

## **Do epigenetic clocks provide explanations for sex differences in lifespan? A cross-sectional twin study**

Anna Kankaanpää<sup>1</sup>, Asko Tolvanen<sup>2</sup>, Pirkko Saikkonen<sup>1</sup>, Aino Heikkinen<sup>3</sup>, Eija K. Laakkonen<sup>1</sup>,  
Jaakko Kaprio<sup>3</sup>, Miina Ollikainen<sup>3,4</sup>, Elina Sillanpää<sup>1,3</sup>

<sup>1</sup>Gerontology Research Center (GEREC), Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

<sup>2</sup>Methodology Centre for Human Sciences, University of Jyväskylä, Jyväskylä, Finland

<sup>3</sup>Institute for Molecular Medicine Finland (FIMM), HiLife, University of Helsinki, Helsinki, Finland

<sup>4</sup>Department of Public Health, University of Helsinki, Helsinki, Finland

Correspondence to:

Anna Kankaanpää, MSc, Doctoral Researcher

Gerontology Research Center, Faculty of Sport and Health Sciences

University of Jyväskylä, PO Box 35 (VIV)

FI-40014 Jyväskylä, Finland

E-mail: [anna.k.kankaanpaa@jyu.fi](mailto:anna.k.kankaanpaa@jyu.fi)

Tel: +358 40 589 5977

## **ABSTRACT**

### **Background**

The sex gap in life expectancy has been narrowing in Finland over the past four to five decades; however, on average, women still live longer than men. Epigenetic clocks are markers for biological aging that predict lifespan. In this study, we examined the mediating role of lifestyle factors on the association between sex and biological aging in younger and older adults.

### **Methods**

Our sample included same-sex younger and older twins (21-42-y, n = 1110; 50-76-y, n = 763) and younger opposite-sex twins (21-30-y, n = 302). Blood-based DNA methylation (DNAm) was used to compute epigenetic age acceleration by four epigenetic clocks as a measure of biological aging. Path modelling was used to study whether the association between sex and biological aging is mediated through lifestyle-related factors, i.e. education, body mass index, smoking, alcohol use, and physical activity.

### **Findings**

In comparison to women, men were biologically older and, in general, they had unhealthier life habits. The effect of sex on biological aging was partly mediated by smoking, but only in older twins. Sex was directly associated with biological aging and the association was stronger in older twins.

### **Interpretation**

Previously reported sex differences in lifespan are also evident in biological aging. Declining smoking prevalence among men is a plausible explanation for the narrowing of the difference in life expectancy between the sexes. Data generated by the epigenetic clocks may help in estimating the effects of lifestyle and environmental factors on aging and in predicting aging in future generations.

**Key words:** DNA methylation, sex gap, lifestyle, lifespan, biological age, twin study

## INTRODUCTION

Both sexes have experienced tremendous increases in life expectancy over the 20<sup>th</sup> century. However, through all historical periods, women have had a longer life expectancy than men. The sex gap in life expectancy varies across time and country.<sup>1</sup> In Finland, the sex gap increased greatly in the first half of the 20<sup>th</sup> century. That gap was greatest in the mid-1970s (nine years); since then, it has narrowed to 5.4 years.<sup>2</sup>

It has been suggested that sex differences in lifespan are caused by a complex combination of biological (genetic, hormonal) and non-biological (behavioural, economic, social, environmental, and cultural) factors.<sup>3</sup> Investigating sex differences in cause-specific mortality increases the understanding of the mechanisms underlying the sex differences in overall mortality. In comparison to women, men experience a higher risk of death from almost all causes.<sup>4</sup> External causes of deaths, such as traffic accidents, trauma, alcohol intoxication, illicit drug overdoses, and suicides, are more common among men. However, at most, these factors typically explain a modest fraction of all premature deaths. The majority of premature deaths are caused by non-communicable diseases (for example, cardiometabolic diseases, lung diseases, cancers, mental disorders, and dementia).<sup>4</sup> The biological and behavioural factors predisposing an individual to these diseases are predominantly the most important drivers of male-to-female differences in mortality.

Overweight and obesity are dramatically increasing worldwide, predisposing both men and women to several non-communicable diseases.<sup>5</sup> However, data regarding how obesity trends affect the sex gap in life expectancy are limited. Of the health-hazardous behavioural factors, tobacco smoking has been seen as the predominant driver of both the trend and the extent of sex differences in life expectancy. A recent study suggested that increasing smoking-related mortality among women and decreasing smoking-related mortality among men may account for as much as 40% of the narrowing sex gap in life expectancy over the last two decades.<sup>6</sup> In general, men consume more alcohol than women. In Finland, the risk for alcohol-related death is three-times higher in men than women.<sup>2</sup> Globally, men tend to be more physically active than women at all ages,<sup>7</sup> and leisure-time physical activity is known to be associated with a lower risk of premature death.<sup>8</sup> Therefore, leisure-time physical activity is expected to diminish the sex gap in life expectancy. Socioeconomic factors might also impact the sex gap, such as differences in education, income, and physical demands of work between the sexes.

For example, the sex gap is probably diminishing because many deaths related to trauma and toxication that previously occurred in male-dominated occupations are much rarer nowadays.<sup>5</sup>

In addition to societal factors, differences in innate biology may also have a role in the survival gap between the sexes. Genetic and physiological differences between the sexes include progressive skewing of X chromosome inactivation, telomere attrition, maternally inherited mitochondrial inheritance and hormonal and cellular differences in inflammatory and immunological responses and in substrate metabolism.<sup>9</sup> The biological longevity advantage of women may also result from oestrogen-associated greater resistance to oxidative damage.<sup>9</sup> However, in women, sex hormone levels change drastically during menopause, potentially contributing to the reduction of age-related sex differences in health outcomes such as cardiovascular risk factors.

Life expectancy may not always be a reliable proxy for how fast the population is aging, as it is the most distal outcome of ageing processes. To better monitor population health, more sensitive methods are needed to track changes in ageing. Novel biological clocks, i.e. epigenetic clocks, may help track and understand the individual aging process and offer insights into sex differences in biological age and how lifestyle may counteract the aging process.<sup>10-13</sup> These composite measures have been developed to quantify an individual's biological age, and they may enable accurate estimation of the pace of aging in all age groups. The first published results on biological age determined by epigenetic clocks have shown that men tend to be biologically older than women.<sup>14-17</sup>

This study aimed to examine sex differences in biological age measured by novel epigenetic clocks in age groups under and over 50 years, with 50 being a proxy for menopausal age.<sup>18</sup> Moreover, we aimed to assess whether the potential difference in biological aging between the sexes is mediated by different lifestyle factors, and whether age modifies these associations.

