

Early-life exposure to secondhand smoke and polycyclic aromatic hydrocarbons: associations with epigenetic aging among children

Sneha S. Patil¹, Anne K. Bozack², Ajay Pillarisetti³, Smita S. Patil⁴, Kyle Steenland⁵, Kalpana Balakrishnan¹, Lance A. Waller⁶, Shirin Jabbarzadeh⁶, Jennifer Peel⁷, Thomas F. Clasen⁵, Dana Boyd Barr⁵, Naveen Puttaswamy^{1,†}, Andres Cardenas^{8,9,*}

¹Department of Environmental Health Engineering, ICMR Center for Advanced Research on Air Quality, Climate and Health, Sri Ramachandra Institute for Higher Education and Research (Deemed University), Chennai, TN 600116, India

²Department of Population Health, NYU Grossman School of Medicine, New York, NY 10016, United States

³Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94704, United States

⁴General Physician, Private Practitioner, Pune, MH 411006, India

⁵Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, United States

⁶Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322, United States

⁷Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO 80523, United States

⁸Department of Epidemiology and Population Health, Stanford University, Stanford, CA 94305, United States

⁹Department of Pediatrics, Stanford University, Stanford, CA 94305, United States

*Corresponding author. Department of Epidemiology and Population Health, Stanford University, Stanford, CA 94305, USA. E-mail: andresca@stanford.edu

[†]These authors have contributed equally to this work and share senior authorship.

Abstract

Epigenetic clocks are valuable tools for assessing biological age. Exposure to secondhand smoke (SHS) and polycyclic aromatic hydrocarbons (PAHs) has been linked with epigenetic age deviation (EAD) in adults. However, associations in children remain largely unexplored. We investigated relationships between exposure—spanning prenatal, early postnatal, current (36–60 months, time of sample collection) periods—and EAD in preschool-aged children. DNA methylation was measured in buccal cells from 43 children (mean age: 4.1 years) in the CAries Risk from exposure to Environmental tobacco Smoke—Household Air Pollution Intervention Network cohort. SHS exposure was assessed using urinary cotinine and 3-hydroxycotinine (3-HC), while PAHs exposure was assessed via urinary 2-naphthol (2-NAP) and 1-hydroxypyrene (1-PYR). Nicotine equivalents were calculated as molar sum of cotinine + 3-HC. EAD was estimated using Horvath, Skin&Blood, Pediatric Buccal Epigenetic (PedBE), and DNA methylation-based telomere length (DNAmTL) clocks. Associations between exposures and EAD were evaluated using linear regression, adjusting for study site, mother's education, child's body mass index z-score, age, and sex. Chronological age significantly correlated with all epigenetic clocks ($r = 0.37$ – 0.78) and with DNAmTL estimator ($r = -0.38$). Current SHS exposure, as measured by urinary cotinine, was associated with Horvath EAD ($B = 0.23$ years, $P = .05$). Similarly, current 3-HC and nicotine equivalents were positively associated with PedBE EAD ($B = 0.08$ years, $P = .04$; $B = 0.11$ years, $P = .05$, respectively). Among PAH metabolites, current 2-NAP levels were associated with principal component-based Skin&Blood EAD ($B = 0.21$ years, $P = .02$) and PedBE EAD ($B = 0.13$ years, $P = .05$). Findings suggest early-life SHS and PAHs exposure may contribute to accelerated epigenetic aging in children.

Keywords: epigenetics; children; epigenetic aging; secondhand smoke; polycyclic aromatic hydrocarbons

Introduction

Environmental exposures account for 70%–90% of the global disease burden and nearly a quarter of all deaths, disproportionately affecting children under five and those in low- and middle-income countries (LMICs) [1]. These exposures are diverse and pervasive, with many pollutants capable of crossing the placenta and disrupting fetal cellular programming. The Developmental Origins of Health and Disease hypothesis posits that early-life environmental insults can trigger long-term physiological alterations, heightening disease susceptibility later in life [2].

Among modifiable environmental exposures, secondhand smoke (SHS) is a ubiquitous public health issue [3]. Despite a decline in smoking rates over the past half-century, 37% of the world's population is still exposed to SHS, with a higher prevalence in children, and health risks exacerbated by racial and economic disparities [3, 4]. SHS contains over 7000 chemicals, including 69 known carcinogens [5, 6]. Of particular concern are polycyclic aromatic hydrocarbons (PAHs), a class of chemicals generated from the incomplete combustion of tobacco, charred food, coal, and wood [7]. Epidemiological and clinical studies have corroborated associations between SHS [5, 6] and PAHs

Received 17 June 2025; revised 8 October 2025; accepted 12 October 2025

© The Author(s) 2025. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons

Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com

Table 1. Characteristics of CARES-HAPIN children included in the analyses (2023–2025, Tamil Nadu, India)

Characteristic	Overall (N = 43)
Children's demographic information	
Age, mean (SD), years	4.12 (0.53)
Sex	
Boy	21 (48.8%)
Girl	22 (51.2%)
BMI, mean (SD)	13.48 (1.95)
BMI z-score, mean (SD)	−1.65 (1.65)
Parent's demographic information	
Mother's age, mean (SD), y	20.44 (3.53)
Mother's education	
No formal education	12 (27.9%)
Primary school completed	16 (37.2%)
Secondary school or college completed	15 (34.9%)

[8, 9] exposure and inflammation, oxidative stress, epigenetic modifications, and downstream adverse health outcomes, underscoring the need for reliable biomarkers to predict disease risk and elucidate the underlying mechanisms.

Epigenetic modifications, particularly DNA methylation, represent a promising and persistent biomarker of exposure that may reflect a mechanistic link between SHS, PAHs, and lagged health effects. DNA methylation, primarily occurring at cytosine residues upstream of guanine (CpG sites), is influenced by environmental factors from the intrauterine period onward [10], with prenatal smoking having a strong signature in umbilical cord blood DNA methylation [11]. Growing evidence highlights the roles of direct DNA damage and epigenetic mechanisms in SHS and PAHs-related diseases [12, 13].

