

RESEARCH

Open Access



# Epigenetic age acceleration mediates the association between smoking and diabetes-related outcomes

Xue-Yong Chang<sup>1</sup> and Wan-Yu Lin<sup>1,2,3\*</sup>

## Abstract

**Background** Smoking can lead to the deterioration of lung function and susceptibility to diabetes. Recently, smoking was found to induce DNA methylation (DNAm) changes in some cytosine-phosphate-guanine sites (CpGs). As linear combinations of DNAm levels of aging-related CpGs, five measures of epigenetic age acceleration (EAA) have received extensive attention: HannumEAA, IEAA, PhenoEAA, GrimEAA, and DunedinPACE. It is of interest to explore whether some measures of EAA can mediate the associations of smoking with diabetes-related outcomes and indices of ventilatory lung function.

**Methods and results** In this study, we included self-reported smoking variables (smoking status, the number of pack-years, and years since smoking cessation), seven DNAm markers (HannumEAA, IEAA, PhenoEAA, GrimEAA, DNAm-based smoking pack-years, DNAm plasminogen activator inhibitor 1 [PAI-1] levels, and DunedinPACE), and four health outcomes (fasting glucose, hemoglobin A1C, forced expiratory volume in 1.0 s [FEV1], and forced vital capacity [FVC]) from 2474 Taiwan Biobank participants. Mediation analyses were conducted while adjusting for chronological age, sex, body mass index, drinking status, regular exercise status, educational attainment, and five cell-type proportions. We demonstrated that GrimEAA, DNAm-based smoking pack-years, DNAm PAI-1 levels, DunedinPACE, and PhenoEAA mediated smoking associations with diabetes-related outcomes. Moreover, current and former smoking both had an adverse indirect effect on FVC through DNAm PAI-1 levels. For former smokers, a long time since smoking cessation had a positive indirect impact on FVC through GrimEAA and on FEV1 through PhenoEAA.

**Conclusions** This is one of the first studies to comprehensively investigate the role of five measures of EAA in mediating the associations of smoking with the health outcomes of an Asian population. The results showed that the second-generation epigenetic clocks (GrimEAA, DunedinPACE, and PhenoEAA) significantly mediated the associations between smoking and diabetes-related outcomes. In contrast, the first-generation epigenetic clocks (HannumEAA and IEAA) did not significantly mediate any associations of smoking variables with the four health outcomes. Cigarette smoking can, directly and indirectly, deteriorate human health through DNAm changes in aging-related CpG sites.

**Keywords** Aging, Chronological age, DNA methylation, Epigenetic clock, Methylation clock

\*Correspondence:

Wan-Yu Lin

linwy@ntu.edu.tw

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Cigarette smoking can relieve stress and decrease subjective anxiety [1–5]. However, nicotine, the main chemical in cigarettes, affects lung function and leads to chronic obstructive pulmonary disease (COPD) [6, 7]. Furthermore, smoking is also associated with diabetes because of its deleterious effects on pancreatic functions [8–10]. The risks of developing diabetes and COPD are correlated with cigarette consumption [11].

The underlying link between smoking and diseases (such as diabetes [12] and COPD [13]) remains unclear. Smoking leads to two dynamic selection mechanisms [14]. First, smoking increases the risks of diabetes and the deterioration of lung function. Second, smokers with diabetes or worsening lung function are more likely to quit smoking [15]. Therefore, it is necessary to investigate the relationship between cigarettes and glycemic control (or pulmonary function).

As indicated by several epigenome-wide association studies (EWAS) [16–23], smoking is associated with DNA methylation (DNAm) changes at cytosine-phosphate-guanine (CpG) sites across the genome [24, 25]. In the last decade, some CpGs were integrated to predict human biological aging. Five epigenetic clocks have received widespread attention: HannumEAA [26], IEAA [27], PhenoEAA [28], GrimEAA [29, 30], and DunedinPACE [31], where EAA represents “epigenetic age acceleration”.

HannumEAA [26] and IEAA [27] are the so-called “first-generation epigenetic clocks”, as they were developed to predict chronological age. The other three clocks [28–31] are called the “second-generation epigenetic clocks”, which were derived to estimate physiological conditions and aging rates [32]. For example, PhenoEAA [28] is composed of 513 CpGs that can predict phenotypic age, which is a combination of ten indicators for immune, inflammation, metabolic, liver, and kidney conditions. GrimEAA comprises 1,030 CpGs associated with smoking pack-years and seven plasma proteins [29]. Among these plasma proteins, DNAm plasminogen activation inhibitor 1 (PAI-1) (estimated by 211 CpGs) has been shown to outperform GrimEAA in predicting several chronic diseases [29]. Recently, version 2 of GrimEAA [30] additionally included high-sensitivity C-reactive protein and hemoglobin A1C (HbA<sub>1c</sub>) to the original measure [29]. DunedinPACE [31] was derived from 173 CpGs related to declines in organ-system integrity.

GrimEAA and DNAm PAI-1 levels have been reported to be significantly associated with lung function decline (measured by forced expiratory volume in 1.0 s [FEV1] and forced vital capacity [FVC]) [33, 34]. Moreover, DunedinPACE and GrimEAA can reflect diabetes in Taiwanese adults [35]. These findings highlighted that EAA

might be a bridge linking smoking to diabetes and lung function reduction.

One mechanism by which active cigarette smoking may influence health is through EAA, a DNAm-based biomarker of aging [36]. Recently, based on a sample of 2978 participants representative of the U.S. population, the three “second-generation epigenetic clocks” were found to mediate a portion of the effects of smoking pack-years on mortality, heart disease, blood pressure levels, and cancer [36]. By analyzing Taiwan Biobank (TWB) DNAm data, we explored whether the five epigenetic clocks can mediate the effects of smoking on fasting glucose (FG), HbA<sub>1c</sub>, FEV1, and FVC. In addition to the five measures of EAA, two GrimAge components were investigated as mediators: DNAm-based smoking pack-years and DNAm PAI-1 levels. The former is more related to smoking, while the latter has been shown to outperform GrimEAA in predicting several age-related traits [29].

## Results

### Basic characteristics

Table 1 shows the primary characteristics of the 2474 TWB participants stratified by smoking status. All the individuals were between 30 and 70 years when participating in the TWB survey. Only 63.7% of the 2474 TWB participants underwent lung function examinations. The measurements of FEV1 and FVC were based on 1576 individuals (1190 non-smokers, 201 former smokers, and 185 current smokers).

According to the TWB questionnaire, current smokers were defined as those who “had smoked cigarettes for at least 6 months and had not quit smoking when participating in TWB.” In total, 235 male and 48 female participants were current smokers. Their average number of smoking pack-years was 20.8 (standard deviation [SD]=19.9).

Former smokers were defined as those who “had previously smoked cigarettes for at least 6 months but had quit smoking for at least 6 months when participating in TWB.” In total, 272 male and 40 female participants were former smokers. Their average number of smoking pack-years was 12.9 (SD=14.1), and the average time since smoking cessation was 12.9 (SD=10.0) years.

On average, former smokers (mean age=52.9, SD=10.2 years) were older than current smokers (mean age=48.8, SD=10.6 years) and non-smokers (mean age=49.4, SD=11.2 years) (Kruskal–Wallis test  $p$ -value=3.9E-7). There was a strong association between sex and smoking status (Chi-square test  $p$ -value=3.8E-84). Over 80% of current and former smokers were males, while more than 60% of non-smokers were females.

Drinking was defined as those who “had a weekly intake of more than 150 mL of alcoholic beverages for at

**Table 1** Basic characteristics of the 2474 participants stratified by smoking status

	Non-smokers	Former smokers <sup>a</sup>	Current smokers <sup>b</sup>	<i>p</i> -value <sup>c</sup>
Total	1879 (75.95%)	312 (12.61%)	283 (11.44%)	
Age (year)	49.4 (11.2)	52.9 (10.2)	48.8 (10.6)	3.9E-07
Age range	30–70	30–70	30–70	
Sex (male)	736 (39.2%)	272 (87.2%)	235 (83.0%)	3.8E-84
Drinking <sup>d</sup>	72 (3.8%)	30 (9.6%)	69 (24.4%)	1.2E-36
Former smokers' pack-years	–	12.9 (14.1)	–	
Current smokers' pack-years	–	–	20.8 (19.9)	
Years since smoking cessation	–	12.9 (10.0)	–	
Regular exercise <sup>e</sup>	819 (43.6%)	178 (57.1%)	95 (33.6%)	3.0E-08
Educational attainment <sup>f</sup>	5.6 (0.9)	5.6 (0.9)	5.4 (0.8)	2.2E-04
BMI (kg/m <sup>2</sup> )	24.1 (3.7)	25.3 (3.2)	25.4 (3.7)	1.3E-14
FG (mg/dL)	94.4 (16.5)	98.9 (20.3)	102.2 (34.0)	8.0E-16
HbA <sub>1c</sub> (%)	5.7 (0.7)	5.8 (0.7)	5.9 (1.0)	1.7E-04
FEV1 (L) <sup>g</sup>	2.2 (0.8)	2.4 (0.8)	2.5 (0.8)	1.0E-09
FVC (L) <sup>g</sup>	3.0 (0.9)	3.4 (0.7)	3.5 (0.7)	3.9E-28

The characteristics are presented in *n* (%) or mean ± standard deviation

<sup>a</sup> Former smokers were defined as those who “had previously smoked cigarettes for at least 6 months but had quit smoking for at least 6 months when participating in TWB”

<sup>b</sup> Current smokers were defined as those who “had smoked cigarettes for at least 6 months and had not quit smoking when participating in TWB”

<sup>c</sup> The *p*-value of testing the mean or proportion difference among non-smokers, former smokers, and current smokers was based on a Kruskal–Wallis test for continuous variables or a Chi-square test for sex, drinking, and regular exercise

<sup>d</sup> Drinking was defined as those who “had a weekly intake of more than 150 mL of alcoholic beverages for at least 6 months and had not stopped drinking when participating in TWB”

<sup>e</sup> Regular exercise was defined as “performing exercise for 30 min thrice a week”. “Exercise” included leisure-time activities such as jogging, swimming, cycling, yoga, resistance training, hiking, etc.