## **METHODS**

### **Study population**

The Finnish Twin Cohort (FTC) includes three large cohort studies: 1) The older FTC includes twins born before 1958, 2) Finntwin16 includes twins born in 1975-1979, and 3) Finntwin12 includes twins born in 1983-1987.<sup>19-21</sup> The twins selected from all three cohorts

were invited to participate in clinical in-person studies. Data on lifestyle- and health-related traits were collected using questionnaires and interviews. After written informed consent was obtained, blood samples for DNA analyses were collected. Twins (age range from 21 to 76) who had participated in clinical in-person studies with sampling for whole blood DNA and subsequent DNA methylation (DNAm) analyses and who had the relevant phenotype data were included in the current study. Further information on data collection is provided in the appendix (p 1). The analysis sample included monozygotic (MZ) and dizygotic (DZ) same-sex twins (N = 1893, 54% MZ) as well as opposite-sex twin pairs (302 twin individuals).

The FTC data collections were approved by the ethics committees of the University of Helsinki (113/E3/01 and 346/E0/05) and Helsinki University Central Hospital (270/13/03/01/2008 and 154/13/03/00/2011).

## **Main variables**

### **DNA methylation and assessment of biological age**

DNAm profiles were obtained using Illumina's Infinium HumanMethylation450 BeadChip or the Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA). DNAm-based epigenetic age estimates, obtained by Horvath's<sup>10</sup> and Hannum's<sup>11</sup> clocks and by PhenoAge<sup>12</sup> and GrimAge<sup>13</sup> estimators, were calculated using a publicly available online calculator (<https://dnamage.genetics.ucla.edu/new>). The age acceleration (AA) of each clock was defined as the residual from regressing the estimated biological age on chronological age (AA<sub>Horvath</sub>, AA<sub>Hannum</sub>, AA<sub>Pheno</sub>, and AA<sub>Grim</sub>, respectively). A more detailed description of the pre-processing and normalising of the DNAm data, as well as additional information on the utilised clocks, is provided in the appendix (pp 1-2).

### **Potential mediating variables**

We surmised that differences in the covariates between the sex groups are more likely the factors that underlie the sex differences rather than being confounders. The potential mediators included body mass index (BMI), smoking, alcohol use, physical activity, and educational attainment, which is a key component of socio-economic status.

*Educational attainment* was assessed as the number of years of full-time education.

*Body mass index* (BMI), measured as kg/m<sup>2</sup>, can be used as an estimate of healthy diet and sufficient energy intake. A high BMI describes excess fat in the body; thus, it is a

consequence of a long-term imbalance between energy intake and expenditure. We measured height in cm using a stadiometer and body mass in kg using a beam scale in kg.

*Smoking* was self-reported and classified as never, former, and current smoker.

*Alcohol use* was measured based on self-reported quantity and frequency of use and the content of the alcohol beverages. These data were transformed into 100% alcohol grams per day.

*Physical activity* was assessed using the Baecke Questionnaire.<sup>22</sup> The questionnaire has three sections: sports participation, leisure-time physical activity excluding sports, and work- or school-related physical activity. The questionnaire includes four questions on sports activity and leisure-time activity, excluding sports, and eight questions on occupational physical load scored on a five-point scale. A sport index, a non-sport leisure-time (leisure) index, and a work index, respectively, were based on the mean scores of each section as described by Baecke et al.<sup>22</sup> and Mustelin et al. for the FinnTwin12 study.<sup>23</sup>

### **Statistical analysis**

To compare differences in the study variables between men and women, we used regression analysis for the continuous variables and the Chi-square test for the categorical variables. In the regression models, the within-pair dependency of twin individuals was taken into account using the cluster option in the analysis.

The shape of the association between age and AA was studied using polynomial models of age as the continuous variable. To study whether sex differences in AA varied by age, the interaction effects of sex and age were also included in the regression models.

Mediation models were used to test whether the association between sex and AA is direct or mediated through lifestyle factors in the same-sex twins and opposite-sex twin pairs. First, the single mediation models were fitted. These models included indirect paths from sex to AA through one lifestyle factor at a time as well as the direct effect of sex on AA. In the same-sex twins, we further studied whether these associations differed according to age group, that is, whether age moderated the associations. The single mediation models included the interaction effect of sex and age group on the mediator variable and directly on AA. Furthermore, the interaction effect between the mediator variable and age group on AA was tested for significance. Second, a multiple mediation model was fitted to assess the mediation effect of the different lifestyle factors simultaneously. The mediators as well as the

interactions were included in the final multiple mediation model based on the results of the single mediator models.

The standard errors were corrected for nested sampling using the special option in Mplus (TYPE=COMPLEX). The models for the opposite-sex twin pairs were fitted using multilevel modelling, and the mediation models were specified at the within-twin pairs level. The approach controls for shared childhood environmental factors and partly for genetic factors.

The indirect effects were calculated as the product of the regression coefficients. For the same-sex twins, the age-specific indirect effects were calculated using the parameters of the model. The parameters of the models were estimated using the full information maximum likelihood (FIML) method with robust standard errors. For the models including the ordinal mediator variable smoking status, the estimation was conducted using a robust weighted least squares (WLSMV) estimator. In that case, the mediator is assumed to be the continuous latent variable underlying ordinal smoking status. Descriptive statistics and differences in the study variables were calculated and tested using Stata 16 software (StataCorp, Inc. College Station, TX, USA), and further modelling was conducted with the Mplus statistical package (version 8.2).<sup>24</sup>

## RESULTS

### **Sex differences in lifestyle factors and epigenetic aging**

The characteristics of the younger and older same-sex twins and the opposite-sex twin pairs included in this study are presented in Table 1. In both age groups of the same-sex twins, there were fewer men than women. Among the older twins, the men were younger and better educated than the women. The men belonging to an opposite-sex twin pair had a lower level of education in comparison to their twin sisters. The men had a higher BMI in young adulthood than the women. Among the same-sex twins, there were more current smokers among the men in comparison to the women; there was no sex difference in smoking among the opposite-sex twin pairs. In all the groups, the men consumed more alcohol than the women. Moreover, the men had a lower level of leisure index in all the groups in comparison to the women, but the men in the younger same-sex twin group had a higher level of sport index than women.

Overall, the men had higher AA than the women, and the sex difference in AA tended to increase with age in the same-sex twins (Table 1 and Figure 1). Interestingly, when DNAm PhenoAge was used to assess AA, the men were epigenetically younger than the women in the younger age, which was in contrast with the AA estimates derived from the other clocks. When controlling for shared childhood environment and partly for genetic factors among the opposite-sex twin pairs, the male twins had higher  $AA_{\text{Hannum}}$  and  $AA_{\text{Grim}}$  in comparison to their sisters, but there were no significant differences in  $AA_{\text{Horvath}}$  and  $AA_{\text{Pheno}}$  between the sexes (Table 1). A more detailed description of the sex differences in the epigenetic aging is provided in the appendix (pp 2-8).