In recent years, epigenetic clocks derived from DNA methylation patterns have emerged as powerful machine-learning tools for estimating biological or chronological age and predicting age-related disease and overall health. The difference between epigenetic age (estimated via DNA methylation) and chronological age, known as epigenetic age deviation (EAD), offers a novel lens to evaluate health status and the impact of early-life exposures on aging trajectories [14]. In children, greater EAD has been linked with multiple detrimental health outcomes, such as asthma, obesity, cognitive impairment, cancer, and other morbidities—conditions robustly associated with SHS and PAHs exposure [5, 6, 15–18].

Although EAD has been linked to environmental exposures such as air pollution, adult tobacco smoke, and cadmium, studies have focused on the prenatal period or adult populations, with sparse and inconsistent evidence among exposed children [15–21]. Furthermore, prior work has largely examined single exposures and relied on blood-derived DNA methylation, leaving accessible, non-invasive tissues like buccal cells underexplored. To address these gaps, our study investigates epigenetic age biomarkers in buccal cells among children exposed to SHS and PAHs across prenatal, early-life, and current periods. We hypothesize that, similar to adult smoking, early-childhood exposure to SHS and PAHs would accelerate epigenetic biomarkers of aging.

Results

Study population

Table 1 presents a summary of participant characteristics (N = 43). This sample was designed to reflect the sociodemographic and environmental diversity of the larger cohort. The mean chrono-

logical age of children at buccal DNA collection was 4.12 years (SD = 0.53). There were 21 male (48.8%) and 22 female (51.2%) children. The mean BMI of the children was 13.48 (SD = 1.95). Geometric means of biomarkers (μg/g creatinine; 95% CI) are presented in [Supplementary Table S1](#).

Epigenetic age

Scatter plots of chronological age and epigenetic age are shown in [Fig. 1](#). The Pediatric Buccal Epigenetic (PedBE) clock exhibited the highest correlation with chronological age ($r = 0.77$, $P < .001$) and the lowest mean absolute error (median absolute error (MAE) = 0.4 years). The principal component (PC) versions of the Horvath clocks also performed well, with the PC Horvath and PC Skin&Blood clocks exhibiting correlations of $r = 0.78$ and $r = 0.76$, respectively (both $P < .001$). In comparison, the CpG-based Horvath and Skin&Blood clocks showed weaker correlations ($r = 0.37$, $P = .013$ and $r = 0.62$, $P < .001$, respectively). The DNA methylation telomere length (DNAmTL) and PC DNAmTL estimators, as expected, were negatively correlated with age ($r = -0.38$, $P = .013$ and $r = -0.35$, $P = .021$, respectively), indicating shortening of telomere length with increasing age.

SHS exposure and EAD

In adjusted models ([Table 2](#)), a two-fold increase in average prenatal cotinine concentration was inversely associated with Horvath EAD [B (95% CI) = −0.22 years (−0.44, −0.01), $P = .04$]. Conversely, a doubling of average postnatal and current cotinine concentrations was positively associated with Horvath EAD, with effect estimates of 0.23 years (95% CI = −0.00, 0.44, $P = .05$) and 0.18 years (95% CI = −0.00, 0.35, $P = .05$), respectively. Although associations with other clocks were not statistically significant, effect estimates for Skin&Blood and PedBE EAD were consistently positive but small. Average prenatal and average postnatal 3-hydroxycotinine (3-HC) concentrations showed no meaningful relationships. However, a two-fold rise in current 3-HC concentration was significantly associated with both longer DNAmTL (B (95% CI) = 0.02 kbs (0.00, 0.04), $P = .02$), and increased PedBE EAD (B (95% CI) = 0.08 years (0.00, 0.16), $P = .04$).

A doubling of average postnatal urinary nicotine equivalents showed a marginal 0.22 years increase in Horvath EAD (95% CI = −0.02, 0.45, $P = .08$). Similarly, an increase by a factor of two in current nicotine equivalents concentration corresponded with longer DNAmTL (B (95% CI) = 0.03 kbs (0.00, 0.05), $P = .03$) and accelerated epigenetic age via PedBE (B (95% CI) = 0.11 years (0.00, 0.21), $P = .05$). Associations with other clocks for nicotine equivalents were null.

In sensitivity analyses adjusting for estimated buccal cell-type proportions, effect estimates for cotinine and Horvath were slightly attenuated but retained their direction (prenatal: B (95% CI) = −0.22 years (−0.43, −0.01), $P = .04$; postnatal: B (95% CI) = 0.17 years (−0.02, 0.34), $P = .08$; current: B (95% CI) = 0.20 years (−0.03, 0.44), $P = .09$; [Supplementary Table S2](#)). Associations for 3-HC and nicotine equivalents with PedBE and DNAmTL were robust to cell-type adjustment, albeit with slightly reduced precision.

PAHs exposure and EAD

Average prenatal 2-naphthol (2-NAP) associations with EAD were positive but imprecise. In contrast, a doubling of current 2-NAP concentration was positively associated with PC Skin&Blood EAD (B (95% CI) = 0.21 years (0.03, 0.39), $P = .02$) and showed a borderline positive association with PedBE EAD (B (95% CI) = 0.13 years (0.00, 0.26), $P = .05$) ([Table 3](#)). For 1-hydroxypyrene