<sup>f</sup> Educational attainment was an integer ranging from 1 to 7: 1 “illiterate”, 2 “no formal education but literate”, 3 “primary school graduate”, 4 “junior high school graduate”, 5 “senior high school graduate”, 6 “college graduate”, and 7 “Master’s or higher degree”

<sup>g</sup> Only 63.7% of the 2,474 TWB participants underwent lung function examinations. The measurements of FEV1 and FVC were based on 1576 individuals (1190 non-smokers, 201 former smokers, and 185 current smokers)

least 6 months and had not stopped drinking when participating in TWB”. The percentage of drinking was largest in current smokers (24.4%) compared with former smokers (9.6%) and non-smokers (3.8%) (Chi-square test *p*-value = 1.2E-36).

Regular exercise was defined as “performing exercise for 30 min thrice a week”. “Exercise” included leisure-time activities such as jogging, swimming, cycling, yoga, resistance training, hiking, etc. The percentage of individuals performing regular exercise was highest in former smokers (57.1%) compared with current smokers (33.6%) and non-smokers (43.6%) (Chi-square test *p*-value = 3.0E-8).

Educational attainment was an integer ranging from 1 to 7: 1 “illiterate”, 2 “no formal education but literate”, 3 “primary school graduate”, 4 “junior high school graduate”, 5 “senior high school graduate”, 6 “college graduate”, and 7 “Master’s or higher degree”. The average educational attainment scores were 5.6 (SD=0.9), 5.6 (SD=0.9), and 5.4 (SD=0.8) for non-smokers, former smokers, and current smokers, respectively. This

indicates that the average educational attainment for the 2,474 TWB participants was between “senior high school graduate” and “college graduate”. However, on average, non-smokers and former smokers had higher educational attainment than current smokers (Kruskal–Wallis test *p*-value = 0.00022).

Because of the different male–female proportions across the three groups, current and former smokers demonstrated larger body mass index (BMI) and higher levels of four outcomes (FG, HbA<sub>1c</sub>, FEV1, and FVC) than non-smokers. Given that smoking prevalence and health outcomes significantly vary by sex, we also performed a sensitivity analysis to test the mediation effects in males only (there was insufficient power to implement such tests in females due to their low prevalence of smoking).

#### Exposure-mediator relationship (X→M)

Boxplots of the seven DNAm markers are shown in Figure S1 (Additional file 1). Before performing mediation analysis, we excluded 7, 1, 2, 5, 54, and 1 extreme outliers of HannumEAA, IEAA, PhenoEAA, GrimEAA,

DNA-based smoking pack-years, and DunedinPACE, respectively. Extreme outliers were defined by values smaller than  $Q_1 - 3 \times (Q_3 - Q_1)$  or larger than  $Q_3 + 3 \times (Q_3 - Q_1)$ , where  $Q_1$  and  $Q_3$  are the 25th and 75th percentiles of an EAA, respectively. We detected more extreme outliers in DNA-based smoking pack-years because it demonstrated the most right-skewed distribution among the seven markers (skewness=1.65, Additional file 1: Figure S1). In contrast, we found no extreme outliers in DNAm PAI-1 levels (skewness=0.12, Additional file 1: Figure S1).

After removing the extreme outliers, we presented the heatmaps of Pearson's correlation coefficients among variables in Additional file 1: Figures S2 (both males and females) and S3 (only in males). Because we obtained four measures of EAA (HannumEAA, IEAA, PhenoEAA, and GrimEAA) as the residuals of regressing epigenetic age on chronological age, their correlations with chronological age were close to 0 (Additional file 1: Figure S2). DunedinPACE, DNA-based smoking pack-years, and DNAm PAI-1 levels were positively associated with chronological age (Additional file 1: Figures S2–S3). The analysis focusing on males (Additional file 1: Figure S3) showed that FEV1 and FVC were inversely associated with the second-generation epigenetic clocks (DunedinPACE, GrimEAA, and PhenoEAA) and the two components of GrimEAA (DNA-based smoking pack-years and DNAm PAI-1 levels).

We then evaluated the exposure–mediator relationship ( $X \rightarrow M$ ). Table 2 shows the results of regressing each DNAm marker ( $M$ ) on smoking variables ( $X$ ). Table 1 shows that chronological age (in years), sex (male vs. female), BMI (in  $\text{kg}/\text{m}^2$ ), drinking status (yes vs. no), performing regular exercise (yes vs. no), and educational attainment (integer ranging from 1 to 7) are different across the three smoking groups. Therefore, we adjusted for these six covariates in the regression models. Moreover, we also adjusted for five cell-type proportions (B lymphocytes,  $\text{CD4}^+$  T cells,  $\text{CD8}^+$  T cells, monocytes, and natural killer cells) estimated by the Houseman deconvolution method [37]. Cell-type composition is critical because the TWB acquired the DNAm data from peripheral blood rather than other bulk tissues [38]. We had seven DNAm markers and five smoking variables; therefore,  $p$ -values  $< 0.05/(7 \times 5) = 0.0014$  were considered significant.

As shown in Table 2, GrimEAA and DNA-based smoking pack-years were associated with all five smoking variables ( $p < 0.0014$ ). PhenoEAA was associated with all smoking variables except years since smoking cessation ( $p = 0.0028$ ). DunedinPACE, DNAm PAI-1 levels, and HannumEAA were positively associated with current smoking status and current smokers' pack-years. IEAA

was not associated with any of the five smoking variables (Table 2). Therefore, IEAA is not a plausible mediator and will be omitted from the following mediation analysis.

Current smokers, on average, were 5.650 years older in GrimEAA ( $p = 1.0\text{E-}150$ ), 1.930 years older in PhenoEAA ( $p = 2.8\text{E-}10$ ), 1.205 years older in HannumEAA ( $p = 2.1\text{E-}07$ ), and had a faster pace of biological aging of 0.084 years per chronological year than non-smokers ( $p = 7.9\text{E-}38$ ). Each additional current smokers' pack-year was associated with 0.174 years larger in GrimEAA ( $p = 2.4\text{E-}127$ ), 0.043 years larger in PhenoEAA ( $p = 2.2\text{E-}5$ ), 0.027 years larger in HannumEAA ( $p = 0.0003$ ), and a faster pace of biological aging of 0.002 years per chronological year ( $p = 5.0\text{E-}29$ ). DNAm-based smoking pack-years and DNAm PAI-1 levels were also significantly associated with current smoking status and current smokers' pack-years.

Former smoking status and former smokers' pack-years were associated with DNAm-based smoking pack-years, GrimEAA, and PhenoEAA. Former smokers' years since smoking cessation were negatively associated with GrimEAA and DNAm-based smoking pack-years. One more year since smoking cessation was associated with a decreased GrimEAA by 0.108 years ( $p = 2.0\text{E-}8$ ) and a decreased DNAm-based smoking pack-years by 0.313 ( $p = 1.2\text{E-}13$ ).

#### Mediation analysis results ( $X \rightarrow M \rightarrow Y$ )

Mediation analysis results are shown in Tables 3, 4, 5, 6, 7 and Figs. 1, 2, 3, 4, 5. We first tested the statistical significance of 20  $X$ – $Y$  associations. The total effects of  $X$  (smoking variable) on  $Y$  (health outcome) were considered significant given the  $p$ -values  $< 0.05$ . We did not adjust for multiple testing in this stage because the total effects of  $X$  on  $Y$  were not the main objective of this study. Four of the 20  $X$ – $Y$  associations had significant  $p$ -values of  $< 0.05$ : current smoking status on FG ( $p = 0.00012$ , Table 4) and  $\text{HbA}_{1c}$  ( $p = 0.048$ , Table 4); current smokers' pack-years on FG ( $p = 1.4\text{E-}7$ , Table 6) and  $\text{HbA}_{1c}$  ( $p = 0.003$ , Table 6).

The significance of the  $X$ – $Y$  association is not a requirement for the following  $X \rightarrow M \rightarrow Y$  mediation analysis [39, 40] because (1) when the effect size is small, the sample size may not be sufficient to detect the effect of  $X$  on  $Y$  [40]; and (2) two or more indirect effects with opposite directions may cancel each other out, producing an insignificant total effect [41].

Former smoking status, former smokers' pack-years, and years since smoking cessation were not significantly associated with health outcomes (Tables 3, 5, and 7). However, our analysis of all 173,807 TWB participants (Additional file 1: Table S1) showed that 18 of the 20  $X$ – $Y$  associations were significant ( $p < 0.05$ ; Additional file 1:

**Table 2** Results of regressing each DNAm marker on smoking variables (based on models 1, 3, and 5)

	HannumEAA (in years)		IEAA (in years)		PhenoEAA (in years)		GrimEAA (in years)		DNAm-based smoking pack-years		DNAm PAI-1 (in pg/mL)		DunedinPACE	
	Effect (S.E.)	p-value	Effect (S.E.)	p-value	Effect (S.E.)	p-value	Effect (S.E.)	p-value	Effect (S.E.)	p-value	Effect (S.E.)	p-value	Effect (S.E.)	p-value
Former smoking status (yes vs. no) (based on model 1)	0.426 (0.216)	0.0489	0.214 (0.243)	0.3781	<b>1.011 (0.284)</b>	<b>0.0004</b>	<b>1.446 (0.187)</b>	<b>1.5E-14</b>	<b>4.705 (0.338)</b>	<b>2.0E-42</b>	263.389 (131.291)	0.0449	0.019 (0.006)	0.0020
Current smoking status (yes vs. no) (based on model 1)	<b>1.205 (0.231)</b>	<b>2.1E-07</b>	0.367 (0.260)	0.1582	<b>1.930 (0.305)</b>	<b>2.8E-10</b>	<b>5.650 (0.201)</b>	<b>1.0E-150</b>	<b>14.916 (0.389)</b>	<b>9.6E-252</b>	<b>1002.309 (140.563)</b>	<b>1.3E-12</b>	<b>0.084 (0.006)</b>	<b>7.9E-38</b>
Former smokers' pack-years (based on model 3)	0.021 (0.011)	0.0480	0.021 (0.012)	0.0764	<b>0.051 (0.014)</b>	<b>0.0003</b>	<b>0.087 (0.009)</b>	<b>5.0E-20</b>	<b>0.255 (0.018)</b>	<b>1.9E-44</b>	12.694 (6.549)	0.0527	0.001 (3.0E-04)	0.0017
Current smokers' pack-years (based on model 3)	<b>0.027 (0.008)</b>	<b>0.0003</b>	0.003 (0.009)	0.6907	<b>0.043 (0.010)</b>	<b>2.2E-05</b>	<b>0.174 (0.007)</b>	<b>2.4E-127</b>	<b>0.567 (0.018)</b>	<b>2.8E-185</b>	<b>28.240 (4.629)</b>	<b>1.2E-09</b>	<b>0.002 (2.1E-04)</b>	<b>5.0E-29</b>
Years since smoking cessation (based on model 5)	-0.029 (0.024)	0.2339	-0.038 (0.027)	0.1555	-0.091 (0.030)	0.0028	<b>-0.108 (0.019)</b>	<b>2.0E-08</b>	<b>-0.313 (0.040)</b>	<b>1.2E-13</b>	-16.028 (12.937)	0.2164	-0.001 (0.001)	0.2770