### **Mediation models in the twins from the same-sex pairs**

Education, BMI, smoking, alcohol use, sport index, leisure index and work index were considered to be the potential mediator variables. Here, smoking was assumed to be a continuous latent variable underlying ordinal smoking status. The estimation results of the single mediator models revealed that the association of sex with education and sport index differed between the age groups (appendix, pp 10-11, mediator regressed on sex  $\times$  age). There were also some differences in the associations of the potential mediator variables (including education, smoking and alcohol use) with AA between the age groups. The results of the associations between lifestyle-related factors and AA are presented in the appendix (pp 9-11).

The estimated age-specific indirect associations of male sex on AA through the potential mediator variables are shown in Table 2. Male sex was associated with higher  $AA_{\text{Horvath}}$ ,  $AA_{\text{Pheno}}$  and  $AA_{\text{Grim}}$  through higher BMI, but only in the younger twins. Smoking partly mediated the association of male sex with higher  $AA_{\text{Hannum}}$ ,  $AA_{\text{Pheno}}$  and  $AA_{\text{Grim}}$  in the older twins. Greater alcohol use partly mediated the association of male sex with higher  $AA_{\text{Pheno}}$  in the older twins and with  $AA_{\text{Grim}}$  in both age groups. Moreover, in younger twins, male sex was associated with lower  $AA_{\text{Grim}}$  through a higher sport index. Because education, leisure index and work index did not mediate the sex difference in AA, these variables were dropped from further modelling.

Based on the results of the single mediator models, the final multiple mediator model included BMI, smoking, alcohol use, and sport index as the mediators (Figure 2). The model also included the interaction effect between sex and age on smoking and sport index, the



interaction effect between smoking and age on AA, and the interaction effect between alcohol use and age on AA<sub>Pheno</sub>. When the lifestyle factors were controlled for each other, smoking partly mediated the association between sex and AA (AA<sub>Hannum</sub>, AA<sub>Pheno</sub>, and AA<sub>Grim</sub>) but only in the older twins (Table 2). Alcohol use partly mediated the sex difference in AA<sub>Pheno</sub> in the older twins. Male sex was still found to have a direct positive effect on AA<sub>Horvath</sub>, AA<sub>Hannum</sub>, and AA<sub>Grim</sub> after all the adjustments, and the association was stronger in the older cohort (Figure 2). Moreover, male sex was found to have a positive direct effect on AA<sub>Pheno</sub>, but only in the older cohort.

### **Mediation models in the opposite-sex twin pairs**

Information on the association between the lifestyle-related factors and AA is provided in appendix (pp 12-13). Based on the estimation results of the single mediator models, male sex was associated with accelerated AA<sub>Horvath</sub> through higher BMI in the opposite-sex twin pairs (Table 3). Otherwise, there were no significant indirect effects. Similar to the models for the same-sex twins, BMI, smoking, alcohol use, and sport index were included in the multiple mediator model as the mediator variables (Figure 3). A significant indirect association of male sex on AA<sub>Horvath</sub> through higher BMI was also observed after controlling for other lifestyle factors (Table 3). Otherwise, lifestyle factors did not mediate the differences in AA between the men and their female twin sisters. A direct effect of male sex on higher AA<sub>Hannum</sub> and AA<sub>Grim</sub> was also observed among the opposite-sex twin pairs (Figure 3).

## **DISCUSSION**

Our findings suggest that previously reported sex differences in life expectancy can be seen in biological aging when measured with epigenetic clocks (namely Horvath's clock, Hannum's clock, DNAm PhenoAge, and GrimAge). Sex difference was already evident in young adulthood and it increased with age; on average, the men were 1.2 to 1.5 years older than women in the younger same-sex twins (21 to 42 years of age) and 3.2 to 4.3 years older in the older same-sex twins (50 to 76 years of age). The only exception was observed in the younger same-sex twins when PhenoAge was used to assess biological aging; the men were epigenetically 1.3 years younger than the women, but the sex difference in biological age reversed in the older adult twins. According to previous studies, sex difference in biological

aging measured with epigenetic clocks seems to appear in adolescence,<sup>25</sup> and men are epigenetically one-to-two years older than women in adulthood.<sup>26</sup>

Opposite-sex twins provide a natural setting for studying sex differences while maximally controlling for genetic factors and shared childhood environmental factors. Epigenetic aging is highly heritable.<sup>27</sup> Although the share of genes is 50% in male-female dizygotic twins, the mean difference by sex of epigenetic aging within these twin pairs (21 to 30 years of age) was comparable to the sex differences observed in the larger cohort of the same-sex twins.

The observed increase in sex difference with age was mainly due to the fact that epigenetic aging accelerated with age among the men. Changes in the hormonal levels during menopause have a detrimental effect on women's health,<sup>28</sup> thus, sex differences in biological aging would be expected to diminish around and after the age of 50. Our analysis studying the shape of the association between chronological age and epigenetic aging did not find any evidence that this is the case in biological aging. This is in line with a recently published study investigating sex differences in the longitudinal trajectories of epigenetic aging from midlife onwards (50 to 90 years of age).<sup>16</sup> According to the study, men were biologically older than women when Horvath's and Hannum's epigenetic clocks and GrimAge were used, and the difference remained constant across the age span.<sup>16</sup>

To the best of our knowledge, this is the first study that has tested the mechanisms underlying sex differences in biological aging measured with epigenetic clocks. We found that several lifestyle-related factors partly mediated the association of sex with biological aging in the same-sex twins, when the mediation of these factors was assessed one at a time. However, after controlling for each of the health-related behaviours in the multiple mediator models, only smoking consistently mediated the sex difference in the older twins. Smoking was associated with accelerated biological aging, and the association was stronger in the older twins; this suggests the cumulative effect of smoking on biological aging. Moreover, the sex difference in smoking behaviour was larger in the older twins; in fact, the difference in the proportion of never-smokers between men and women was wider in the older twins (45% vs 73%) in comparison to the younger twins (45% vs 53%). A previous population-based study investigating long-term trends in smoking in Finland has shown that the prevalence of daily smoking has steadily decreased among men since the late 1970s (37% to 17%); in contrast, the current prevalence among women is about the same as it was four decades ago (~15%).<sup>29</sup>

Together, these findings support recent studies suggesting that the narrowing of the sex differences in smoking probably partly explains the declining sex gap in lifespan.<sup>6,30</sup>

We observed some differences in the associations between lifestyle-related factors and epigenetic aging across the utilised clocks. These inconsistencies are probably due to differences in the procedures used to develop these epigenetic age estimators. The first-generation clocks, namely Horvath's clock and Hannum's clock, were trained to predict chronological age. More novel estimators are supposed to also capture CpG sites whose DNAm levels correlate with the deviation of biological age from chronological age.