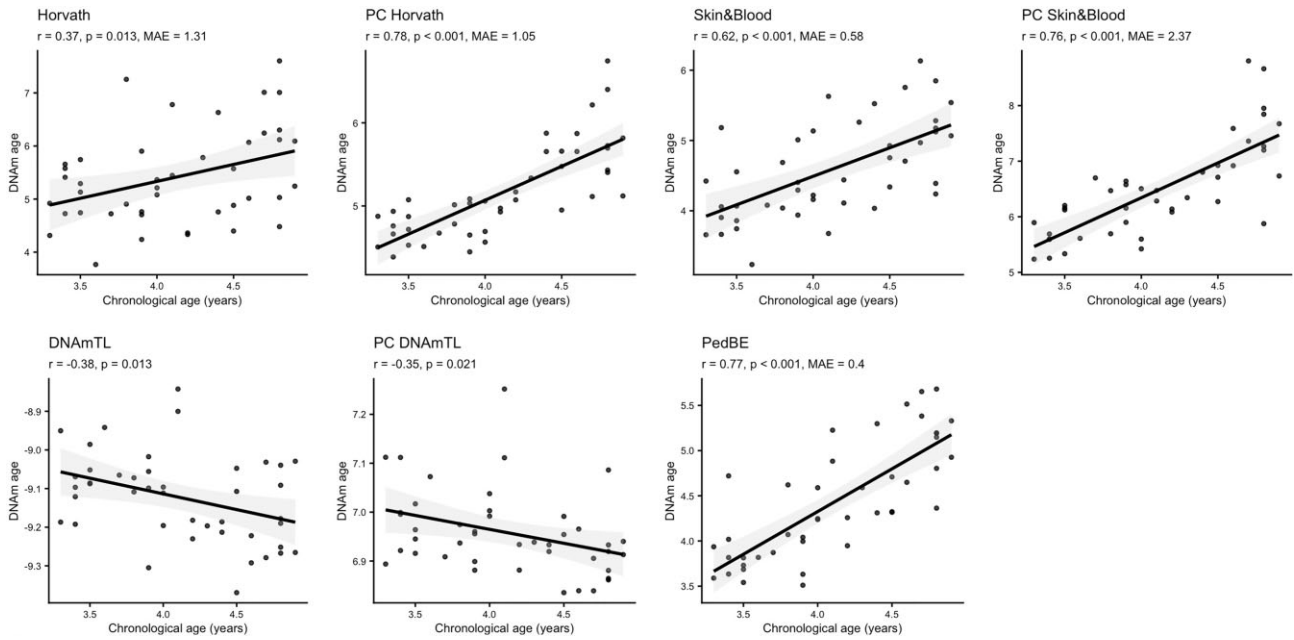


Figure 1. Scatter plots of chronological age and DNA methylation age (CARES-HAPIN, 2023–2025, Tamil Nadu, India). EA was estimated using the Horvath, Skin&Blood, DNAmTL, and PedBE clocks in buccal cells ($N = 43$). The upper panels show the Pearson's correlation coefficient (r), P -value (p), and MAE between each pair of variables. Linear trendline and 95% CI plotted as a solid line and shaded area.

Table 2. Associations of urinary nicotine biomarkers with EAD in CARES-HAPIN children (2023–2025, Tamil Nadu, India)

Outcome	Average prenatal		Average postnatal		Current	
	B (95% CI)	P-value	B (95% CI)	P-value	B (95% CI)	P-value
Cotinine						
Horvath	−0.22 (−0.44, −0.01)	0.04	0.18 (−0.00, 0.35)	0.05	0.23 (−0.00, 0.44)	0.05
PC Horvath	−0.05 (−0.13, 0.04)	0.26	0.06 (−0.01, 0.12)	0.09	0.02 (−0.07, 0.11)	0.64
Skin&Blood	−0.06 (−0.21, 0.09)	0.40	0.05 (−0.07, 0.17)	0.41	0.06 (−0.09, 0.21)	0.44
PC Skin&Blood	−0.06 (−0.21, 0.09)	0.40	0.06 (−0.06, 0.18)	0.35	0.04 (−0.11, 0.20)	0.58
DNAmTL	0.01 (−0.01, 0.04)	0.23	0.01 (−0.01, 0.03)	0.53	0.00 (−0.02, 0.03)	0.76
PC DNAmTL	0.00 (−0.02, 0.02)	0.88	0.01 (−0.01, 0.02)	0.36	0.01 (−0.01, 0.03)	0.29
PedBE	−0.02 (−0.13, 0.08)	0.67	0.04 (−0.05, 0.13)	0.36	0.06 (−0.04, 0.17)	0.25
3-hydroxycotinine						
Horvath	−0.09 (−0.40, 0.21)	0.55	0.14 (−0.09, 0.38)	0.22	0.05 (−0.13, 0.22)	0.60
PC Horvath	−0.05 (−0.16, 0.07)	0.40	0.05 (−0.04, 0.14)	0.24	−0.05 (−0.11, 0.02)	0.13
Skin&Blood	−0.01 (−0.20, 0.19)	0.96	0.03 (−0.12, 0.19)	0.67	−0.01 (−0.12, 0.11)	0.92
PC Skin&Blood	0.09 (−0.11, 0.29)	0.37	0.02 (−0.14, 0.17)	0.83	0.05 (−0.07, 0.16)	0.40
DNAmTL	0.01 (−0.03, 0.04)	0.75	0.01 (−0.02, 0.03)	0.48	0.02 (0.00, 0.04)	0.02
PC DNAmTL	−0.02 (−0.05, 0.01)	0.12	0.01 (−0.01, 0.03)	0.27	0.01 (−0.01, 0.02)	0.21
PedBE	−0.01 (−0.16, 0.13)	0.86	0.03 (−0.09, 0.14)	0.65	0.08 (0.00, 0.16)	0.04
Nicotine equivalents (cotinine + 3-HC)						
Horvath	−0.23 (−0.53, 0.08)	0.15	0.22 (−0.02, 0.45)	0.08	0.09 (−0.14, 0.32)	0.41
PC Horvath	−0.06 (−0.17, 0.06)	0.33	0.06 (−0.03, 0.15)	0.16	−0.06 (−0.14, 0.03)	0.17
Skin&Blood	−0.03 (−0.23, 0.18)	0.77	0.06 (−0.11, 0.22)	0.50	0.00 (−0.15, 0.15)	0.98
PC Skin&Blood	0.04 (−0.17, 0.24)	0.76	0.05 (−0.11, 0.21)	0.51	0.07 (−0.08, 0.22)	0.35
DNAmTL	0.01 (−0.02, 0.04)	0.51	0.01 (−0.02, 0.04)	0.43	0.03 (0.00, 0.05)	0.03
PC DNAmTL	−0.01 (−0.04, 0.01)	0.30	0.01 (−0.01, 0.03)	0.32	0.01 (−0.01, 0.03)	0.20
PedBE	−0.03 (−0.18, 0.12)	0.67	0.05 (−0.07, 0.17)	0.39	0.11 (0.00, 0.21)	0.05

Results from linear models adjusted for study site, mother's education, child's body mass index z-score, age, and sex. Highlighted P -values denote significance.

