	42	1.1	14	1.2
Occasional (%)	224	6.1	74	6.2
Low (%)	2,878	77.7	814	68.1
Medium (%)	199	5.4	115	9.6
High (%)	62	1.7	46	3.9
Very high (%)	22	0.6	26	2.2
Physical activity				
Mean PA 1975–1990 (MET h · week ⁻¹)	3,702	19.1 (15.9)	1,195	17.0 (14.1)
Meeting PA guidelines (avg MET h · week ⁻¹ 1975–1990)	2,983	80.6	893	74.7

398 Values are means and standard deviations for continues variables and proportions of all participants for others. Data from 1990 if not else
399 specified; a,Data from 1981; BMI, Body mass index; PA, Physical activity; MET, Metabolic equivalents; Meeting PA guidelines, average of
400 MET h · week⁻¹ during 1975–1990 \geq 7.5.

401 Table 2. Associations of polygenic risk scores for coronary heart disease and physical activity with all-cause mortality after step-by-step
 402 adjustments for other healthy lifestyle factors.

All-cause mortality	Model 1	Model 2	Model 3	Model 4	Model 5
Cases / N	1,366 / 5,415	1,355 / 5,381	1,294 / 5,236	1,292 / 5,230	1,195 / 4,897
PRS CHD (ZScore)	1.07 (0.99 to 1.14)	1.07 (0.99 to 1.14)	1.06 (0.98 to 1.14)	1.06 (0.98 to 1.14)	1.07 (0.99 to 1.15)
Meeting PA guidelines (avg MET h · week⁻¹ 1975–1990)	0.85 (0.72 to 0.99)	0.86 (0.73 to 1.01)	0.90 (0.76 to 1.06)	0.92 (0.78 to 1.08)	0.93 (0.78 to 1.11)
BMI (kg · m⁻²)		1.02 (1.00 to 1.05)	1.02 (1.00 to 1.05)	1.02 (1.00 to 1.04)	1.02 (0.99 to 1.04)
Smoking					
Never			ref.	ref.	ref.
Occasional			1.28 (0.83 to 1.97)	1.28 (0.83 to 1.97)	1.24 (0.78 to 1.97)
Former			1.38 (1.14 to 1.68)	1.38 (1.14 to 1.67)	1.38 (1.12 to 1.69)
Light			1.81 (1.24 to 2.65)	1.80 (1.23 to 2.64)	1.77 (1.18 to 2.65)
Medium			2.23 (1.75 to 2.85)	2.23 (1.74 to 2.84)	2.16 (1.66 to 2.81)
Heavy			4.01 (3.14 to 5.13)	3.95 (3.09 to 5.05)	3.54 (2.69 to 4.64)
Fruit and vegetable consumption per day^a					
Not once				ref.	ref.
Once or twice				0.81 (0.65 to 1.01)	0.86 (0.68 to 1.10)
Three times or more				0.76 (0.54 to 1.09)	0.81 (0.56 to 1.17)
Alcohol consumption					

Lifetime abstainer	ref
Former	1.00 (0.51 to 1.96)
Occasional	0.85 (0.58 to 1.26)
Low	0.93 (0.70 to 1.24)
Medium	1.30 (0.88 to 1.93)
High	1.69 (0.99 to 2.86)
Very high	2.16 (1.22 to 3.82)

403 Values are Hazard Ratios with 99% Confidence intervals; All models additionally adjusted for ten principal components, age, sex, and educational attainment, health status and clustered based on the twin structure of
404 the data. PRS, polygenic risk score, CHD, coronary heart disease; ZScore, Standardized score; Meeting PA guidelines, the average of self-reported physical activity from 1975–1990 ≥ 7.5 metabolic equivalent hours
405 per week (MET h · week⁻¹); Statistically significant associations (P<0.01) highlighted with **bold**.

406 Table 3. Associations of polygenic risk scores for coronary heart disease and physical activity with cardiovascular mortality after step-by-step
 407 adjustments for other healthy lifestyle factors.