Significant results with  $p < 0.05$  ( $7 \times 5$ ) = 0.0014 are highlighted in bold font (7: seven DNAm markers; 5: five smoking variables)

**Table 3** Results of six DNAm markers in mediating the associations between former smoking status and four health outcomes

Outcome <sup>a</sup>	Total effect	95% Confidence interval		p-value	Sample size		
FG	0.0442	- 0.0761	0.1644	0.471	2469		
HbA <sub>1c</sub>	0.0253	- 0.0941	0.1448	0.677	2469		
FEV1	- 0.0820	- 0.2155	0.0516	0.229	1576		
FVC	0.0427	- 0.0667	0.1520	0.444	1576		
Outcome	Mediator	Mediation effect <sup>b</sup>	95% confidence interval		FDR <sup>c</sup>	Proportion mediated <sup>d</sup> (%)	Sample size
FG	HannumEAA	0.0041	- 0.0007	0.0127	0.264	9.3	2462
HbA <sub>1c</sub>		0.0046	- 0.0005	0.0132	0.209	18.2	2462
FEV1		0.0038	- 0.0062	0.0156	0.597	- 4.6	1569
FVC		- 0.0027	- 0.0129	0.0061	0.710	- 6.3	1569
FG	PhenoEAA	0.0086	- 0.0001	0.0217	0.138	19.5	2467
HbA <sub>1c</sub>		<b>0.0143</b>	<b>0.0039</b>	<b>0.0286</b>	<b>0.021</b>	56.5	2467
FEV1		- 0.0039	- 0.0188	0.0101	0.710	4.8	1574
FVC		- 0.0015	- 0.0146	0.0111	0.927	- 3.5	1574
FG	GrimEAA	<b>0.0609</b>	<b>0.0361</b>	<b>0.0916</b>	<b>&lt;0.001</b>	137.8	2464
HbA <sub>1c</sub>		<b>0.0710</b>	<b>0.0443</b>	<b>0.1016</b>	<b>&lt;0.001</b>	280.6	2464
FEV1		- 0.0043	- 0.0246	0.0174	0.816	5.2	1572
FVC		- 0.0184	- 0.0372	- 0.0016	0.108	- 43.1	1572
FG	DNAm-based pack-years	<b>0.0511</b>	<b>0.0133</b>	<b>0.0914</b>	<b>0.028</b>	115.6	2415
HbA <sub>1c</sub>		<b>0.0667</b>	<b>0.0297</b>	<b>0.1075</b>	<b>&lt;0.001</b>	263.6	2415
FEV1		0.0019	- 0.0370	0.0427	0.957	- 2.3	1538
FVC		0.0093	- 0.0196	0.0406	0.707	21.8	1538
FG	DNAm PAI-1	0.0268	0.0005	0.0543	0.120	60.6	2469
HbA <sub>1c</sub>		0.0252	0.0016	0.0513	0.117	99.6	2469
FEV1		- 0.0073	- 0.0210	0.0026	0.294	8.9	1576
FVC		- <b>0.0132</b>	- <b>0.0276</b>	- <b>0.0026</b>	<b>0.026</b>	- 30.9	1576
FG	DunedinPACE	<b>0.0117</b>	<b>0.0021</b>	<b>0.0244</b>	<b>0.024</b>	26.5	2468
HbA <sub>1c</sub>		<b>0.0199</b>	<b>0.0068</b>	<b>0.0365</b>	<b>0.007</b>	78.7	2468
FEV1		0.0002	- 0.0079	0.0090	0.957	- 0.2	1575
FVC		- 0.0048	- 0.0137	0.0009	0.246	- 11.2	1575

<sup>a</sup> Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

<sup>b</sup> Mediation effect is the effect of X (former smoking status) on Y (health outcome) through the mediator (DNAm marker)

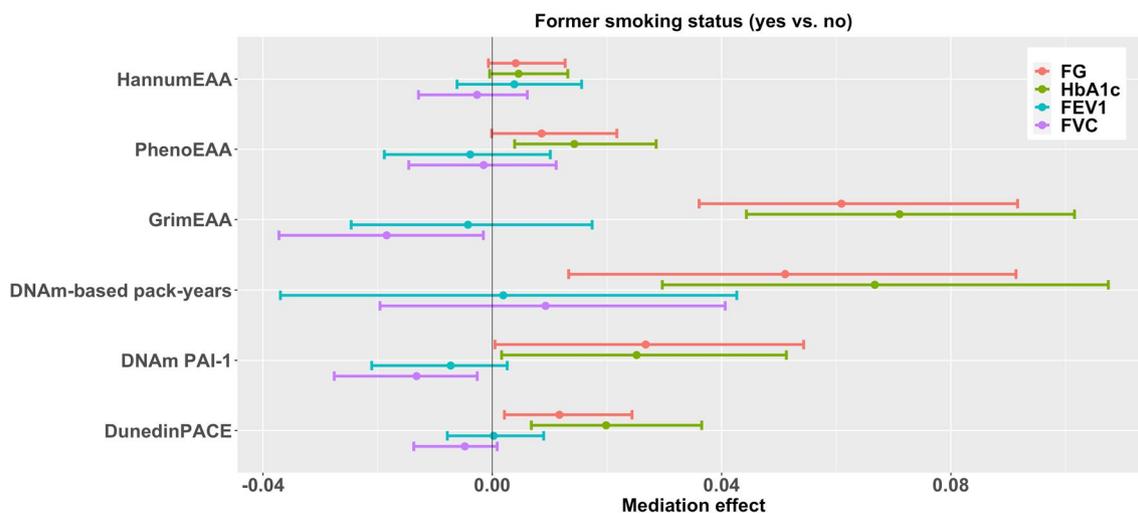
<sup>c</sup> FDR (false discovery rate) represents the FDR-adjusted p-value. Significant effects with FDR < 0.05 are marked in bold

<sup>d</sup> Proportion mediated was calculated by dividing the mediation effect by the total effect

Table S2). Moreover, according to previous studies, the time since smoking cessation is inversely associated with diabetes risk [42]; former smoking is related to lasting damage to lung functions [43]. Our analysis of all 173,807 TWB participants and these prior findings encouraged us to continue the mediation analysis even when the X–Y associations were insignificant in the sample of 2474 individuals.

No significant mediation effects were observed in HannumEAA (false discovery rate [FDR] > 0.05, Tables 3, 4, 5, 6, 7, Figs. 1, 2, 3, 4, 5). The associations between current smoking and diabetes-related outcomes (FG and HbA<sub>1c</sub>) were significantly mediated by GrimEAA, DNAm-based

smoking pack-years, DNAm PAI-1 levels, DunedinPACE, and PhenoEAA (only HbA<sub>1c</sub>) (FDR < 0.05, Table 4 and Fig. 2). GrimEAA mediated 94.0% of current smoking's effect on FG and 214.7% of current smoking's effect on HbA<sub>1c</sub>. DNAm-based smoking pack-years mediated 64.0% of current smoking's impact on FG and 163.6% of current smoking's impact on HbA<sub>1c</sub>. DNAm PAI-1 levels mediated 40.2% of current smoking's effect on FG and 74.1% of current smoking's effect on HbA<sub>1c</sub>. DunedinPACE mediated 20.9% of current smoking's impact on FG and 69.4% of current smoking's impact on HbA<sub>1c</sub>. PhenoEAA mediated 21.1% of current smoking's effect on HbA<sub>1c</sub>.



**Fig. 1** 95% Confidence intervals of the effects of six DNAm markers in mediating the associations between former smoking status and four health outcomes. Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

The proportion mediated was over 100% because the direct and mediation effects were in opposite directions. Taking current smoking status → GrimEAA → HbA<sub>1c</sub> as an example, the mediation effect was 0.2776 (Table 4), the direct effect was -0.1483 (*p*=0.083), and the total effect was 0.2776-0.1483=0.1293 (Table 4).

The associations between current smokers’ pack-years and diabetes-related outcomes (FG and HbA<sub>1c</sub>) were significantly mediated by GrimEAA, DNAm PAI-1 levels, DunedinPACE, and PhenoEAA (only HbA<sub>1c</sub>) (FDR<0.05, Table 6 and Fig. 4). GrimEAA mediated 56.1% of the total effect of current smokers’ pack-years on FG and 120.6% of the total effect of current smokers’ pack-years on HbA<sub>1c</sub>. DNAm PAI-1 mediated 25.4% of the total impact of current smokers’ pack-years on FG and 42.9% of the total impact of current smokers’ pack-years on HbA<sub>1c</sub>. DunedinPACE mediated 12.3% of the total effect of current smokers’ pack-years on FG and 38.1% of the total effect of current smokers’ pack-years on HbA<sub>1c</sub>. PhenoEAA mediated 9.5% of the total impact of current smokers’ pack-years on HbA<sub>1c</sub>.