(1-PYR), an increase by a factor of two in average postnatal levels was suggestively associated with Skin&Blood EAD (B (95% CI) = 0.19 years (−0.01, 0.40), $P = .07$). Additionally, every two-fold rise in current concentration showed a marginal association with PC Skin&Blood EAD (B (95% CI) = 0.12 years (−0.02, 0.26), $P = .08$).

In cell-type-adjusted sensitivity analyses, the association between a two-fold increase in current 2-NAP levels and PC Skin&Blood remained statistically significant (B (95% CI) = 0.23 years (0.05, 0.42), $P = .01$), while the association with PedBE acceleration became less precise (B (95% CI) = 0.10 years (−0.03, 0.23), $P = .13$). Associations between 1-PYR levels and Skin&Blood

Table 3. Associations of urinary PAH metabolites with EAD in CARES-HAPIN children (2023–2025, Tamil Nadu, India)

Outcome	Average prenatal		Average postnatal		Current	
	B (95% CI)	P-value	B (95% CI)	P-value	B (95% CI)	P-value
2-naphthol						
Horvath	−0.09 (−0.39, 0.21)	0.55	0.08 (−0.20, 0.36)	0.57	0.15 (−0.14, 0.44)	0.29
PC Horvath	−0.01 (−0.13, 0.10)	0.81	0.02 (−0.08, 0.13)	0.66	−0.03 (−0.14, 0.08)	0.61
Skin&Blood	−0.01 (−0.21, 0.19)	0.91	−0.02 (−0.21, 0.16)	0.80	0.05 (−0.14, 0.24)	0.58
PC Skin&Blood	0.00 (−0.19, 0.20)	0.96	0.02 (−0.16, 0.21)	0.79	0.21 (0.03, 0.39)	0.02
DNAmTL	−0.02 (−0.05, 0.01)	0.28	−0.00 (−0.03, 0.03)	0.95	0.02 (−0.01, 0.05)	0.19
PC DNAmTL	−0.02 (−0.04, 0.01)	0.14	−0.01 (−0.03, 0.02)	0.72	0.01 (−0.01, 0.04)	0.33
PedBE	−0.05 (−0.19, 0.09)	0.47	−0.05 (−0.18, 0.08)	0.46	0.13 (0.00, 0.26)	0.05
1-hydroxypyrene						
Horvath	−0.06 (−0.32, 0.20)	0.63	0.23 (−0.10, 0.55)	0.17	0.06 (−0.17, 0.27)	0.62
PC Horvath	0.04 (−0.06, 0.14)	0.40	0.05 (−0.08, 0.17)	0.46	−0.03 (−0.11, 0.05)	0.41
Skin&Blood	−0.01 (−0.18, 0.16)	0.93	0.19 (−0.01, 0.40)	0.07	0.01 (−0.13, 0.16)	0.86
PC Skin&Blood	0.05 (−0.12, 0.22)	0.55	0.04 (−0.18, 0.26)	0.70	0.12 (−0.02, 0.26)	0.08
DNAmTL	−0.01 (−0.04, 0.01)	0.38	0.00 (−0.03, 0.04)	0.90	0.01 (−0.02, 0.03)	0.64
PC DNAmTL	−0.01 (−0.03, 0.01)	0.43	0.00 (−0.03, 0.03)	0.89	0.01 (−0.01, 0.03)	0.52
PedBE	−0.01 (−0.13, 0.11)	0.88	0.07 (−0.09, 0.22)	0.37	0.07 (−0.03, 0.17)	0.16

Results from linear models adjusted for study site, mother's education, child's body mass index z-score, age, and sex. Highlighted P-values denote significance.

remained directionally positive but did not meet statistical significance after cell-type adjustment (average postnatal: B (95% CI) = 0.19 years (−0.02, 0.39), $P = .08$; [Supplementary Table S3](#)).

Discussion

Our study is among the first to apply multiple DNA methylation-based aging biomarkers, including the Horvath, Skin&Blood, PedBE, and DNAmTL clocks, to buccal samples from children in LMIC settings born to non-smoking mothers. Notably, epigenetic age estimates were significantly correlated with chronological age, reinforcing the applicability of the clocks to buccal-derived DNA in the pediatric South Asian population. Our results, particularly for current exposures, support the hypothesis that tobacco-related toxicants contribute to epigenetic aging signatures in early childhood. Average postnatal levels and current cotinine concentrations were significantly associated with Horvath EAD, while current 3-HC and nicotine equivalents correlated with PedBE EAD (B (95% CI) = 0.11 (0.00, 0.21), $P = .05$) and longer DNAmTL estimates (B (95% CI) = 0.03 (0.00, 0.05), $P = .03$). Current 2-NAP levels were linked to increased PC Skin&Blood EAD. These novel findings contribute to the evidence on how early-life environmental exposures shape epigenetic aging trajectories in children.

While previous studies have shown that prenatal maternal smoking is associated with accelerated epigenetic aging in cord blood, at birth [15–17], we observed a deceleration in buccal epigenetic aging with average prenatal SHS and PAHs exposure. This divergence may reflect tissue-specific responses, the timing of sample collection, or *in-utero* adaptations that mitigate toxicant effects or influence developmental maturity throughout childhood, though the long-term implications remain to be completely elucidated. Discrepancies in exposure source and intensity (active smoking vs. secondhand exposure) may also underlie these inconsistencies. By focusing on SHS and PAHs exposure during the early childhood period, our study expands the literature, emphasizing the role of pediatric epigenetic clocks in environmental insults beyond the prenatal window [16, 17].