Cardiovascular mortality	Model 1	Model 2	Model 3	Model 4	Model 5
Cases / N	449 / 5,415	447 / 5,381	426 / 5,236	424 / 5,230	389 / 4,897
PRS CHD (ZScore)	1.17 (1.04 to 1.32)	1.16 (1.03 to 1.31)	1.15 (1.02 to 1.30)	1.15 (1.02 to 1.30)	1.17 (1.03 to 1.33)
Meeting PA guidelines (avg MET h · week⁻¹ 1975–1990)	0.78 (0.59 to 1.04)	0.83 (0.63 to 1.10)	0.84 (0.63 to 1.12)	0.87 (0.65 to 1.17)	0.92 (0.67 to 1.25)
BMI (kg · m⁻²)		1.06 (1.03 to 1.10)	1.06 (1.02 to 1.10)	1.06 (1.02 to 1.10)	1.06 (1.02 to 1.10)
Smoking					
Never			ref.	ref.	ref.
Occasional			1.32 (0.67 to 2.62)	1.33 (0.67 to 2.65)	1.31 (0.66 to 2.62)
Former			1.36 (0.98 to 1.88)	1.36 (0.98 to 1.89)	1.41 (0.99 to 2.00)
Light			1.47 (0.72 to 3.00)	1.45 (0.71 to 2.98)	1.26 (0.55 to 2.89)
Medium			2.12 (1.31 to 3.42)	2.13 (1.33 to 3.43)	2.12 (1.27 to 3.55)
Heavy			3.84 (2.39 to 6.17)	3.74 (2.33 to 6.02)	3.50 (2.06 to 5.95)
Fruit and vegetable consumption per day^a					
Not once				ref.	ref.
Once or twice				0.72 (0.49 to 1.07)	0.72 (0.48 to 1.08)
Three times or more				0.55 (0.27 to 1.09)	0.52 (0.25 to 1.07)
Alcohol consumption					

Lifetime abstainer	ref.
Former	0.94 (0.37 to 2.39)
Occasional	0.68 (0.35 to 1.35)
Low	0.80 (0.50 to 1.28)
Medium	1.11 (0.55 to 2.25)
High	1.12 (0.43 to 2.91)
Very high	1.42 (0.40 to 5.08)

408 Values are Hazard Ratios with 99% Confidence intervals; All models additionally adjusted for ten principal components, age, sex, and educational attainment, health status and clustered based on the twin structure of
409 the data. PRS, polygenic risk score, CHD, coronary heart disease; ZScore, Standardized score; Meeting PA guidelines, the average of self-reported physical activity from 1975–1990 \geq 7.5 metabolic equivalent hours
410 per week (MET h · week⁻¹); Statistically significant associations (P<0.01) highlighted with **bold**.

Descriptives	Twins not meeting the PA guidelines	Twins meeting the PA guidelines	<i>P</i>
N	180	180	
BMI	25.3 (3.9)	25.1 (3.8)	0.350
Educational attainment ^a			
Low	73 (40.6 %)	68 (37.8 %)	ref.
Middle	82 (45.6 %)	83 (46.1 %)	0.546
High	23 (12.8 %)	22 (12.2 %)	0.869
Healthy ^a	149 (82.8 %)	153 (85.0 %)	0.557
Current smoker	43 (23.9 %)	46 (25.6 %)	0.655
Healthy diet ^a	0 (0 %)	0 (0 %)	NA
High alcohol consumption	0 (0 %)	3 (1.7 %)	NA
Mean MET·h week ⁻¹ 1975–1990	4.9 (1.7)	15.8 (7.8)	<0.001

411
 412 Figure 1. Descriptives of the monozygotic twin pairs, in which one twin meets the PA guidelines and
 413 the other twin does not. Values are means and standard deviations or proportions (%). Meeting PA
 414 guidelines, the average of self-reported physical activity from 1975–1990 ≥ 7.5 metabolic equivalent
 415 hours per week (MET h · week⁻¹); Data from 1990 if not else specified; a, Data from 1981; BMI,
 416 Body mass index; *P*, statistical difference between the twins who meet the PA guidelines or not; NA,
 417 Not applicable due low number of twins. Statistically significant associations ($P < 0.01$) highlighted
 418 with **bold**.

419
 420 Table 5. Within-pair risk of mortality in more active twins compared to their less active twin for 360
 421 twins from 180 discordant pairs

Risk of mortality	Cases/N	Twin meeting PA guidelines
		HR (99 % CI)
All-cause mortality		
Among all twin-pairs [#]	87/360	0.74 (0.37 to 1.48)
By genetic liability for CHD		
High	36/120	0.69 (0.23 to 2.11)
Intermediate	22/120	0.55 (0.15 to 2.01)
Low	29/120	1.00 (0.30 to 3.37)
Cardiovascular mortality		
Among all twin-pairs [#]	36/360	0.39 (0.10 to 1.46)

By genetic liability for CHD

High	16/120	0.43 (0.07 to 2.54)
Intermediate	10/120	0.50 (0.05 to 4.65)
Low	10/120	0.25 (0.01 to 4.45)

422 In a co-twin control design with monozygotic twins, twin pairs are controlled for their genetic factors (genome), and early life
423 environmental factors (intrauterine and childhood). #, Models are additionally adjusted for BMI and smoking status. CHD, coronary heart
424 disease; Meeting PA guidelines, the average of self-reported physical activity from 1975–1990 ≥ 7.5 metabolic equivalent hours per week.
425 Statistically significant associations ($P < 0.01$) highlighted with **bold**.