Of the epigenetic age estimators employed in our study, DNAm GrimAge is the one that has been most recently published, and it outperforms other estimators in terms of predicting mortality.<sup>13,16</sup> It utilises information on chronological age, sex and seven DNAm-based surrogates for seven plasma proteins and for smoking pack-years. Although sex difference in GrimAge is in-built, reflecting differences in mortality, the observed sex differences were very similar to the corresponding ones measured with Hannum's clock, which is purely based on CpG sites with their DNAm levels correlating with chronological age. To further understand the sex differences in biological aging, we studied the DNAm-based surrogates included in the GrimAge estimator (see appendix pp 2, 4-8). We observed a significantly higher level of DNAm-based plasminogen activator inhibitor 1 (PAI-1) among men in comparison to women. Moreover, in the men, the level of DNAm PAI-1 drastically increased with age. This DNAm-based surrogate predicts morbidity better than DNAm GrimAge, and it associates with hypertension, type 2 diabetes, and coronary heart disease.<sup>13</sup> Therefore, higher levels of DNAm PAI-1 in men may play a role in the sex differences in cardiovascular mortality observed in previous studies.<sup>4</sup>

Our study has several strengths. We utilised recently published epigenetic clocks that are shown to predict mortality.<sup>16</sup> The large sample size of our study enabled us to use complex mediator models. Because data from opposite-sex twin pairs were available, we were also able to control the analyses for shared childhood environmental factors and partly for genetic factors. This study also had some limitations. Most of the studied lifestyle-related factors were self-reported. Furthermore, our data were cross-sectional, and our analysis did not rule out the possibility of reversed causality when studying the associations between lifestyle-related factors and epigenetic aging.

Our results deepen the understanding of the association between sex-dependent lifestyle factors and the aging process. Progress in this area may result in methods that are able to determine individual trajectories in aging that already occur in early adulthood. These methods would enable researchers to investigate the effects of environmental and societal changes and lifestyle interventions on biological aging.

### **Contributors**

ES conceived the idea for the study. AH and AK pre-processed the DNAm data. AK, ES, and AH accessed and verified the dataset. AK, AT, and ES designed the statistical analysis, and AK and ES performed the analyses. JK and MO have designed and collected the FTC dataset and participated in designing the analysis of this manuscript. AK and ES drafted the first version of the manuscript. All authors contributed to the interpretation of results and drafting or revising the manuscript.

### **Declaration of interests**

The authors declare that there are no conflicts of interest.

### **Data sharing**

The dataset used in the current study will be located in the Biobank of the National Institute for Health and Welfare, Finland. All the biobanked data are publicly available for use by qualified researchers following a standardised application procedure.

### **Acknowledgements**

This work was supported by the Academy of Finland (213506, 265240, 263278, 312073 to JK, and 297908 to MO), EC FP5 GenomEUtwin (JK), National Institutes of Health/National Heart, Lung, and Blood Institute (grant HL104125), EC MC ITN Project EPITRAIN (JK and MO), the University of Helsinki Research Funds (MO), Sigrid Juselius Foundation (JK and MO), Yrjö Jahnsson Foundation (6868), and Juho Vainio Foundation (ES).

The authors thank all the FTC study participants and research team members. The Gerontology Research Center is a joint effort between the University of Jyväskylä and the University of Tampere.

## Supplementary Material

### Supplementary appendix

#### REFERENCES:

- 1 Dicker D, Nguyen G, Abate D, *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; **392**: 1684–735.
- 2 Official Statistics of Finland (OSF): Causes of death. ISSN=1799-5078. [http://www.stat.fi/til/ksyyt/index\\_en.html](http://www.stat.fi/til/ksyyt/index_en.html) (accessed Jan 23, 2020).
- 3 Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, *et al.* Sex and gender: modifiers of health, disease, and medicine. *Lancet* 2020; **396**: 565–82.
- 4 Roth GA, Abate D, Abate KH, *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; **392**: 1736–88.
- 5 Stanaway JD, Afshin A, Gakidou E, *et al.* Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study. *Lancet* 2018; **392**: 1923–94.
- 6 Östergren O, Martikainen P. The contribution of smoking-related deaths to the gender gap in life expectancy in Sweden between 1997 and 2016. *Scand J Public Health* 2020; **48**: 346–9.
- 7 Hallal PC, Andersen LB, Bull FC, *et al.* Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet* 2012; **380**: 247–57.
- 8 Arem H, Moore SC, Patel A, *et al.* Leisure time physical activity and mortality: A detailed pooled analysis of the dose-response relationship. *JAMA Intern Med* 2015; **175**: 959–67.
- 9 Seifarth JE, McGowan CL, Milne KJ. Sex and life expectancy. *Gend Med* 2012; **9**: 390–401.
- 10 Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol* 2013; **14**: R115.

- 11 Hannum G, Guinney J, Zhao L, *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol Cell* 2013; **49**: 359–67.
- 12 Levine ME, Lu AT, Quach A, *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 2018; **10**: 573–91.
- 13 Lu AT, Quach A, Wilson JG, *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)* 2019; **11**: 303–27.
- 14 Jylhävä J, Pedersen NL, Hägg S. Biological Age Predictors. *EBioMedicine* 2017; **21**: 29–36.
- 15 Simpkin AJ, Hemani G, Suderman M, *et al.* Prenatal and early life influences on epigenetic age in children: A study of mother-offspring pairs from two cohort studies. *Hum Mol Genet* 2016; **25**: 191–201.
- 16 Li X, Ploner A, Wang Y, *et al.* Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *Elife* 2020; **9**: 1–20.
- 17 Crimmins EM, Thyagarajan B, Levine ME, Weir DR, Faul J. Associations of Age, Sex, Race/Ethnicity and Education with 13 Epigenetic Clocks in a Nationally Representative US Sample: The Health and Retirement Study. *Journals Gerontol - Ser A Biol Sci Med Sci* 2021; glab016, doi: <http://dx.doi.org/10.1080/07853890.2020.1840620>.
- 18 Dratva J, Real FG, Schindler C, *et al.* Is age at menopause increasing across Europe? Results on age at menopause and determinants from two population-based studies. *Menopause* 2009; **16**: 385–94.
- 19 Kaidesoja M, Aaltonen S, Bogl LH, *et al.* FinnTwin16: A Longitudinal Study from Age 16 of a Population-Based Finnish Twin Cohort. *Twin Res Hum Genet* 2019; **22**: 530–9.
- 20 Rose RJ, Salvatore JE, Aaltonen S, *et al.* FinnTwin12 Cohort: An Updated Review. *Twin Res Hum Genet* 2019; **22**: 302–11.
- 21 Kaprio J, Bollepalli S, Buchwald J, *et al.* The older Finnish Twin Cohort: 45 years of follow-up. *Twin Res Hum Genet* 2019; **22**: 240–54.
- 22 Baecke JAH, Burema J, Fritjters JER. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr* 1982; **96**: 932–42.
- 23 Mustelin L, Joutsu J, Latvala A, Pietiläinen KH, Rissanen A, Kaprio J. Genetic influences on physical activity in young adults: A twin study. *Med Sci Sports Exerc* 2012; **44**: 1293–301.
- 24 Muthén LK, Muthén BO. Mplus User’s Guide. Eighth Ed. Los Angeles, CA. Available

from: <https://www.statmodel.com>.