Elevated average postnatal levels and current cotinine concentrations were positively associated with Horvath EAD in concordance with Simpkin et al. and de Prado-Bert et al. [15, 17]. Addi-

tionally, current levels of 3-HC and nicotine equivalents were related to PedBE acceleration and longer DNAmTL, suggesting that SHS exposure may influence multiple molecular markers of aging. Although an increase in DNAmTL may seem counterintuitive, telomere dynamics in early childhood are complex. Telomeres shorten most rapidly from birth to around age three, followed by relative stabilization; however, lengthening has also been documented in some children, potentially reflecting telomerase activity in hematopoietic stem cells or shifts in leukocyte cell composition during growth [22]. Thus, increases in DNAmTL in our study may reflect altered cellular turnover or compensatory proliferative responses in buccal epithelial tissue rather than simple telomere attrition. Prior research suggests that DNAmTL captures broader aspects of cellular aging independent of telomere attrition and telomerase activity [23]. Increased DNAmTL has also been reported in certain cancer contexts, further underscoring the complexity of interpreting these epigenetic markers in relation to environmental exposures. Parallel associations were discerned between PAH metabolites: 2-NAP and postnatal 1-PYR and acceleration in Skin&Blood and PedBE clocks, implicating the putative role of PAHs as epigenetic modulators during critical developmental periods.

The application of epigenetic clocks in children remains fairly novel. While first-generation clocks primarily predict chronological age, newer clocks trained on phenotypic aging or mortality aim to capture broader biological aging processes but are mostly trained in adults [24]. Among those tailored for pediatric populations, the PedBE clock, trained on 1032 individuals aged 2–20 years, demonstrated the strongest correlation with age in our data as expected. The Horvath and Skin&Blood clocks also yielded significant, albeit weaker correlations with chronological age ($P < .001$), likely due to their multi-tissue training sets, which included buccal cells and the broader age range represented in their training data. We excluded second-generation clocks developed from adult phenotypes, as these may be less relevant to early-life biology. Despite promising patterns, it remains unclear how variation in pediatric epigenetic clocks translates to long-term health outcomes. The limited CpG overlap among clocks implies that each may index distinct biological domains, such as transcriptional regulation, telomere biology, cellular activation, and senes-

cence [25]. While significant associations were observed with the traditional Horvath clock, these findings should be interpreted with caution. The PC Horvath clock, which mitigates missing CpGs and technical variability, did not demonstrate comparable significance, suggesting that the Horvath clock associations may in part reflect measurement noise in our relatively small sample. Replication in larger cohorts is imperative to substantiate these findings. Persistent associations observed across multiple clocks in our study suggest that early-life SHS and PAHs exposure may broadly perturb interconnected aging processes, potentially via oxidative stress and immune-inflammatory dysregulation, although the precise molecular intermediaries remain elusive.

Congruous with our *a priori* hypothesis, adjustment for buccal cell-type composition modestly attenuated associations with Horvath EAD, suggesting partial confounding via immune or inflammatory-mediated pathways. Conversely, associations with PedBE and Skin&Blood clocks persisted, indicating either reduced sensitivity to tissue heterogeneity or enhanced specificity to exposure-related molecular alterations. Together, these results accentuate the complexity of environmentally driven epigenetic aging during early childhood. While buccal DNA methylation profiling offers a non-invasive and scalable alternative for pediatric research, the lack of clocks calibrated exclusively for buccal tissue and interindividual cellular heterogeneity may introduce measurement imprecision for certain EAD metrics. Nonetheless, our work demonstrates the feasibility of conducting high-dimensional epigenetic analyses using buccal-derived DNA from young children in LMIC settings. This represents a distinctive strength given the logistical, infrastructural, and ethical challenges of pediatric biospecimen collection in such contexts.

Additional strengths include the detailed assessment of environmental exposures from the prenatal to early-life period, key windows of rapid biological plasticity, and use of sensitivity analyses that collectively augmented the robustness of our findings. However, some limitations merit consideration. The relatively small sample size for DNA methylation analyses may have limited our ability to detect small to modest associations. Despite covariate adjustment, residual confounding from unmeasured or imprecisely measured variables remains possible. Moreover, though we assessed exposures across early-life, the full complexity of the exposome, or all potential contributors to EAD was not captured. Our analyses focused exclusively on established epigenetic clocks, which are calculated using DNA methylation levels at CpG sites representing a small subset of the epigenome. Consequently, other exposure-related methylation changes may not have been detected. While this approach was appropriate for our hypothesis-driven focus on age- and development-related pathways, it does not encompass the full epigenome. Future epigenome-wide analyses with larger sample sizes and appropriate replication may provide broader insights. Finally, although the LMIC context amplifies the public health relevance of these findings, it may limit generalizability to other populations with different environmental or genetic backgrounds. Validation in other cohorts is warranted.

In conclusion, our study identified early-life exposure to SHS and PAHs as modifiable determinants of EAD in children, adding to growing evidence that epigenetic clocks serve as sensitive biomarkers of environmental exposures in early-life spanning from the prenatal period to childhood. While some associations were modest and did not attain significance, the consistency in directional patterns, particularly for current exposures, implicates these environmental pollutants as potential modulators of biological aging processes during crucial developmental windows. Future prospective investigations are needed to determine whether

such early-life exposures develop epigenetic alterations over time and, in turn, whether these changes confer increased disease risk in later life.

Methods

Study population

The CARES Risk from exposure to Environmental tobacco Smoke (CARES) study is nested within the Indian cohort of the multi-country Household Air Pollution Intervention Network (HAPIN) trial. Details of the study design and participants are provided elsewhere [26–28]. Between April and September 2023, 712 children aged 3–5 years were followed up in Tamil Nadu. Of these, 48 children were selected for DNA methylation analyses using a stratified multistage cluster sampling design. Stratification was based on district (Villupuram and Nagapattinam) and caregiver-reported environmental tobacco smoke (ETS) exposure. Within each stratum, clusters—defined as villages or administrative blocks—were randomly drawn to ensure representation across the ETS exposure spectrum.