Moreover, years since smoking cessation positively affected FVC through GrimEAA (FDR=0.026) and FEV1 through PhenoEAA (FDR=0.026, Table 7 and Fig. 5). Former smoking had an adverse indirect effect on diabetes-related outcomes through GrimEAA, DNAm-based smoking pack-years, DunedinPACE, and PhenoEAA (only HbA<sub>1c</sub>), and former smoking exerted an indirect negative impact on FVC through DNAm PAI-1 levels (FDR<0.05, Table 3 and Fig. 1). Longer former smokers’ pack-years had an adverse indirect effect on diabetes-related outcomes through GrimEAA, DunedinPACE

(only HbA<sub>1c</sub>), and PhenoEAA (only HbA<sub>1c</sub>) (FDR<0.05, Table 5 and Fig. 3).

Among the 2474 individuals, 1243 were males. The primary characteristics of the 1243 male participants stratified by smoking status are shown in Additional file 1: Table S3. The exposure-mediator relationship (X→M, Additional file 1: Table S4) and the mediation analysis results for male participants (X→M→Y, Additional file 1: Tables S5–S9 and Figures S4–S8) are similar to those obtained from both males and females (Tables 2, 3, 4, 5, 6, 7 and Figs. 1, 2, 3, 4, 5).

### Discussion

Previous research suggests that DNAm is a plausible biological pathway linking smoking exposure to health span [36]. By analyzing the DNAm data of 2978 participants from the Health and Retirement Study (HRS) 2016 Venous Blood Study, a recent study showed that the three second-generation epigenetic clocks (PhenoEAA, GrimEAA, and DunedinPoAm38 [the previous version of DunedinPACE]) [28, 29, 31, 44] mediated the effects of smoking pack-years on cancer, heart disease, high blood pressure, and mortality [36]. Moreover, GrimEAA and DunedinPoAm38 also mediated the association between smoking and lung diseases such as emphysema or chronic bronchitis [36]. The first-generation epigenetic clocks (HannumEAA and IEAA) [26, 27] were not mediators of the association between smoking exposure and any health outcomes in that study (cancer, heart disease, high blood pressure, lung disease, and mortality) [36]. These results were based on 2,978 U.S. adults over

**Table 4** Results of six DNAm markers in mediating the associations between current smoking status and four health outcomes

Outcome <sup>a</sup>	Total effect	95% confidence interval		p-value	Sample size
FG	<b>0.2533</b>	<b>0.1245</b>	<b>0.3820</b>	<b>1.2E-04</b>	2469
HbA <sub>1c</sub>	<b>0.1293</b>	<b>0.0014</b>	<b>0.2572</b>	<b>0.048</b>	2469
FEV1	0.0208	- 0.1229	0.1645	0.777	1576
FVC	0.0826	- 0.0350	0.2003	0.169	1576

Outcome	Mediator	Mediation effect <sup>b</sup>	95% confidence interval		FDR <sup>c</sup>	Proportion mediated <sup>d</sup> (%)	Sample size
FG	HannumEAA	0.0116	- 0.0013	0.0267	0.183	4.6	2462
HbA <sub>1c</sub>		0.0130	- 0.0011	0.0296	0.145	10.1	2462
FEV1		0.0053	- 0.0076	0.0210	0.590	25.5	1569
FVC		- 0.0037	- 0.0160	0.0080	0.707	- 4.5	1569
FG	PhenoEAA	0.0164	- 0.0007	0.0350	0.145	6.5	2467
HbA <sub>1c</sub>		<b>0.0273</b>	<b>0.0091</b>	<b>0.0483</b>	<b>0.013</b>	21.1	2467
FEV1		- 0.0058	- 0.0269	0.0160	0.767	- 27.9	1574
FVC		- 0.0023	- 0.0206	0.0152	0.927	- 2.8	1574
FG	GrimEAA	<b>0.2381</b>	<b>0.1457</b>	<b>0.3392</b>	<b>&lt; 0.001</b>	94.0	2464
HbA <sub>1c</sub>		<b>0.2776</b>	<b>0.1868</b>	<b>0.3744</b>	<b>&lt; 0.001</b>	214.7	2464
FEV1		- 0.0172	- 0.1029	0.0746	0.840	- 82.7	1572
FVC		- 0.0745	- 0.1468	- 0.0040	0.108	- 90.2	1572
FG	DNAm-based pack-years	<b>0.1620</b>	<b>0.0448</b>	<b>0.2918</b>	<b>0.028</b>	64.0	2415
HbA <sub>1c</sub>		<b>0.2115</b>	<b>0.0899</b>	<b>0.3360</b>	<b>0.007</b>	163.6	2415
FEV1		0.0062	- 0.1168	0.1360	0.957	29.8	1538
FVC		0.0302	- 0.0671	0.1316	0.710	36.6	1538
FG	DNAm PAI-1	<b>0.1018</b>	<b>0.0665</b>	<b>0.1406</b>	<b>&lt; 0.001</b>	40.2	2469
HbA <sub>1c</sub>		<b>0.0958</b>	<b>0.0634</b>	<b>0.1357</b>	<b>&lt; 0.001</b>	74.1	2469
FEV1		- 0.0176	- 0.0439	0.0070	0.285	- 84.6	1576
FVC		- <b>0.0320</b>	- <b>0.0575</b>	- <b>0.0103</b>	<b>0.016</b>	- 38.7	1576
FG	DunedinPACE	<b>0.0529</b>	<b>0.0114</b>	<b>0.0928</b>	<b>0.045</b>	20.9	2468
HbA <sub>1c</sub>		<b>0.0897</b>	<b>0.0496</b>	<b>0.1347</b>	<b>&lt; 0.001</b>	69.4	2468
FEV1		0.0013	- 0.0379	0.0410	0.957	6.3	1575
FVC		- 0.0260	- 0.0600	0.0066	0.254	- 31.5	1575

<sup>a</sup> Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

<sup>b</sup> Mediation effect is the effect of X (current smoking status) on Y (health outcome) through the mediator (DNAm marker)

<sup>c</sup> FDR (false discovery rate) represents the FDR-adjusted p-value. Significant effects with FDR < 0.05 are marked in bold

<sup>d</sup> Proportion mediated was calculated by dividing the mediation effect by the total effect

50 years old (mean age = 67.65; SD = 9.61 years; range: 50–100). 73% of the sample were non-Hispanic White, 12% were non-Hispanic Black, 11% were Hispanic, and 4% were other non-Hispanic ethnicities.

In our study, by analyzing the DNAm data of 2474 TWB participants, we showed that the associations between current smoking and diabetes-related indicators (FG and HbA<sub>1c</sub>) were mediated by GrimEAA, DunedinPACE, and PhenoEAA (Tables 4 and 6, Figs. 2 and 4). A plausible mediator should meet two requirements: (1) the association between smoking (X) and EAA (M); and (2) the association between EAA (M) and the health outcome (Y, here, diabetes-related indicators) while

controlling for smoking status (X). GrimEAA was derived from 1030 CpGs associated with smoking pack-years and nine plasma proteins [30]. GrimEAA and its component (DNAm-based smoking pack-years) are the most related to the five smoking variables (Table 2). They are good mediators of the association between smoking and diabetes-related indicators. To our knowledge, this is the first study linking smoking with diabetes through EAA.

We also show that former smoking adversely affects diabetes-related outcomes through the second-generation epigenetic clocks and FVC through DNAm PAI-1 levels (Table 3). However, quitting smoking earlier positively affects FVC through GrimEAA and FEV1 through

**Table 5** Results of six DNAm markers in mediating the associations between former smokers' pack-years and four health outcomes

Outcome <sup>a</sup>	Total effect	95% confidence interval		p-value	Sample size
FG	0.0046	- 0.0013	0.0106	0.129	2452
HbA <sub>1c</sub>	0.0050	- 0.0010	0.0109	0.102	2452
FEV1	- 0.0017	- 0.0084	0.0051	0.629	1563
FVC	- 0.0010	- 0.0065	0.0045	0.726	1563

Outcome	Mediator	Mediation effect <sup>b</sup>	95% confidence interval		FDR <sup>c</sup>	Proportion mediated <sup>d</sup> (%)	Sample size
FG	HannumEAA	0.0002	- 3.0E-05	0.0006	0.264	4.3	2445
HbA <sub>1c</sub>		0.0002	- 0.0001	0.0006	0.259	4.0	2445
FEV1		0.0002	- 0.0003	0.0008	0.575	- 11.8	1556
FVC		- 0.0001	- 0.0006	0.0004	0.767	10.0	1556
FG	PhenoEAA	0.0004	- 1.0E-05	0.0010	0.138	8.7	2450
HbA <sub>1c</sub>		<b>0.0007</b>	<b>0.0002</b>	<b>0.0013</b>	<b>0.013</b>	14.0	2450
FEV1		- 0.0002	- 0.0010	0.0005	0.767	11.8	1561
FVC		1.0E-05	- 0.0007	0.0007	0.968	- 1.0	1561
FG	GrimEAA	<b>0.0032</b>	<b>0.0018</b>	<b>0.0048</b>	<b>&lt; 0.001</b>	69.6	2448
HbA <sub>1c</sub>		<b>0.0038</b>	<b>0.0023</b>	<b>0.0057</b>	<b>&lt; 0.001</b>	76.0	2448
FEV1		0.0001	- 0.0013	0.0014	0.957	- 5.9	1560
FVC		- 0.0004	- 0.0016	0.0006	0.590	40.0	1560
FG	DNAm-based pack-years	0.0020	- 0.0001	0.0042	0.145	43.5	2399
HbA <sub>1c</sub>		0.0020	0.0001	0.0040	0.117	40.0	2399
FEV1		0.0006	- 0.0014	0.0028	0.710	- 35.3	1526
FVC		0.0017	0.0001	0.0035	0.108	- 170.0	1526
FG	DNAm PAI-1	0.0013	0.0002	0.0025	0.062	28.3	2452
HbA <sub>1c</sub>		0.0012	0.0002	0.0025	0.070	24.0	2452
FEV1		- 0.0003	- 0.0008	0.0001	0.336	17.6	1563
FVC		- 0.0004	- 0.0011	- 4.0E-05	0.101	40.0	1563
FG	DunedinPACE	0.0005	0.0001	0.0012	0.062	10.9	2451
HbA <sub>1c</sub>		<b>0.0010</b>	<b>0.0003</b>	<b>0.0018</b>	<b>0.016</b>	20.0	2451
FEV1		0.0001	- 0.0004	0.0006	0.906	- 5.9	1562
FVC		- 0.0002	- 0.0007	0.0002	0.501	20.0	1562