- 25 Simpkin AJ, Howe LD, Tilling K, *et al*. The epigenetic clock and physical development during childhood and adolescence: longitudinal analysis from a UK birth cohort. *Int J Epidemiol* 2017; **46**: 549–58.
- 26 Horvath S, Gurven M, Levine ME, *et al*. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol* 2016; **17**: 0–22.
- 27 Kankaanpää A, Tolvanen A, Bollepalli S, *et al*. Leisure-Time and Occupational Physical Activity Associates Differently with Epigenetic Aging. *Med Sci Sports Exerc* 2020; doi:10.1249/MSS.0000000000002498. [Epub ahead of print]
- 28 Mauvais-Jarvis F, Clegg DJ, Hevener AL. The role of estrogens in control of energy balance and glucose homeostasis. *Endocr Rev* 2013; **34**: 309–38.
- 29 Ruokolainen O, Heloma A, Jousilahti P, *et al*. Thirty-eight-year trends of educational differences in smoking in Finland. *Int J Public Health* 2019; **64**: 853–60.
- 30 Janssen F. Changing contribution of smoking to the sex differences in life expectancy in Europe, 1950–2014. *Eur J Epidemiol* 2020; **35**: 835–41.

Table 1. Sex differences in lifestyle-related factors, DNA methylation age and age acceleration (AA) estimates according to age group in the same-sex twins and in opposite-sex twin pairs.

	Same-sex twins						Opposite-sex twin pairs					
	21 to 42 yr-old twins (N = 1130)				Over 50 yr-old twins (N = 763)				21 to 30 yr-old twin pairs (N =151)			
	women	men	sex difference		women	men	sex difference		women	men	sex difference	
		mean	<i>p</i>			mean	<i>p</i>			mean	<i>p</i>	
N	613 <sup>a</sup>	517 <sup>b</sup>			621 <sup>c</sup>	142 <sup>d</sup>			151 <sup>e</sup>	151 <sup>e</sup>		
Zygoty, mz/dz	349/264	274/243			322/299	86/56						
Education, years	16.6 (3.6)	16.6 (3.6)	0.0	0.943	9.6 (3.3)	12.3 (4.1)	2.6	<0.001	17.3 (0.3)	16.3 (0.3)	-1.0	0.003
Age, years	24.7 (3.8)	25.1 (3.5)	0.4	0.181	66.6 (4.7)	62.0 (3.8)	-4.5	<0.001	23.9 (2.2)			
DNAmAge, est. years												
Horvath	31.3 (5.8)	32.8 (5.1)	1.5	0.001	65.5 (6.1)	64.2 (5.4)	-1.2	0.057	31.1 (4.6)	31.7 (4.9)	0.6	0.076
Hannum	19.9 (4.7)	21.0 (4.4)	1.2	0.001	55.7 (5.9)	55.0 (5.4)	-0.7	0.311	19.1 (3.4)	20.6 (4.0)	1.4	<0.001
PhenoAge	16.0 (7.2)	15.1 (6.3)	-0.9	0.070	55.7 (7.7)	56.7 (7.2)	1.0	0.231	14.2 (5.4)	13.8 (5.8)	-0.4	0.459
GrimAge	26.6 (4.9)	28.0 (4.7)	1.3	<0.001	58.6 (5.0)	59.2 (5.8)	0.6	0.348	25.8 (3.1)	27.1 (3.7)	1.3	<0.001
DNAmAge acceleration												
AA <sub>Horvath</sub>	-0.5 (3.9)	0.7 (3.6)	1.2	<0.001	-0.8 (3.6)	3.3 (5.2)	4.1	<0.001	-0.1 (3.5)	0.6 (3.9)	0.6	0.076
AA <sub>Hannum</sub>	-0.5 (3.5)	0.3 (3.1)	0.8	0.001	-0.7 (4.7)	2.6 (4.2)	3.2	<0.001	-0.6 (3.0)	0.8 (3.6)	1.4	<0.001
AA <sub>Pheno</sub>	0.6 (5.9)	-0.7 (5.0)	-1.3	0.001	-1.2 (7.2)	4.3 (6.0)	5.5	<0.001	-0.7 (5.4)	-1.1 (5.7)	-0.4	0.458
AA <sub>Grim</sub>	-0.5 (3.6)	0.5 (3.5)	1.0	<0.001	-0.8 (3.6)	3.3 (5.2)	4.1	<0.001	-0.7 (2.9)	0.6 (3.4)	1.3	<0.001
<b>Lifestyle-related variables</b>												
BMI, kg/m <sup>2</sup>	23.6 (5.0)	24.4 (3.9)	0.8	0.016	27.6 (4.9)	28.0 (4.5)	0.4	0.408	22.5 (3.5)	23.9 (3.5)	1.4	<0.001
Smoking, n (%)												
never	322 (52.6)	232 (44.9)		0.019	480 (77.3)	63 (44.4)		<0.001	66 (43.7)	61 (40.4)		0.646
former	103 (16.8)	89 (17.2)			90 (14.5)	53 (37.3)			50 (33.1)	48 (31.8)		
current	187 (30.6)	196 (37.9)			51 (8.2)	26 (18.3)			35 (23.2)	42 (27.8)		
Alcohol, g/day	7.2 (9.4)	15.4 (18.4)	8.1	<0.001	4.1 (7.4)	11.3 (19.7)	7.2	<0.001	7.8 (9.3)	16.4 (19.9)	8.5	<0.001
Physical activity												
work index	2.8 (0.7)	2.7 (0.8)	0.0	0.488	2.3 (1.0)	2.5 (0.98)	0.2	0.161	2.7 (0.7)	2.8 (0.8)	0.1	0.379



sport index	2.9 (0.8)	3.1 (0.8)	0.2	0.004	3.1 (0.8)	3.0 (0.83)	-0.1	0.386	3.0 (0.8)	2.9 (0.8)	0.0	0.942
leisure index	3.0 (0.6)	2.9 (0.6)	-0.1	0.005	2.9 (0.6)	2.7 (0.61)	-0.2	0.026	3.0 (0.6)	2.8 (0.6)	-0.3	0.001

DNAmAge, DNA methylation age; BMI, body mass index.

Values are means and standard deviations (BMI, alcohol, physical activity indexes) or numbers and percentages (smoking).

Between sex difference (regression analysis adjusted with family relatedness or Pearson's Chi square is significant) when  $p < 0.050$ .

In physical activity indices N: <sup>a</sup> 426-430, <sup>b</sup> 336-339, <sup>c</sup> 177-184, <sup>d</sup> 139-141, and <sup>e</sup> 91-93.

Table 2. Standardised indirect effects of sex (male) on epigenetic age acceleration (AA) through the potential mediator variables in the same-sex twins.