The study protocol was approved by the institutional review boards at Stanford School of Medicine and Sri Ramachandra Institute of Higher Education and Research (IEC-NI22/JUL/83/82). Written informed consent was obtained from the parents or legal guardians of all participants. Trained interviewers administered structured questionnaires to collect data on sociodemographic factors, maternal education, dietary diversity, and child characteristics. The child's body mass was measured using a foot-to-foot bioimpedance scale, and height was measured in triplicate using a stadiometer and averaged. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. ETS exposure was assessed via caregiver-reported household smoking behaviours. Ambiguous or missing responses were resolved through follow-up calls.

Exposure assessment

Urine samples were collected from pregnant women in each trimester and from children at postnatal visits (3, 6, 12, and 36–60 months). Biomarker concentrations were normalized to urinary creatinine to account for urine dilution [29], which was measured by photometric detection using the Jaffé method [30].

Urinary SHS biomarkers

Urinary concentrations of cotinine and 3-HC were measured using liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) [29, 31]. A 200 μ L urine sample was spiked with an isotopically labelled internal standard, mixed with cold acetone, refrigerated, and centrifuged to precipitate proteins. The supernatant was processed, and 10 μ L was injected for analysis. All analyses were performed without glucuronidase deconjugation in order to quantify free (unconjugated) analytes, which represent the biologically active fraction most likely to mediate epigenetic effects. Quality control samples, including blanks, were analyzed concurrently. The limit of detection (LOD) for both cotinine and 3-HC was 0.05 ng/mL, with mean accuracy ranging from 81.7% to 88.7% (Supplementary Table S4). Nicotine equivalents were calculated as the sum of cotinine + 3-HC.

Urinary PAH metabolites

Urinary levels of monohydroxy PAH metabolites (2-NAP and 1-PYR) were measured using a validated method [32]. A 2 mL urine aliquot was incubated with β -glucuronidase (from *Helix pomatia*,

type HP-2) and sodium acetate buffer (pH 5.5). Analyses were conducted using high-performance liquid chromatography (Shimadzu LC-2010A HPLC system, Kyoto, Japan) equipped with RF-10AXL fluorescence detector. Chromatographic separation was performed on a Gemini C6-Phenyl RP column. The mobile phase consisted of phosphate buffer and acetonitrile. The LODs were 0.02 ng/mL for 2-NAP and 0.06 ng/mL for 1-PYR, with mean accuracies of 102.8% and 104.6%, respectively. Quality control samples were included for each batch ([Supplementary Table S4](#)).

Microarray quantification of DNA methylation

Buccal cell samples were collected using Oragene kits (OC-175, DNA Genotek), and genomic DNA was extracted using QIAGEN kits per manufacturer's protocols. DNA yield was assessed using the NanoDrop 2000 Spectrophotometer (Thermo Scientific). Bisulfite modification of genomic DNA was performed using the Zymo DNA Methylation kit (Zymo Research, Irvine, CA, USA), and samples were hybridized to Infinium MethylationEPIC v2.0 BeadChip arrays (Illumina, San Diego, CA, USA).

Samples were randomized across chips prior to measurement. DNA methylation data were imported into R (version 4.4.1) and subsequently preprocessed using the SeSAMe package [33]. Non-linear dye-bias correction and background subtraction were performed using the normal-exponential out-of-band (noob) method [34]. Sample-level QC included dropping samples based on low intensity (<10.5) ($n = 0$), inadequate bisulfite conversion ($n = 0$), or sex mismatches ($n = 5$). *ComBat* from the *sva* package was used to remove residual batch effects from the sample row in the BeadChip [35]. Density distributions were visualized at each processing step, and principal component analysis was conducted to evaluate associations of technical, biological, and measured traits with global DNA methylation variation. The DNA methylation status for each CpG locus was calculated as a Beta value, ranging from zero to one representing % methylation. Buccal cell proportions—epithelial, fibroblast, and immune cells, were estimated using *EpiDISH* [36].

Calculation of epigenetic aging biomarkers

Epigenetic clocks were calculated using the *methscore* [37] function in the *ENmix* package [38], including Horvath, Skin&Blood, PedBE, and DNAmTL due to their performance in children [23, 39–41]. Due to the missingness of clock CpGs on the EPIC v2.0 BeadChip, we also calculated PC-based versions of the Horvath clock, Skin&Blood, and DNAmTL [42]. This approach is less influenced by the missingness of individual CpGs and array-type differences. We estimated the outcome of interest, EAD, as the residuals from regressing each clock on the participants' chronological age. A positive EAD indicates advanced biological aging relative to chronological age, and a negative age-adjusted DNAmTL reflects shorter telomere length. [Supplementary Table S5](#) summarizes these biomarkers, their health outcomes predicted, and the tissues used in training.

Statistical analyses

Descriptive statistics were calculated as mean and standard deviations (SDs) for continuous variables and percentages for categorical variables. To meet assumptions of normality and heteroscedasticity, we used $\log_2$ -transformed exposure concentrations in the analyses, where a one-unit increase reflects a doubling of the original levels. Average (i.e. cumulative mean) prenatal and postnatal levels were calculated as the mean of concentrations measured across the three trimesters and at 3, 6, and 12 months,

respectively. Current exposure (36–60 months) reflects levels at sample collection. To evaluate the performance of the epigenetic clocks, we calculated Pearson correlation coefficients and the MAE between estimated epigenetic age and chronological age. Potential confounders were identified a priori and included study site (Villupuram or Nagapattinam), maternal education (no formal education, primary school completed, or secondary school/college completed), child's age- and sex-specific BMI z-score (derived from WHO growth standards), age (continuous), and biological sex. Associations between exposure and EAD were evaluated using linear regression models, adjusting for these covariates. Sensitivity analyses included additional adjustments for estimated cell-type proportions (epithelial, fibroblast, and immune cells). Results are reported as point estimates with 95% confidence intervals (CIs) and P-values from covariate-adjusted models.