<sup>a</sup> Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

<sup>b</sup> Mediation effect is the effect of X (former smokers' pack-years) on Y (health outcome) through the mediator (DNAm marker)

<sup>c</sup> FDR (false discovery rate) represents the FDR-adjusted p-value. Significant effects with FDR < 0.05 are marked in bold

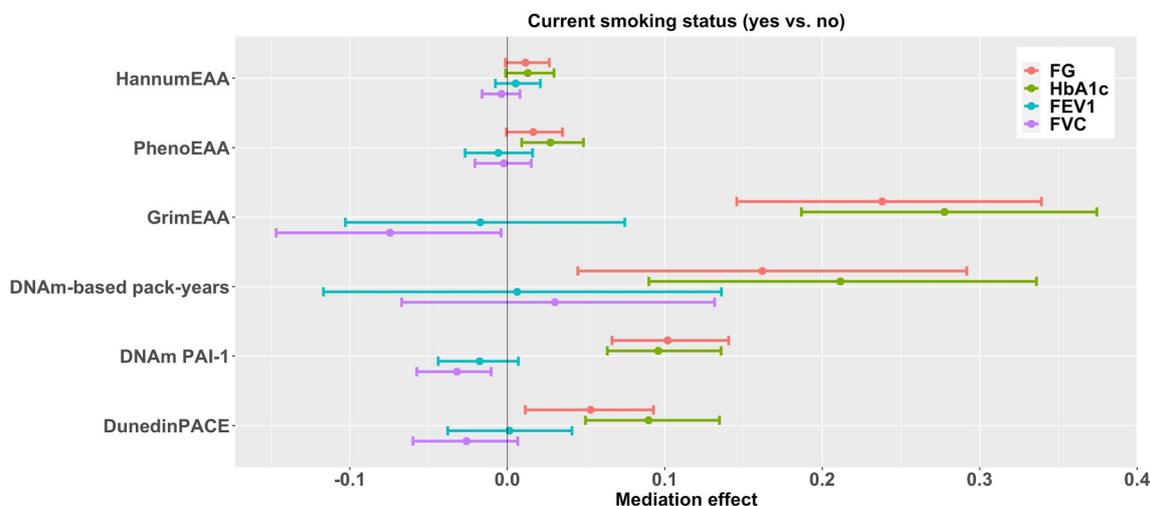
<sup>d</sup> Proportion mediated was calculated by dividing the mediation effect by the total effect

PhenoEAA (Table 7). Nonetheless, this is a cross-sectional study, and no temporal relationship exists between EAA and health outcomes. We cannot exclude the possibility that adverse health outcomes (i.e., diabetes or reduced ventilatory lung function) lead to altered EAA.

Although we here evaluate different outcomes, our results are consistent with previous findings [36] that the second-generation epigenetic clocks are mediators of the association between smoking exposure and health outcomes (cancer, heart disease, high blood pressure, lung disease, and mortality in [36]; FG, HbA<sub>1c</sub>, FEV1, and FVC in the present study). European data have suggested that GrimEAA and DNAm PAI-1 levels are inversely

associated with FVC [33]. This result aligns with our finding that smoking-related variables indirectly affect FVC through GrimEAA (Table 7) and DNAm PAI-1 levels (Tables 3, 4, and 6).

Biomarkers such as EAA can help figure out the mechanism linking smoking to health outcomes. Different from the first-generation epigenetic clocks (HannumEAA and IEAA) [26, 27], the second-generation epigenetic clocks (PhenoEAA, GrimEAA, and DunedinPACE) were derived from two-stage approaches. The first stage was to build some physiological measures, and the following step was to select CpGs that can predict the physiological measures [28–31]. Therefore,



**Fig. 2** 95% Confidence intervals of the effects of six DNAm markers in mediating the associations between current smoking status and four health outcomes. Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

compared with the first-generation epigenetic clocks, the second-generation clocks were more associated with health outcomes while controlling for the smoking variables, meeting the second requirement of a mediator.

In this work, GrimEAA indicated the EAA based on version 2 of DNAm GrimAge [30]. In the 20 mediation analyses ( $X \rightarrow \text{GrimEAA} \rightarrow Y$ , where X and Y are five smoking variables and four health outcomes, respectively), version 2 of GrimEAA [30] exhibited improvements over version 1 of GrimEAA [29]. Specifically, version 2 of GrimEAA showed eleven lower mediation *p*-values, two larger mediation *p*-values, and seven identical *p*-values < 0.001 (Additional file 1: Figure S9) compared with version 1 of GrimEAA. We also observed similar results when the analysis focused on males (Additional file 1: Figure S10). Version 2 of GrimEAA is a better mediator than version 1 of GrimEAA regarding the research topic in this study.

This is one of the first studies to show that the second-generation epigenetic clocks can mediate smoking associations with diabetes-related outcomes, based on rarely reported Asian data. Compared with other studies treating EAA as a mediator [36, 45], our study was unique in analyzing an East Asian sample. Moreover, our study sample was relatively younger (mean age = 49.76; SD = 11.08 years; range: 30–70) than that derived from the HRS 2016 Venous Blood Study [36, 45]. Finally, we investigated different health outcomes from previous studies [36, 45]. This work can fill the research gap regarding the role of EAA as a mediator between smoking and health.

## Conclusions

In this study, we comprehensively investigated the ability of the five measures of EAA to mediate the associations of smoking with four health outcomes of an East Asian population. The results showed that the second-generation epigenetic clocks (GrimEAA, DunedinPACE, and PhenoEAA) significantly mediated the associations between smoking and diabetes-related outcomes. Former smoking adversely affected diabetes-related outcomes through the second-generation epigenetic clocks and FVC through DNAm PAI-1 levels. However, quitting smoking earlier positively affected FVC through GrimEAA and FEV1 through PhenoEAA. In contrast, the first-generation epigenetic clocks (HannumEAA and IEAA) did not significantly mediate any effects of smoking variables on the four health outcomes (FG, HbA<sub>1c</sub>, FEV1, and FVC). Active cigarette smoking can, directly and indirectly, deteriorate human health through DNAm changes in aging-related CpG sites.

## Methods

### Study participants

Since October 2012, the TWB has recruited 173,807 community-based volunteers living in Taiwan. According to Taiwan's male–female proportion and the population size in each county, 2474 individuals were randomly selected from all TWB participants to undergo DNAm analysis. In this study, we analyzed the DNAm data of these 2474 TWB participants. Written informed consent and blood and urine samples were provided by each individual when participating in the study. The TWB collected lifestyle factors such

**Table 6** Results of six DNAm markers in mediating the associations between current smokers' pack-years and four health outcomes

Outcome <sup>a</sup>	Total effect	95% confidence interval		p-value	Sample size		
FG	<b>0.0114</b>	<b>0.0072</b>	<b>0.0156</b>	<b>1.4E-07</b>	2452		
HbA <sub>1c</sub>	<b>0.0063</b>	<b>0.0021</b>	<b>0.0105</b>	<b>0.003</b>	2452		
FEV1	- 0.0006	- 0.0052	0.0039	0.782	1563		
FVC	- 0.0011	- 0.0048	0.0026	0.565	1563		
Outcome	Mediator	Mediation effect <sup>b</sup>	95% confidence interval		FDR <sup>c</sup>	Proportion mediated <sup>d</sup>	Sample size
FG	HannumEAA	0.0003	- 4.0E-05	0.0007	0.194	2.6	2445
HbA <sub>1c</sub>		0.0003	- 4.0E-05	0.0007	0.178	4.8	2445
FEV1		0.0001	- 0.0002	0.0006	0.599	- 16.7	1556
FVC		- 0.0001	- 0.0004	0.0002	0.767	9.1	1556
FG	PhenoEAA	0.0003	- 2.0E-05	0.0008	0.145	2.6	2450
HbA <sub>1c</sub>		<b>0.0006</b>	<b>0.0002</b>	<b>0.0011</b>	<b>0.016</b>	9.5	2450
FEV1		- 0.0001	- 0.0006	0.0004	0.767	16.7	1561
FVC		0.0000	- 0.0004	0.0004	0.957	0.0	1561
FG	GrimEAA	<b>0.0064</b>	<b>0.0038</b>	<b>0.0094</b>	<b>&lt; 0.001</b>	56.1	2448
HbA <sub>1c</sub>		<b>0.0076</b>	<b>0.0049</b>	<b>0.0109</b>	<b>&lt; 0.001</b>	120.6	2448
FEV1		0.0001	- 0.0024	0.0025	0.957	- 16.7	1560
FVC		- 0.0008	- 0.0029	0.0013	0.622	72.7	1560
FG	DNAm-based pack-years	0.0043	- 0.0003	0.0091	0.145	37.7	2399
HbA <sub>1c</sub>		0.0044	0.0001	0.0087	0.120	69.8	2399
FEV1		0.0013	- 0.0031	0.0062	0.744	- 216.7	1526
FVC		0.0034	0.0002	0.0072	0.108	- 309.1	1526
FG	DNAm PAI-1	<b>0.0029</b>	<b>0.0017</b>	<b>0.0043</b>	<b>&lt; 0.001</b>	25.4	2452
HbA <sub>1c</sub>		<b>0.0027</b>	<b>0.0016</b>	<b>0.0041</b>	<b>&lt; 0.001</b>	42.9	2452
FEV1		- 0.0004	- 0.0012	0.0002	0.332	66.7	1563
FVC		- <b>0.0007</b>	- <b>0.0015</b>	- <b>0.0002</b>	<b>0.024</b>	63.6	1563
FG	DunedinPACE	<b>0.0014</b>	<b>0.0003</b>	<b>0.0026</b>	<b>0.045</b>	12.3	2451
HbA <sub>1c</sub>		<b>0.0024</b>	<b>0.0014</b>	<b>0.0037</b>	<b>&lt; 0.001</b>	38.1	2451
FEV1		0.0001	- 0.0010	0.0012	0.927	- 16.7	1562
FVC		- 0.0005	- 0.0015	0.0004	0.501	45.5	1562

<sup>a</sup> Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

<sup>b</sup> Mediation effect is the effect of X (current smokers' pack-years) on Y (health outcome) through the mediator (DNAm marker)

<sup>c</sup> FDR (false discovery rate) represents the FDR-adjusted p-value. Significant effects with FDR < 0.05 are marked in bold

<sup>d</sup> Proportion mediated was calculated by dividing the mediation effect by the total effect

as smoking status through a face-to-face interview with health care professionals. Serum glucose and HbA<sub>1c</sub> were measured with a Hitachi LST008 analyzer (Hitachi High-Technologies, Tokyo, Japan) and a Trinity Biotech Premier Hb9210 analyzer (Bray, Ireland/Kansas City, MO) after a fast of at least 6 h. FEV1 and FVC were measured through pulmonary function test machines. FVC is the amount of air forcibly exhaled from the lungs after taking the deepest breath, and FEV1 is the volume of air exhaled in the first second [46].