	Mediator <sup>a, b, c</sup>													
	Education		Body mass index		Smoking		Alcohol use		Sport index		Leisure index		Work index	
	B (SE)	<i>p</i>	B (SE)	<i>p</i>	B (SE)	<i>p</i>	B (SE)	<i>p</i>	B (SE)	<i>p</i>	B (SE)	<i>p</i>	B (SE)	<i>p</i>
<b>Single mediator models</b>														
Indirect associations in the younger twins														
AA <sub>Horvath</sub>	0.000 (0.001)	0.977	<b>0.012 (0.005)</b>	<b>0.027</b>	-0.023 (0.021)	0.271	-0.009 (0.009)	0.304	0.004 (0.005)	0.463	-0.001 (0.005)	0.884	0.000 (0.001)	0.773
AA <sub>Hannum</sub>	0.000 (0.009)	0.978	0.002 (0.003)	0.513	0.002 (0.005)	0.693	0.003 (0.010)	0.722	0.003 (0.005)	0.502	-0.004 (0.005)	0.398	-0.001 (0.002)	0.596
AA <sub>Pheno</sub>	0.000 (0.002)	0.974	<b>0.012 (0.005)</b>	<b>0.024</b>	0.010 (0.008)	0.219	0.010 (0.007)	0.161	-0.008 (0.005)	0.165	-0.002 (0.005)	0.753	0.000 (0.001)	0.778
AA <sub>Grim</sub>	0.001 (0.009)	0.948	<b>0.009 (0.004)</b>	<b>0.038</b>	0.048 (0.028)	0.087	<b>0.046 (0.009)</b>	<b>&lt;0.001</b>	<b>-0.019 (0.008)</b>	<b>0.016</b>	0.008 (0.005)	0.117	-0.005 (0.008)	0.529
Indirect associations in the older twins														
AA <sub>Horvath</sub>	-0.002 (0.009)	0.801	-0.001 (0.008)	0.877	0.023 (0.028)	0.407	-0.002 (0.008)	0.819	0.000 (0.003)	0.917	0.012 (0.017)	0.486	0.005 (0.006)	0.439
AA <sub>Hannum</sub>	0.009 (0.010)	0.379	0.006 (0.008)	0.475	<b>0.028 (0.009)</b>	<b>0.001</b>	0.011 (0.008)	0.192	0.000 (0.004)	0.943	0.009 (0.021)	0.661	-0.003 (0.005)	0.553
AA <sub>Pheno</sub>	0.011 (0.010)	0.298	0.000 (0.008)	0.993	<b>0.046 (0.012)</b>	<b>&lt;0.001</b>	<b>0.030 (0.008)</b>	<b>&lt;0.001</b>	0.000 (0.000)	0.905	-0.006 (0.024)	0.814	0.000 (0.004)	0.912
AA <sub>Grim</sub>	-0.011 (0.010)	0.261	0.005 (0.008)	0.486	<b>0.130 (0.028)</b>	<b>&lt;0.001</b>	<b>0.049 (0.009)</b>	<b>&lt;0.001</b>	-0.002 (0.006)	0.722	0.032 (0.023)	0.155	0.001 (0.005)	0.790
<b>Multiple mediator models<sup>d</sup></b>														
Indirect associations in the younger twins														
AA <sub>Horvath</sub>	- <sup>e</sup>		0.008 (0.005)	0.087	-0.005 (0.005)	0.337	-0.001 (0.008)	0.927	0.008 (0.006)	0.195	- <sup>e</sup>		- <sup>e</sup>	
AA <sub>Hannum</sub>			0.003 (0.002)	0.187	0.003 (0.005)	0.632	0.003 (0.008)	0.724	0.008 (0.006)	0.202				
AA <sub>Pheno</sub>			0.008 (0.005)	0.096	0.010 (0.008)	0.217	0.000 (0.010)	0.996	0.004 (0.006)	0.489				
AA <sub>Grim</sub>			0.005 (0.003)	0.105	0.047 (0.026)	0.063	-0.002 (0.005)	0.752	-0.006 (0.004)	0.143				
Indirect associations in the older twins														
AA <sub>Horvath</sub>	- <sup>e</sup>		0.008 (0.005)	0.087	0.008 (0.007)	0.267	-0.001 (0.008)	0.927	0.003 (0.003)	0.317	- <sup>e</sup>		- <sup>e</sup>	
AA <sub>Hannum</sub>			0.003 (0.002)	0.187	<b>0.030 (0.010)</b>	<b>0.002</b>	0.003 (0.008)	0.724	0.003 (0.003)	0.310				
AA <sub>Pheno</sub>			0.008 (0.005)	0.096	<b>0.044 (0.011)</b>	<b>0.001</b>	<b>0.017 (0.007)</b>	<b>0.010</b>	0.001 (0.002)	0.529				
AA <sub>Grim</sub>			0.005 (0.005)	0.105	<b>0.130 (0.028)</b>	<b>&lt;0.001</b>	-0.002 (0.005)	0.777	-0.002 (0.002)	0.252				

---

B, standardised (STDYX) indirect effect; SE, standard error

<sup>a</sup> The model was controlled for zygosity.

<sup>b</sup> SEs were corrected for nested sampling.

<sup>c</sup> Significant regression coefficients at the level 0.05 are presented in bold.

<sup>d</sup> The indirect effects were estimated as equal in the younger and older twins whenever interaction terms were not needed.

<sup>e</sup> Based on the results of the single mediator models the corresponding mediator variable was dropped out from the final multiple mediator model.

Table 3. Standardised indirect effects of sex (male) on epigenetic age acceleration (AA) through the potential mediator variables in the opposite-sex twins at within-twin pair level.

<b>Mediator</b> <sup>a, b</sup>														
	<b>Education</b>		<b>Body mass index</b>		<b>Smoking</b>		<b>Alcohol use</b>		<b>Sport index</b>		<b>Leisure index</b>		<b>Work index</b>	
	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>	<b>B (SE)</b>	<i>p</i>
<b>Single mediator models</b>														
AA <sub>Horvath</sub>	-0.019 (0.016)	0.251	<b>0.054 (0.022)</b>	<b>0.015</b>	-0.003 (0.009)	0.770	-0.026 (0.021)	0.200	0.002 (0.002)	0.847	-0.001 (0.018)	0.957	-0.009 (0.010)	0.385
AA <sub>Hannum</sub>	-0.010 (0.014)	0.480	-0.012 (0.023)	0.605	-0.004 (0.008)	0.628	-0.032 (0.022)	0.159	0.000 (0.010)	0.985	-0.026 (0.020)	0.186	0.001 (0.005)	0.893
AA <sub>Pheno</sub>	-0.017 (0.012)	0.173	0.015 (0.020)	0.458	0.008 (0.011)	0.460	-0.018 (0.020)	0.366	0.001 (0.011)	0.959	-0.025 (0.016)	0.131	0.003 (0.006)	0.640
AA <sub>Grim</sub>	0.025 (0.015)	0.095	-0.006 (0.024)	0.804	0.029 (0.029)	0.307	0.009 (0.023)	0.689	0.000 (0.001)	0.939	0.022 (0.018)	0.221	0.006 (0.008)	0.411
<b>Multiple mediator models</b>														
AA <sub>Horvath</sub>	.. <sup>c</sup>		<b>0.056 (0.025)</b>	<b>0.026</b>	0.001 (0.009)	0.924	-0.033 (0.026)	0.195	0.001 (0.009)	0.941	.. <sup>c</sup>		.. <sup>c</sup>	
AA <sub>Hannum</sub>			-0.009 (0.016)	0.555	-0.004 (0.008)	0.659	-0.029 (0.021)	0.171	-0.001 (0.010)	0.941				
AA <sub>Pheno</sub>			0.027 (0.016)	0.095	0.011 (0.014)	0.414	-0.025 (0.025)	0.319	-0.001 (0.013)	0.941				
AA <sub>Grim</sub>			0.010 (0.017)	0.552	0.030 (0.030)	0.310	0.002 (0.017)	0.932	0.000 (0.001)	0.956				

B, standardised (STDYX) indirect effect; SE, standard error

<sup>a</sup>The model was controlled for age as twin pairs participated in measurements at slightly different ages ( $\pm 1$  year).