Acknowledgements

We gratefully acknowledge the contributions of all HAPIN field staff, laboratory personnel, and participants.

Author contributions

Sneha S Patil (Conceptualization [equal], Formal Analysis [equal], Funding acquisition [equal], Methodology [equal], Visualization [equal], Writing – original draft [lead]), Anne Bozack (Formal Analysis [equal], Funding acquisition [equal], Methodology [equal], Visualization [equal], Writing – review & editing [equal]), Ajay Pillariseti (Writing – review & editing [equal]), Smita S Patil (Data curation [equal], Project administration [equal], Writing – review & editing [equal]), Smita S. Patil (Data curation [equal], Project administration [equal], Writing – review & editing [equal]), Kyle Steenland (Writing – review & editing [equal]), Kalpana Balakrishnan (Writing – review & editing [supporting]), Lance Waller (Data curation [equal], Writing – review & editing [supporting]), Shirin Jabbarzadeh (Data curation [equal], Writing – review & editing [supporting]), Jennifer Peel (Funding acquisition [equal], Writing – review & editing [supporting]), Thomas Clasen (Funding acquisition [equal], Writing – review & editing [equal]), Dana Boyd Barr (Methodology [equal], Writing – review & editing [equal]), Naveen Puttaswamy (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Supervision [equal], Writing – review & editing [equal]), and Andres Cardenas (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Supervision [equal], Writing – review & editing [equal]).

Supplementary data

Supplementary data is available at [EnvEpig](#) online.

Conflict of interest: The authors declare they have no competing interests.

Funding

This work was supported by the Fogarty International Center with participation from the National Institute of Dental and Craniofacial Research under grant D43TW010540. The HAPIN trial is funded by the US National Institutes of Health (cooperative agreement) under grant 1UM1HL134590 in collaboration with the Bill & Melinda Gates Foundation under grant OPP1131279. AKB and AC were partially supported by the National Institutes of Health grant R01ES031259.

Data availability

Data generated and analyzed during the current study are available from the CARES-HAPIN study team upon reasonable request and institutional review board approval.

References

- World Health Organization. 2022. *Preventing disease through healthy environments: A global assessment of the burden of disease from environmental risks* (vol. 2018, p. 17). World Health Organization; <https://www.who.int/publications/i/item/9789241565196> (10 January 2025, date last accessed).
- Barker DJP. The origins of the developmental origins theory. *J Intern Med* 2007;**261**:412–17. <https://doi.org/10.1111/j.1365-2796.2007.01809.x>
- Zhai C, Hu D, Yu G et al. Global, regional, and national deaths, disability-adjusted life years, years lived with disability, and years of life lost for the global disease burden attributable to second-hand smoke, 1990–2019: a systematic analysis for the Global Burden of Disease Study. *Sci Total Environ* 2023;**862**:160677.
- Mbulu L, Palipudi KM, Andes L et al. Secondhand smoke exposure at home among one billion children in 21 countries: findings from the Global Adult Tobacco Survey (GATS). *Tob Control* 2016;**25**:e95–e100. <https://doi.org/10.1136/tobaccocontrol-2015-052693>
- International Agency for Research on Cancer. Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum* 2004;**83**:1–1438.
- Surgeon General. 2014. The health consequences of smoking—50 years of progress: a report of the Surgeon General. *Department of Health and Human Services*. Public Health Service. US Office of the Surgeon General: Rockville. <http://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf> (10 January 2025, date last accessed).
- Agency for Toxic Substances and Disease Registry (ATSDR). 1995. *Toxicological Profile for Polycyclic Aromatic Hydrocarbons*. U.S. Department of Health and Human Services.
- Jr BF, BA R, Souza MCO et al. Polycyclic aromatic hydrocarbons (PAHs): updated aspects of their determination, kinetics in the human body, and toxicity. *J Toxicol Environ Health B Crit Rev* 2023;**26**:28–65.
- Langlois PH, Hoyt AT, Lupo PJ et al. Maternal occupational exposure to polycyclic aromatic hydrocarbons and risk of oral cleft-affected pregnancies. *The Cleft Palate Craniofacial Journal* 2013;**50**:337–46. <https://doi.org/10.1597/12-104>
- Millar D, Holliday R, Grigg G. Five not four: history and significance of the fifth base. In Beck, S, Olek, A (Eds.), *The Epigenome, Molecular Hide and Seek*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2003, 3–20. <https://doi.org/10.1002/3527601511>
- Joubert BR, Felix JF, Yousefi P et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am Hum Genet* 2016;**98**:680–96. <https://doi.org/10.1016/j.ajhg.2016.02.019>
- Baccarelli A, Bollati V. Epigenetics and environmental chemicals. *Curr Opin Pediatr* 2009;**21**:243–51. <https://doi.org/10.1097/MOP.0b013e32832925cc>
- Bakulski KM, Blostein F, London SJ. Linking prenatal environmental exposures to lifetime health with epigenome-wide association studies: state-of-the-science review and future recommendations. *Environ Health Perspect* 2023;**131**:126001. <https://doi.org/10.1289/EHP12956>
- Moqri M, Herzog C, Poganik JR. Biomarkers of Aging Consortium Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell* 2023;**186**:3758–75. <https://doi.org/10.1016/j.cell.2023.08.003>
- 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. <https://doi.org/10.1093/hmg/ddv456>
- Javed R, Chen W, Lin F et al. Infant's DNA methylation age at birth and epigenetic aging accelerators. *Biomed Res Int* 2016;**2016**:1. <https://doi.org/10.1155/2016/4515928>
- de Prado-Bert P, Ruiz-Arenas C, Vives-Usano M et al. The early-life exposome and epigenetic age acceleration in children. *Environ Int* 2021;**155**:106683. <https://doi.org/10.1016/j.envint.2021.106683>
- Fadadu RP, Bozack AK, Cardenas A. Chemical and climatic environmental exposures and epigenetic aging: a systematic review. *Environ Res* 2025;**274**:121347. <https://doi.org/10.1016/j.envres.2025.121347>
- Nwanaji-Enwerem JC, Colicino E, Trevisi L et al. Long-term ambient particle exposures and blood DNA methylation age: findings from the VA normative aging study. *Environ Epigenetics* 2016;**2**:dvw006. <https://doi.org/10.1093/eeep/dvw006>
- Martin EM, Fry RC. Environmental Influences on the epigenome: exposure-associated DNA methylation in human populations. *Annu Rev Public Health* 2018;**39**:309–33. <https://doi.org/10.1146/annurev-publhealth-040617-014629>
- Cardenas A, Ecker S, Fadadu RP et al. Epigenome-wide association study and epigenetic age acceleration associated with cigarette smoking among Costa Rican adults. *Sci Rep* 2022;**12**:4277. <https://doi.org/10.1038/s41598-022-08160-w>
- Cowell W, Tang D, Yu J et al. Telomere dynamics across the early life course: findings from a longitudinal study in children. *Psychoneuroendocrinology* 2021;**129**:105270. <https://doi.org/10.1016/j.psyneuen.2021.105270>
- Lu AT, Seeboth A, Tsai PC et al. DNA methylation-based estimator of telomere length. *Aging* 2019;**11**:5895–923. <https://doi.org/10.18632/aging.102173>
- Bergsma T□, Rogaeva E. DNA methylation clocks and their predictive capacity for aging phenotypes and healthspan. *Neuroscience Insights* 2020;**15**:2633105520942221. <https://doi.org/10.1177/2633105520942221>
- Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of aging. *Nat Rev Genet* 2018;**19**:371–84. <https://doi.org/10.1038/s41576-018-0004-3>
- Clasen T, Checkley W, Peel JL et al. Design and rationale of the HAPIN Study: a multicountry randomized controlled trial to assess the effect of liquefied petroleum gas stove and continuous fuel distribution. *Environ Health Perspect* 2020;**128**:47008. <https://doi.org/10.1289/EHP6407>
- Barr DB, Puttaswamy N, Jaacks LM et al. Design and rationale of the biomarker center of the Household Air Pollution Intervention Network (HAPIN) trial. *Environ Health Perspect* 2020;**128**:47010. <https://doi.org/10.1289/EHP5751>
- Patil SS, Puttaswamy N, Cardenas A et al. Protocol for CARES-HAPIN: an ambidirectional cohort study on exposure to environmental tobacco smoke and risk of early childhood caries. *BMJ Open* 2024;**14**:e083874. <https://doi.org/10.1136/bmjopen-2024-083874>
- Barr DB, Wilder LC, Caudill SP et al. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. *Environ Health Perspect* 2005;**113**:192–200. <https://doi.org/10.1289/ehp.7337>