#### Calculations of EAA (M)

The TWB quantified the blood DNAm levels of the 2474 participants through the Illumina Infinium MethylationEPIC BeadChip covering ~860,000 CpGs. The quality controls and normalization process for the DNAm data have been described elsewhere [35]. The DNAm Age Calculator from Horvath's laboratory, <https://dnamage.genetics.ucla.edu/new>, was used to calculate four measures of EAA and the two components of GrimEAA: HannumEAA [26] (column "AgeAccelerationResidualHannum" of the DNAm Age Calculator output), IEAA [27] (column

**Table 7** Results of six DNAm markers in mediating the associations between years since smoking cessation and four health outcomes

Outcome <sup>a</sup>	Total effect	95% confidence interval		p-value	Sample size		
FG	0.0012	- 0.0123	0.0147	0.859	300		
HbA <sub>1c</sub>	- 0.0018	- 0.0150	0.0115	0.795	300		
FEV1	0.0005	- 0.0143	0.0154	0.947	193		
FVC	- 0.0036	- 0.0147	0.0074	0.516	193		
Outcome	Mediator	Mediation effect <sup>b</sup>	95% confidence interval		FDR <sup>c</sup>	Proportion mediated <sup>d</sup> (%)	Sample size
FG	HannumEAA	- 0.0008	- 0.0032	0.0007	0.566	- 66.7	299
HbA <sub>1c</sub>		- 0.0011	- 0.0037	0.0008	0.447	61.1	299
FEV1		0.0002	- 0.0028	0.0035	0.957	40.0	192
FVC		0.0014	- 0.0001	0.0041	0.191	- 38.9	192
FG	PhenoEAA	- 0.0017	- 0.0049	0.0006	0.273	- 141.7	300
HbA <sub>1c</sub>		- 0.0020	- 0.0055	0.0003	0.194	111.1	300
FEV1		<b>0.0056</b>	<b>0.0013</b>	<b>0.0113</b>	<b>0.026</b>	1120.0	193
FVC		0.0033	0.0006	0.0068	0.055	- 91.7	193
FG	GrimEAA	- 0.0023	- 0.0073	0.0021	0.501	- 191.7	300
HbA <sub>1c</sub>		- 0.0041	- 0.0087	4.0E-05	0.138	227.8	300
FEV1		0.0035	- 0.0018	0.0093	0.336	700.0	193
FVC		<b>0.0050</b>	<b>0.0014</b>	<b>0.0096</b>	<b>0.026</b>	- 138.9	193
FG	DNAm-based pack-years	- 0.0005	- 0.0068	0.0055	0.934	- 41.7	298
HbA <sub>1c</sub>		- 0.0021	- 0.0074	0.0033	0.599	116.7	298
FEV1		- 0.0003	- 0.0076	0.0072	0.957	- 60.0	192
FVC		- 0.0028	- 0.0081	0.0029	0.501	77.8	192
FG	DNAm PAI-1	- 0.0009	- 0.0034	0.0005	0.443	- 75.0	300
HbA <sub>1c</sub>		- 0.0015	- 0.0049	0.0007	0.345	83.3	300
FEV1		- 0.0004	- 0.0029	0.0017	0.767	- 80.0	193
FVC		0.0007	- 0.0006	0.0028	0.512	- 19.4	193
FG	DunedinPACE	- 0.0001	- 0.0015	0.0010	0.934	- 8.3	300
HbA <sub>1c</sub>		- 0.0008	- 0.0030	0.0008	0.501	44.4	300
FEV1		0.0008	- 0.0009	0.0038	0.599	160.0	193
FVC		0.0014	- 0.0011	0.0048	0.467	- 38.9	193

<sup>a</sup> Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

<sup>b</sup> Mediation effect is the effect of X (years since smoking cessation) on Y (health outcome) through the mediator (EAA)

<sup>c</sup> FDR (false discovery rate) represents the FDR-adjusted p-value. Significant effects with FDR < 0.05 are marked in bold

<sup>d</sup> Proportion mediated was calculated by dividing the mediation effect by the total effect

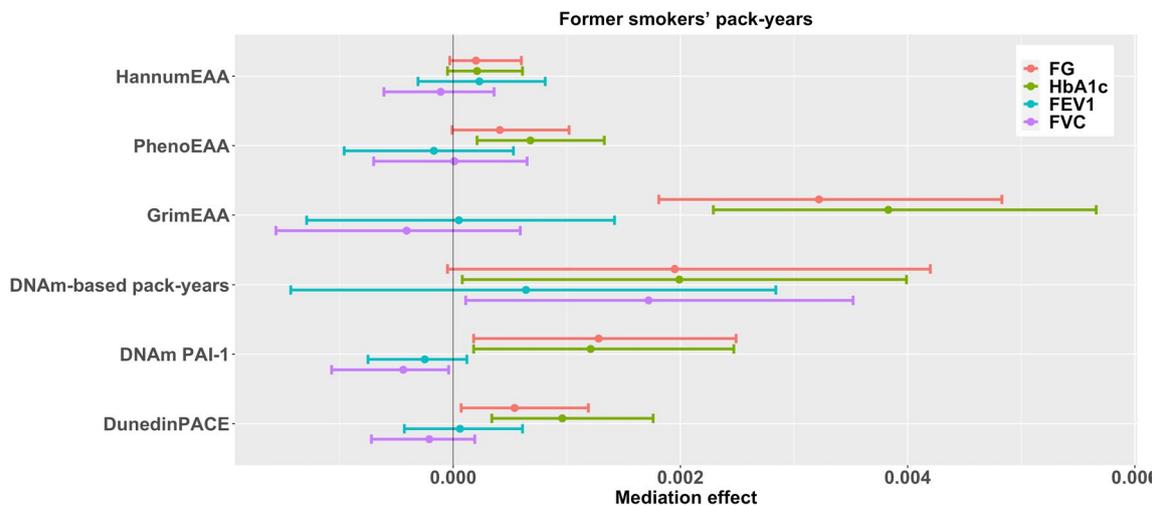
“IEAA”), PhenoEAA [28] (column “AgeAccelPheno”), GrimEAA [29, 30] (column “AgeAccelGrim2”), DNAm-based smoking pack-years [29, 30] (column “DNAm-PACKYRS”), and DNAm PAI-1 levels [29, 30] (column “DNAmPAI1”). DunedinPACE [31] was obtained from the R package DunedinPACE (<https://github.com/danbeisky/DunedinPACE>).

**Statistical analysis**

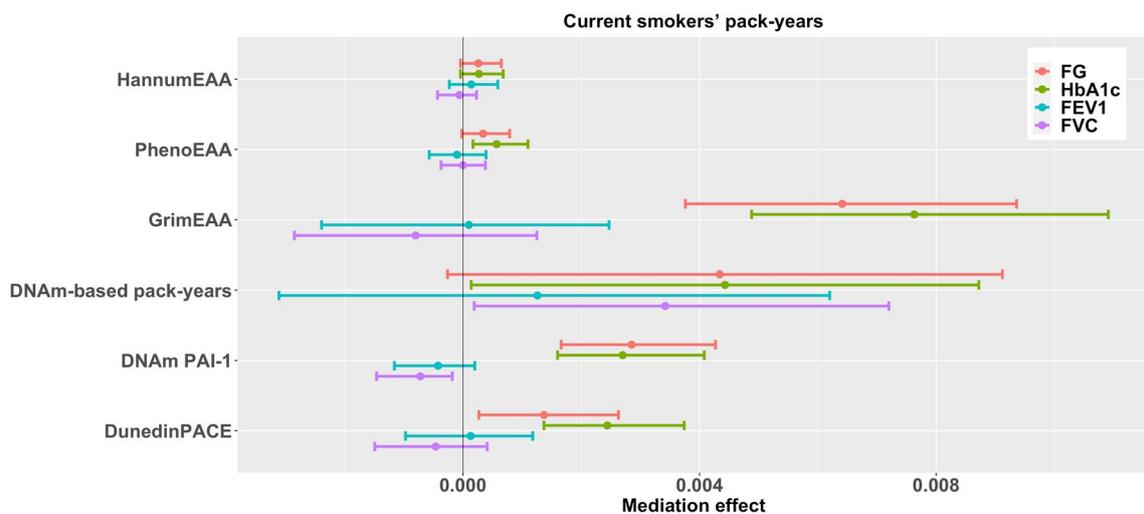
We performed 120 mediation analyses to investigate which of the six epigenetic measures (IEAA not included according to Table 2) plays the best role as a mediator between the five smoking variables and four health

outcomes (6×5 × 4=120). Two right-skewed health outcomes, FG and HbA<sub>1c</sub>, were natural-log transformed to meet the normality assumption for linear models. Furthermore, all four health outcomes were standardized as z-scores. To be considered a plausible mediator, two requirements should be met: (1) the association between smoking (X) and EAA (M); and (2) the association between EAA (M) and the health outcome (Y) while controlling for the smoking variable (X).