<sup>b</sup>Significant indirect effects at the level 0.05 are presented in bold.

<sup>c</sup>Based on the results of the single mediator models in the same-sex and the opposite-sex twin pairs, the corresponding mediator variable was dropped out from the final multiple mediator model.

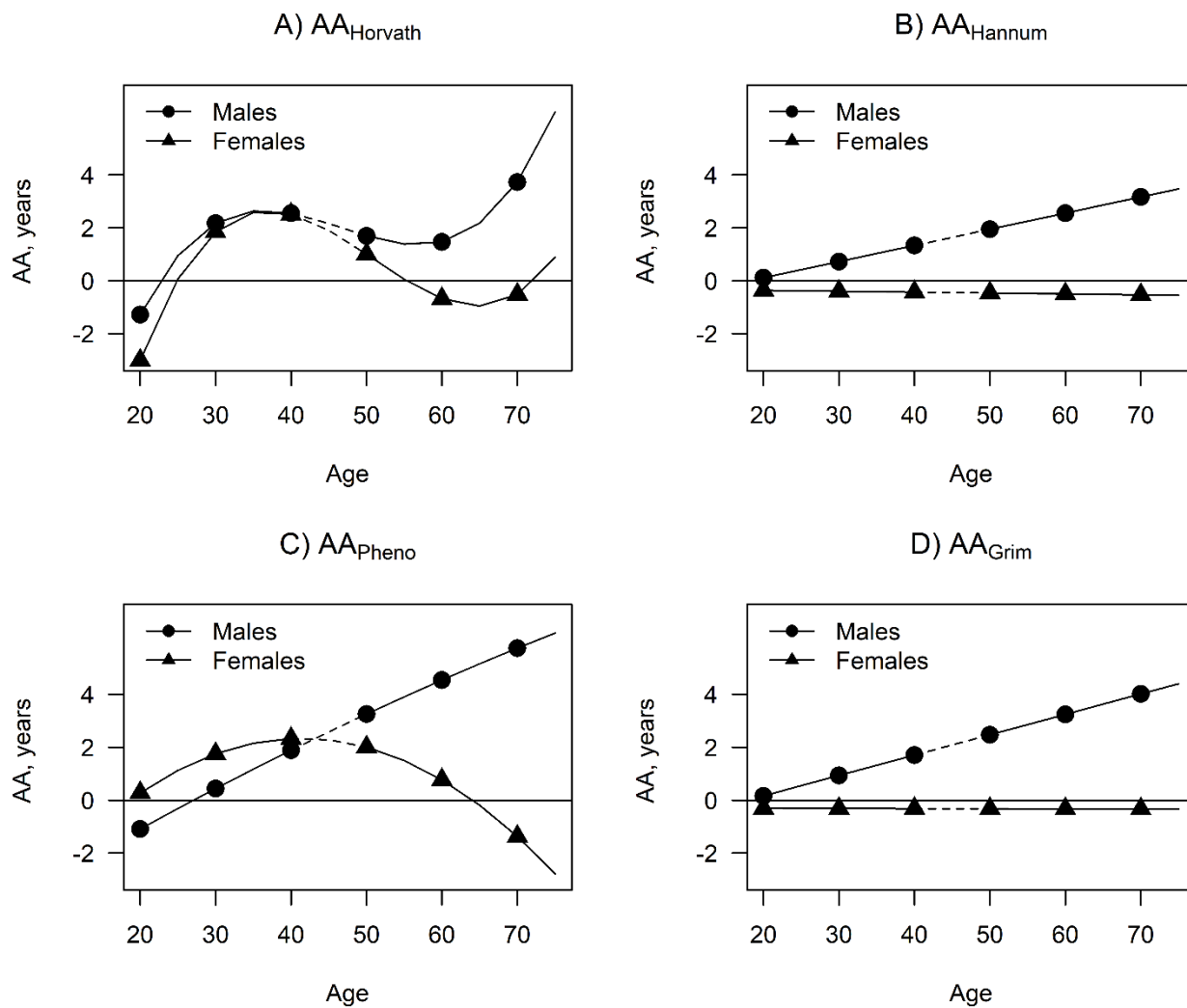


Figure 1. Epigenetic age acceleration (AA) by sex and age in the same-sex twins (n = 1893). Dotted line segments denote the lack of data for 43- to 49-year-olds.