30. Jaffe M. Über den Niederschlag, welchen Pikrinsäure in normalen Harn erzeugt und über eine neue Reaction des Kreatinins. *Z Physiol Chem* 1886;**10**:391–400.
31. McGuffey JE, Wei B, Bernert JT et al. Validation of an LC-MS/MS method for quantifying urinary nicotine, six nicotine metabolites, and the minor tobacco alkaloids—anatabine and anabasine—in smokers' urine. *PLoS One* 2014;**9**:e101816. <https://doi.org/10.1371/journal.pone.0101816>
32. Puttaswamy N, Saidam S, Rajendran G et al. Cross-validation of biomonitoring methods for polycyclic aromatic hydrocarbon metabolites in human urine: results from the formative phase of the Household Air Pollution Intervention Network (HAPIN) trial in India. *J Chromatogr B* 2020;**1154**:122284. <https://doi.org/10.1016/j.jchromb.2020.122284>
33. Zhou W, Jr TTJ, Laird PW et al. SeSAME: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res* 2018;**46**:e123.
34. Triche TJ, Weisenberger DJ, Van Den Berg D et al. Low-level processing of Illumina Infinium DNA methylation BeadArrays. *Nucleic Acids Res* 2013;**41**:e90. <https://doi.org/10.1093/nar/gkt090>
35. Leek JT, Johnson WE, Parker HS et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;**28**:882–83. <https://doi.org/10.1093/bioinformatics/bts034>
36. Zheng SC, Breeze CE, Beck S et al. EpiDISH web server: epigenetic dissection of intra-sample-heterogeneity with online GUI. *Bioinformatics* 2020;**36**:1950–51. <https://doi.org/10.1093/bioinformatics/btz833>
37. Xu Z, Niu L, Kresovich JK et al. methscore: a comprehensive R function for DNA methylation-based health predictors. *Bioinformatics* 2024;**40**:btac302. <https://doi.org/10.1093/bioinformatics/btac302>
38. Xu Z, Niu L, Taylor JA. The ENmix DNA methylation analysis pipeline for Illumina BeadChip and comparisons with seven other preprocessing pipelines. *Clinical Epigenetics* 2021;**13**:216. <https://doi.org/10.1186/s13148-021-01207-1>
39. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol* 2013;**14**:R115. <https://doi.org/10.1186/gb-2013-14-10-r115>
40. Horvath S, Oshima J, Martin GM et al. Epigenetic clock for skin and blood cells applied to Hutchinson-Gilford Progeria Syndrome and ex vivo studies. *Aging* 2018;**10**:1758–75. <https://doi.org/10.18632/aging.101508>
41. McEwen LM, O'Donnell KJ, McGill MG et al. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proc Natl Acad Sci USA* 2020;**117**:23329–35. <https://doi.org/10.1073/pnas.1820843116>
42. Higgins-Chen AT, Thrush KL, Wang Y et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nature Aging* 2022;**2**:644–61. <https://doi.org/10.1038/s43587-022-00248-2>