Former smoking status (Table 3 and Fig. 1) and current smoking status (Table 4 and Fig. 2) were investigated through the following two models (M model and Y model):



**Fig. 3** 95% Confidence intervals of the effects of six DNAm markers in mediating the associations between former smokers' pack-years and four health outcomes. Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores



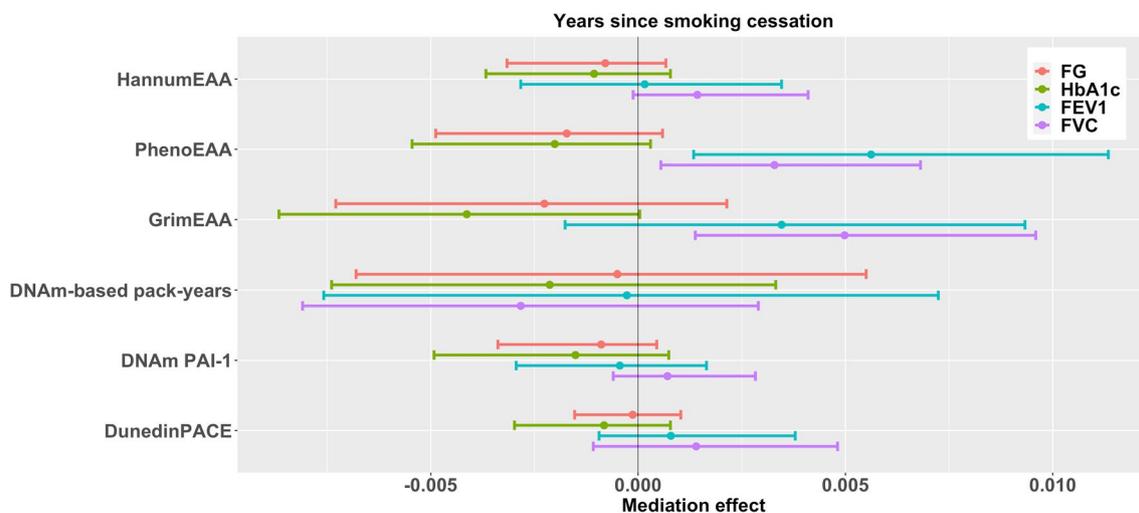
**Fig. 4** 95% Confidence intervals of the effects of six DNAm markers in mediating the associations between current smokers' pack-years and four health outcomes. Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

$$EAA(M) = \beta_{0M} + \beta_{X_1M}FS + \beta_{X_2M}CS + \beta'_{CM}Covariates + \varepsilon_M \tag{1}$$

$$Healthoutcome(Y) = \beta_{0Y} + \beta_{X_1Y}FS + \beta_{X_2Y}CS + \beta_{MY}EAA(M) + \beta'_{CY}Covariates + \varepsilon_Y \tag{2}$$

where *FS* and *CS* are dummy-coded variables for former smokers and current smokers (with non-smokers treated as the reference group); **Covariates** included age,

sex, BMI, drinking status, regular exercise status, educational attainment, and five cell-type proportions (B lymphocytes, CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, monocytes, and natural killer cells).  $\varepsilon_M$  and  $\varepsilon_Y$  are random error terms of the M model and Y model, respectively. The direct effect of current smoking on the health outcome was calculated by  $\hat{\beta}_{X_2Y}$ , and the mediation effect was estimated by  $\hat{\beta}_{X_2M} \cdot \hat{\beta}_{MY}$ , where  $\hat{\beta}$  represents the regression coefficients estimated from models (1) and (2). The total effect of current smoking on the health outcome was the sum of



**Fig. 5** 95% Confidence intervals of the effects of six DNAm markers in mediating the associations between years since smoking cessation and four health outcomes. Before mediation analysis, FG and HbA<sub>1c</sub> were natural-log transformed and then standardized as z-scores; FEV1 and FVC were standardized as z-scores

the direct and mediation effects, i.e.,  $\widehat{\beta}_{X_2Y} + \widehat{\beta}_{X_2M} \cdot \widehat{\beta}_{MY}$ . The effect of former smoking was calculated similarly.

Regarding the analysis for former smokers’ pack-years (Table 5 and Fig. 3) and current smokers’ pack-years (Table 6 and Fig. 4), we considered the following two models:

$$\begin{aligned}
 EAA(M) = & \beta_{0M} + \beta_{X_1M}FPY \\
 & + \beta_{X_2M}CPY \\
 & + \beta'_{CM}Covariates + \varepsilon_M;
 \end{aligned}
 \tag{3}$$

$$\begin{aligned}
 Healthoutcome(Y) = & \beta_{0Y} + \beta_{X_1Y}FPY + \beta_{X_2Y}CPY \\
 & + \beta_{MY}EAA(M) \\
 & + \beta'_{CY}Covariates + \varepsilon_Y,
 \end{aligned}
 \tag{4}$$

where *FPY* and *CPY* are variables coding former smokers’ pack-years and current smokers’ pack-years (with non-smokers coded as 0 in these two variables); **Covariates** have been described for models (1)–(2).

For analysis of years since smoking cessation (Table 7 and Fig. 5), only former smokers were analyzed by the following two models:

$$\begin{aligned}
 EAA(M) = & \beta_{0M} + \beta_{XM}YSC + \beta'_{CM}Covariates + \varepsilon_M;
 \end{aligned}
 \tag{5}$$

$$\begin{aligned}
 Healthoutcome(Y) = & \beta_{0Y} + \beta_{XY}YSC \\
 & + \beta_{MY}EAA(M) \\
 & + \beta'_{CY}Covariates + \varepsilon_Y,
 \end{aligned}
 \tag{6}$$

where *YSC* is former smokers’ years since smoking cessation; **Covariates** included former smokers’ pack-years and the covariates described for models (1)–(2).

All statistical analyses were conducted with R software (version 4.2.3) [47], where the mediation analyses were performed using the R package **mediation** [48, 49]. We used the nonparametric bootstrap for the variance estimation by assigning the “boot = TRUE” option in the “mediate” function. The number of simulations was set at 2000 by setting “sims = 2000”. After obtaining the 120 *p*-values for the testing hypothesis: *H*<sub>0</sub>: mediation effect is 0 versus *H*<sub>1</sub>: mediation effect is not 0, we performed the Benjamini–Hochberg FDR control [50]. Mediation effects with FDR < 0.05 were considered statistically significant.

We calculated variance inflation factor (VIF) values to check multicollinearity. A VIF value larger than 5 is usually considered a threat of multicollinearity [51]. No multicollinearity among the explanatory variables was detected in models (1)–(6) or models in Table 2, for the analyses of 2474 individuals (largest VIF = 2.51) and 1243 male participants (largest VIF = 2.26).

**Abbreviations**

COPD	Chronic obstructive pulmonary disease
CpGs	Cytosine-phosphate-guanine sites
DNAm	DNA methylation
EAA	Epigenetic age acceleration
EWAS	Epigenome-wide association studies
FDR	False discovery rate
FEV1	Forced expiratory volume in 1.0 s
FG	Fasting glucose
FVC	Forced vital capacity

HbA <sub>1c</sub>	Hemoglobin A1C
HRS	Health and Retirement Study
IEAA	Intrinsic epigenetic age acceleration
PAI-1	Plasminogen activation inhibitor 1
SD	Standard deviation
TWB	Taiwan Biobank
VIF	Variance inflation factor

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-023-01512-x>.

**Additional file 1: Table S1.** Basic characteristics of the 173,807 Taiwan Biobank (TWB) participants stratified by smoking status. **Table S2.** (Based on 173,807 Taiwan Biobank participants) Results of regressing health outcomes on smoking variables while adjusting for six covariates (chronological age (in years), sex (male vs. female), BMI (in kg/m<sup>2</sup>), drinking status (yes vs. no), performing regular exercise (yes vs. no), and educational attainment (an integer ranging from 1 to 7)). **Table S3.** (Only in males) Basic characteristics of the 1243 male participants stratified by smoking status. **Table S4.** (Only in males) Results of regressing each DNAm marker on smoking variables (based on models 1, 3, and 5). **Table S5.** (Only in males) Results of six DNAm markers in mediating the associations between former smoking status and four health outcomes. **Table S6.** (Only in males) Results of six DNAm markers in mediating the associations between current smoking status and four health outcomes. **Table S7.** (Only in males) Results of six DNAm markers in mediating the associations between former smokers' pack-years and four health outcomes. **Table S8.** (Only in males) Results of six DNAm markers in mediating the associations between current smokers' pack-years and four health outcomes. **Table S9.** (Only in males) Results of six DNAm markers in mediating the associations between years since smoking cessation and four health outcomes. **Figure S1.** Boxplots of the seven DNA methylation (DNAm) markers. **Figure S2.** (Both males and females) Heatmap of Pearson's correlation coefficients among variables. The numbers in cells are Pearson's correlation coefficients between pairwise variables. **Figure S3.** (Only in males) Heatmap of Pearson's correlation coefficients among variables. The numbers in cells are Pearson's correlation coefficients between pairwise variables. **Figure S4.** (Only in males) 95% confidence intervals of the effects of six DNAm markers in mediating the associations between former smoking status and four health outcomes. **Figure S5.** (Only in males) 95% confidence intervals of the effects of six DNAm markers in mediating the associations between current smoking status and four health outcomes. **Figure S6.** (Only in males) 95% confidence intervals of the effects of six DNAm markers in mediating the associations between former smokers' pack-years and four health outcomes. **Figure S7.** (Only in males) 95% confidence intervals of the effects of six DNAm markers in mediating the associations between current smokers' pack-years and four health outcomes. **Figure S8.** (Only in males) 95% confidence intervals of the effects of six DNAm markers in mediating the associations between years since smoking cessation and four health outcomes. **Figure S9.** (Both males and females) The original p-values (without FDR adjustment) of the mediation effects. **Figure S10.** (Only in males) The original p-values (without FDR adjustment) of the mediation effects.

## Acknowledgements

The authors would like to thank the Editor and anonymous reviewers for their insightful and constructive comments, and the Taiwan Biobank for approving our application to access the data.

## Author contributions

This study was conceived, designed, and interpreted by XYC and WYL. XYC performed statistical analyses and drafted the manuscript. WYL supervised this study and substantively revised the manuscript. Both authors read and approved the final manuscript.

## Funding

This study was supported by the National Science and Technology Council of Taiwan (Grant Number 111-2314-B-002-099 to Wan-Yu Lin).

## Availability of data and materials

The datasets used and analyzed during the current study are available from <https://www.twbiobank.org.tw/>.

## Declarations

### Ethics approval and consent to participate

TWB was approved by the Ethics and Governance Council of Taiwan Biobank and the Institutional Review Board on Biomedical Science Research/IRB-BM, Academia Sinica, Taiwan. Written informed consent was obtained from each participant in accordance with institutional requirements and the principles of the Declaration of Helsinki. Our use of the TWB research data was approved by TWB on February 18, 2020 (application number: TWBR10810-07). This study further received approval from the Research Ethics Committee of the National Taiwan University Hospital (NTUH-REC no. 201805050RINB).

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Room 501, No. 17, Xu-Zhou Road, Taipei 100, Taiwan. <sup>2</sup>Master of Public Health Degree Program, College of Public Health, National Taiwan University, Taipei, Taiwan. <sup>3</sup>Department of Public Health, College of Public Health, National Taiwan University, Taipei, Taiwan.

Received: 8 February 2023 Accepted: 24 May 2023

Published online: 03 June 2023

## References

- Choi D, Ota S, Watanuki S. Does cigarette smoking relieve stress? Evidence from the event-related potential (ERP). *Int J Psychophysiol*. 2015;98(3 Pt 1):470–6.
- Nesbitt PD. Smoking, physiological arousal, and emotional response. *J Pers Soc Psychol*. 1973;25(1):137–44.
- Pomerleau CS, Pomerleau OF. The effects of a psychological stressor on cigarette smoking and subsequent behavioral and physiological responses. *Psychophysiology*. 1987;24(3):278–85.
- Perkins KA, Grobe JE, Fonte C, Breus M. "Paradoxical" effects of smoking on subjective stress versus cardiovascular arousal in males and females. *Pharmacol Biochem Behav*. 1992;42(2):301–11.
- Poureslami I, Shum J, Aran N, Tregobov N. Chinese- and English-speaking adult current smokers' perspectives on smoking and culturally and linguistically appropriate cessation: a qualitative analysis. *Addict Sci Clin Pract*. 2020;15(1):23.
- Munzel T, Hahad O, Kuntic M, Keaney JF, Deanfield JE, Daiber A. Effects of tobacco cigarettes, e-cigarettes, and waterpipe smoking on endothelial function and clinical outcomes. *Eur Heart J*. 2020;41(41):4057–70.
- Mo R, Zhang J, Chen Y, Ding Y. Nicotine promotes chronic obstructive pulmonary disease via inducing pyroptosis activation in bronchial epithelial cells. *Mol Med Rep*. 2022;25(3):1–8.
- Śliwińska-Mossoń M, Milnerowicz H. The impact of smoking on the development of diabetes and its complications. *Diab Vasc Dis Res*. 2017;14(4):265–76.
- Maisonneuve P, Lowenfels AB, Mullhaupt B, Cavallini G, Lankisch PG, Andersen JR, Dimagno EP, Andren-Sandberg A, Domellof L, Frulloni L, et al. Cigarette smoking accelerates progression of alcoholic chronic pancreatitis. *Gut*. 2005;54(4):510–4.
- Ye X, Lu G, Huai J, Ding J. Impact of smoking on the risk of pancreatitis: a systematic review and meta-analysis. *PLoS ONE*. 2015;10(4):e0124075.

11. Campagna D, Alamo A, Di Pino A, Russo C, Calogero AE, Purrello F, Polosa R. Smoking and diabetes: dangerous liaisons and confusing relationships. *Diabetol Metab Syndr*. 2019;11:85.
12. Maddatu J, Anderson-Baucum E, Evans-Molina C. Smoking and the risk of type 2 diabetes. *Transl Res*. 2017;184:101–7.
13. Tantisuwat A, Thaveeratitham P. Effects of smoking on chest expansion, lung function, and respiratory muscle strength of youths. *J Phys Ther Sci*. 2014;26(2):167–70.
14. Darden ME, Kaestner R. Smoking, selection, and medical care expenditures. *J Risk Uncertain*. 2022;64(3):251–85.
15. Bednarek M, Gorecka D, Wielgomas J, Czajkowska-Malinowska M, Regula J, Mieszko-Filipczyk G, Jasionowicz M, Bijata-Bronisz R, Lempicka-Jastrzebska M, Czajkowski M, et al. Smokers with airway obstruction are more likely to quit smoking. *Thorax*. 2006;61(10):869–73.
16. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Latkja E, Adamski J, Peters A, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS ONE*. 2013;8(5): e63812.
17. Besingi W, Johansson A. Smoke-related DNA methylation changes in the etiology of human disease. *Hum Mol Genet*. 2014;23(9):2290–7.
18. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet*. 2011;88(4):450–7.
19. Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, Belvisi MG, Brown R, Vineis P, Flanagan JM. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet*. 2013;22(5):843–51.
20. Allione A, Marcon F, Fiorito G, Guarrera S, Siniscalchi E, Zijno A, Crebelli R, Matullo G. Novel epigenetic changes unveiled by monozygotic twins discordant for smoking habits. *PLoS ONE*. 2015;10(6): e0128265.
21. Dogan MV, Shields B, Cutrona C, Gao L, Gibbons FX, Simons R, Monick M, Brody GH, Tan K, Beach SR, et al. The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. *BMC Genom*. 2014;15:151.
22. Richmond RC, Simpkin AJ, Woodward G, Gaunt TR, Lyttleton O, McArdle WL, Ring SM, Smith AD, Timpson NJ, Tilling K, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet*. 2015;24(8):2201–17.
23. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, Huang Z, Hoyo C, Middttun O, Cupul-Uicab LA, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2012;120(10):1425–31.
24. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenet*. 2015;7:113.
25. Gupta R, van Dongen J, Fu Y, Abdellaoui A, Tyndale RF, Velagapudi V, Boomsma DI, Korhonen T, Kaprio J, Loukola A, et al. Epigenome-wide association study of serum cotinine in current smokers reveals novel genetically driven loci. *Clin Epigenet*. 2019;11(1):1.
26. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M, Fan JB, Gao Y, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49(2):359–67.
27. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
28. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*. 2018;10(4):573–91.
29. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*. 2019;11(2):303–27.
30. Lu AT, Binder AM, Zhang J, Yan Q, Reiner AP, Cox SR, Corley J, Harris SE, Kuo PL, Moore AZ, et al. DNA methylation GrimAge version 2. *Aging (Albany NY)*. 2022;14(23):9484–549.
31. Belsky DW, Caspi A, Corcoran DL, Sugden K, Poulton R, Arseneault L, Baccarelli A, Chamarti K, Gao X, Hannon E, et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife*. 2022;11: e73420.
32. Levine ME. Assessment of epigenetic clocks as biomarkers of aging in basic and population research. *J Gerontol A Biol*. 2020;75(3):463–5.
33. Rezwani FI, Imboden M, Amaral AFS, Wielscher M, Jeong A, Triebner K, Real FG, Jarvelin MR, Jarvis D, Probst-Hensch NM, et al. Association of adult lung function with accelerated biological aging. *Aging (Albany NY)*. 2020;12(1):518–42.
34. Wang C, Just A, Heiss J, Coull BA, Hou L, Zheng Y, Sparrow D, Vokonas PS, Baccarelli A, Schwartz J. Biomarkers of aging and lung function in the normative aging study. *Aging (Albany NY)*. 2020;12(12):11942–66.
35. Lin WY. Epigenetic clocks derived from western samples differentially reflect Taiwanese health outcomes. *Front Genet*. 2023;14:1089819.
36. Klopach ET, Carroll JE, Cole SW, Seeman TE, Crimmins EM. Lifetime exposure to smoking, epigenetic aging, and morbidity and mortality in older adults. *Clin Epigenet*. 2022;14(1):72.
37. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform*. 2012;13(1):86.
38. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol*. 2014;15(2):1–9.
39. Baron RM, Kenny DA. The moderator mediator variable distinction in social psychological-research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173–82.
40. Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychol Methods*. 2002;7(4):422–45.
41. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr*. 2009;76(4):408–20.
42. Luo JH, Rossouw J, Tong E, Giovino GA, Lee CC, Chen C, Ockene JK, Qi LH, Margolis KL. Smoking and diabetes: does the increased risk ever go away? *Am J Epidemiol*. 2013;178(6):937–45.
43. Oelsner EC, Balte PP, Bhatt SP, Cassano PA, Couper D, Folsom AR, Freedman ND, Jacobs DR Jr, Kalhan R, Mathew AR, et al. Lung function decline in former smokers and low-intensity current smokers: a secondary data analysis of the NHLBI Pooled Cohorts Study. *Lancet Respir Med*. 2020;8(1):34–44.
44. Belsky DW, Caspi A, Arseneault L, Baccarelli A, Corcoran DL, Gao X, Hannon E, Harrington HL, Rasmussen LJ, Houts R, et al. Quantification of the pace of biological aging in humans through a blood test, the Dunedin-PoAm DNA methylation algorithm. *Elife*. 2020;9:e54870.
45. Klopach ET, Crimmins EM, Cole SW, Seeman TE, Carroll JE. Accelerated epigenetic aging mediates link between adverse childhood experiences and depressive symptoms in older adults: results from the Health and Retirement Study. *Ssm-Popul Hlth*. 2022;17:101071.
46. Labaki WW, Rosenberg SR. Chronic obstructive pulmonary disease. *Ann Intern Med*. 2020;173(3):ITC17–32.
47. The R Project for Statistical Computing. <https://www.r-project.org/>
48. Imai K, Tingley D, Yamamoto T. Experimental designs for identifying causal mechanisms. *J R Stat Soc A Stat Soc*. 2013;176(1):5–51.
49. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R package for causal mediation analysis. *J Stat Softw*. 2014;59(5):1–38.
50. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57(1):289–300.
51. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiology (Sunnyvale)*. 2016;6(2):227.

## Publisher's Note

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