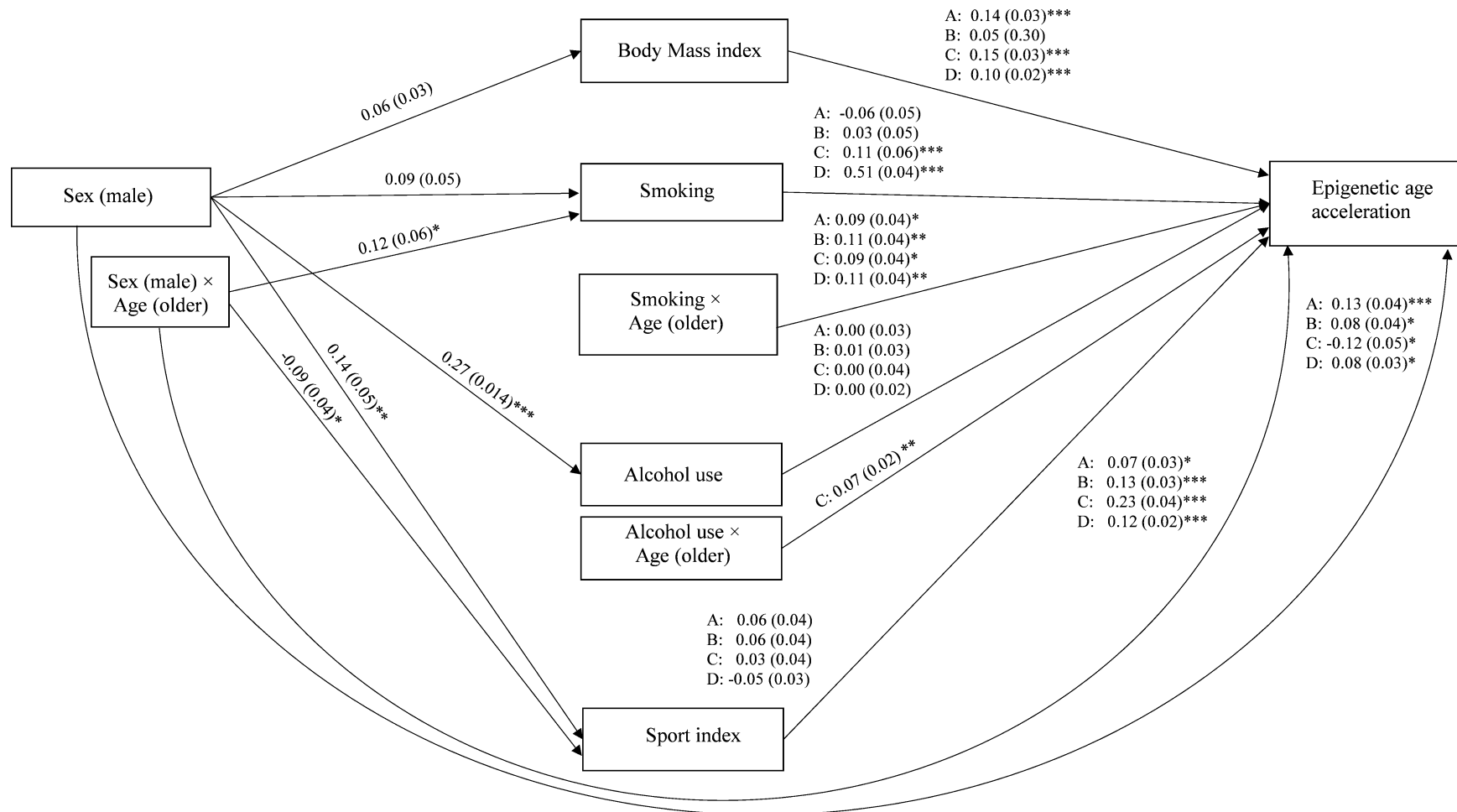


Figure 2. The path diagram of the multiple mediator model in the same-sex twins (n = 1893). Standardised regression coefficients (standard errors) are presented. The modelling was conducted separately for each epigenetic age acceleration (AA) measure: A, AA<sub>Horvath</sub>; B, AA<sub>Hannum</sub>; C, AA<sub>Pheno</sub>; D, AA<sub>Grim</sub>. \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ .

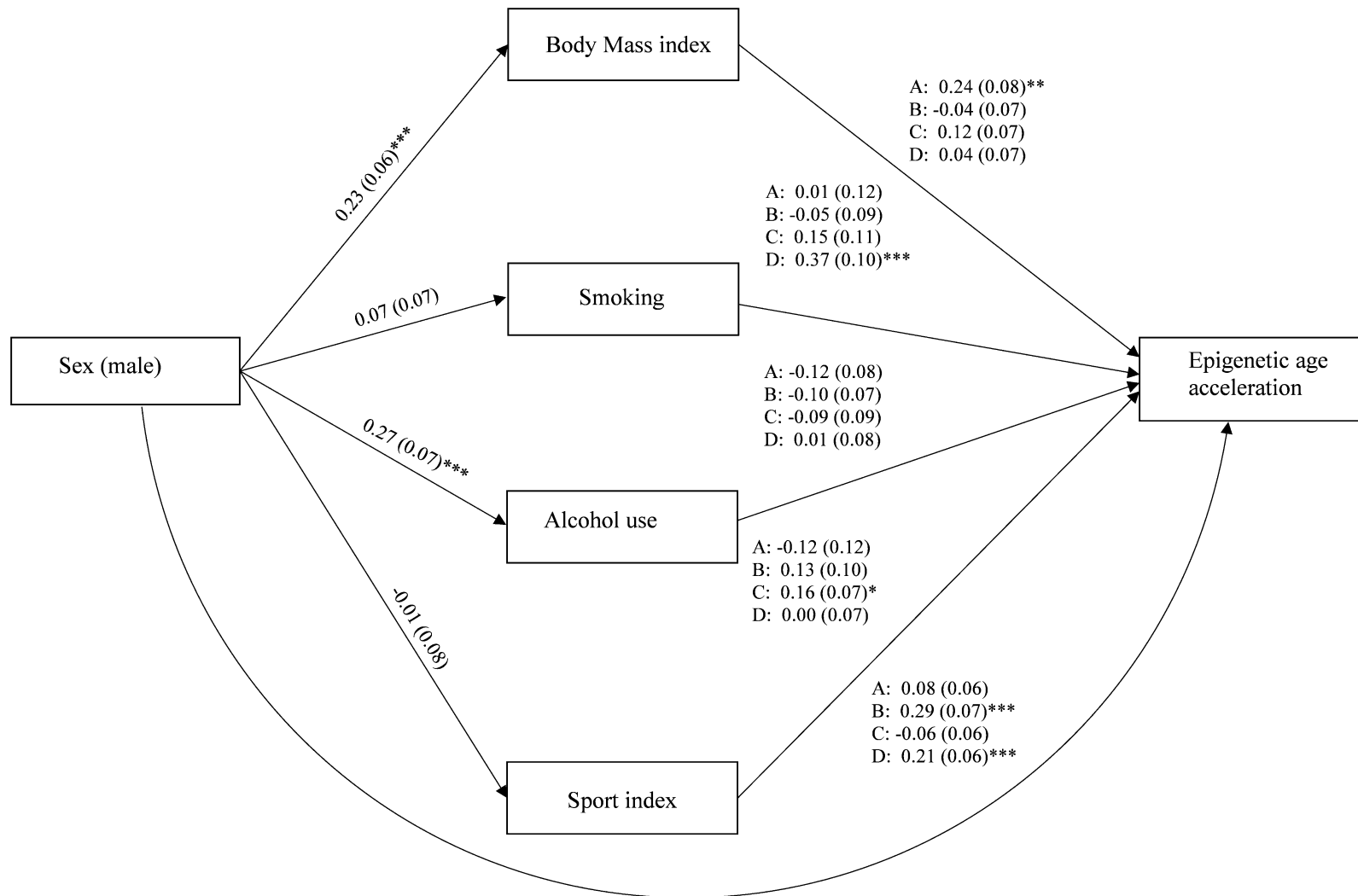


Figure 2. The path diagram of the multiple mediator model in the opposite-sex twin pairs (151 twin pairs). The standardised regression coefficients (standard errors) at the within-twin pair level are presented. The modelling was conducted separately for each epigenetic age acceleration (AA) measure: A, AA<sub>Horvath</sub>; B, AA<sub>Hannum</sub> C, AA<sub>Pheno</sub>; D, AA<sub>Grim</sub>. \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ .