Polymethylation scores results

F blostein

9/30/2021

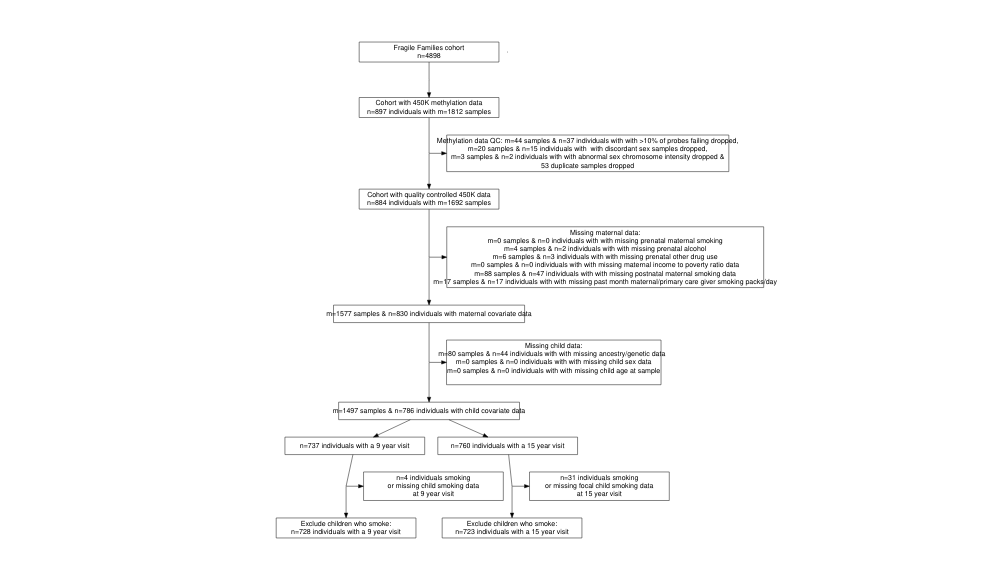
# Target journal and deadline

[Clinical epigenetics](https://www.biomedcentral.com/collections/epibio): Special issue on biomarkers; due date December 1st

"Epigenetic biomarkers have significantly contributed to improved understanding of the origins and progression of disease. Moreover, increasing evidence shows that epigenetic biomarkers have potential for personalized medicine. However, routine use of epigenetic biomarkers in in-vitro diagnostics (IVD) is lagging behind the discoveries of new biomarkers. In this thematic series of Clinical Epigenetics, we discuss current progress in development of epigenetic biomarkers for clinical use in common complex diseases of ageing and welcome both original research and reviews. The sections of the collection include:

* Types of epigenetic biomarkers and readiness for IVD use
  + DNA methylation as biomarker
  + Histone modification as biomarker
  + Upcoming field of circulating free nucleosomes as biomarker
  + Progress towards clinical use of biomarker
* *Biomarkers of environmental exposures and causality of the biomarker*
* Disease predisposition biomarkers
* Biomarkers used for disease detection
* Clinical disease prevention and management: predictive and prognostic
* Monitoring of chronic disease
* *Single versus multiple biomarkers and risk scores*
* Major stages of clinical validation of the biomarker
* Biomarker detection technologies suitable for IVD use and significance of bioinformatics
* Ethics and regulatory aspects of IVD epigenetic biomarkers

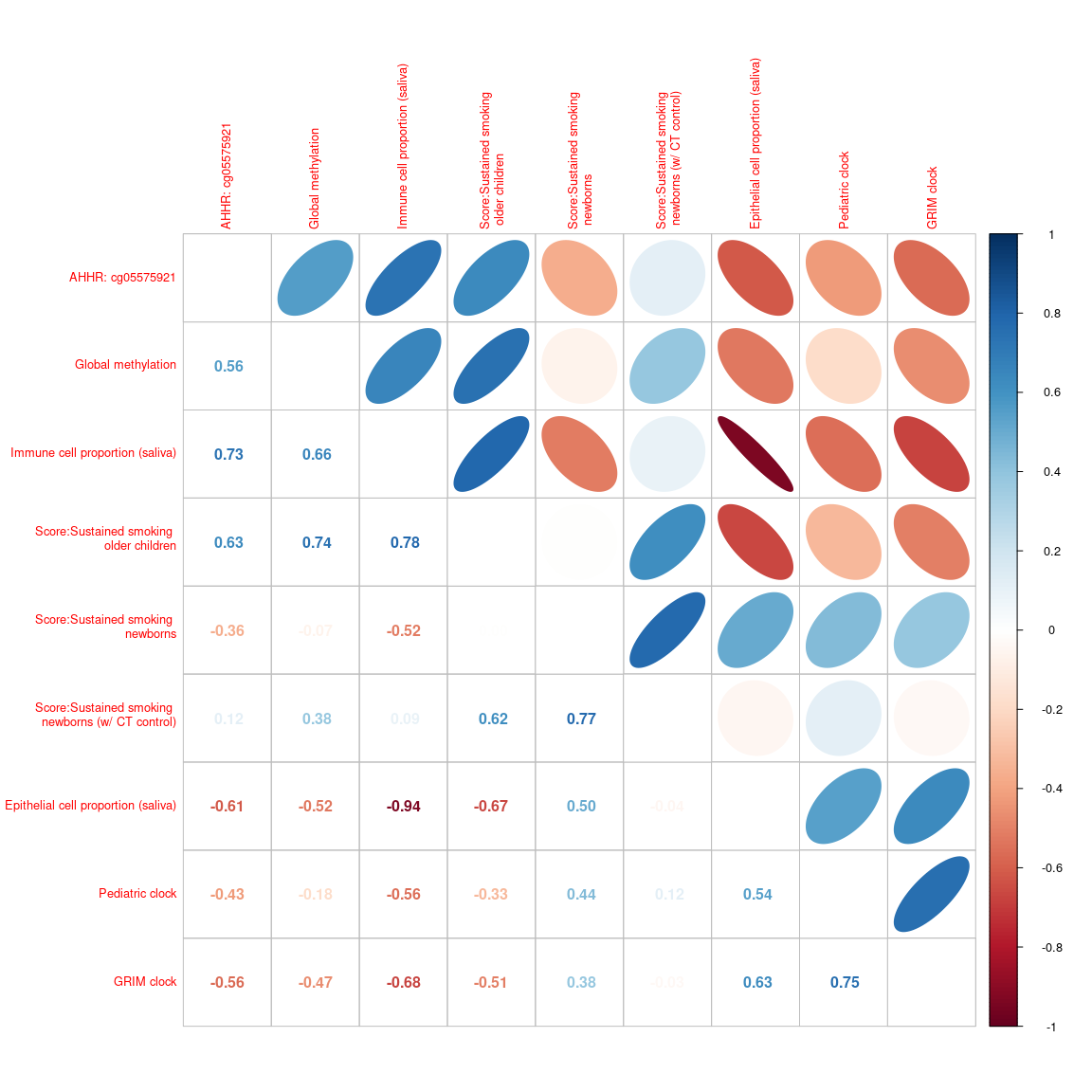
# Figure 1 - Selection of samples from the Fragile Families and Child Wellbeing (FFCW) study into this analysis. N represents the number of individuals at each step in the selection procedure, M represents the number of samples, as individuals with repeated measures can have more than one sample.



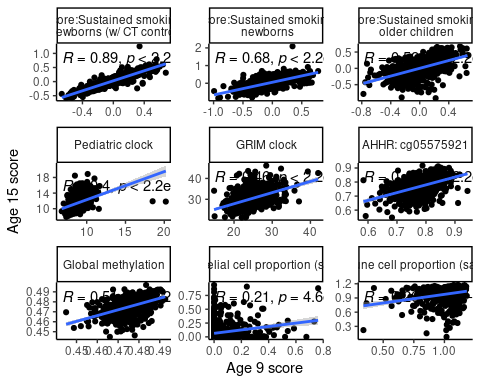
# Supplemental Table 1 - Descriptive statistics of excluded vs included samples with methylation data available from the Fragile and Families and Child Wellbeing study

| Characteristic | N | Final analysis group, N = 1,4511 | Not in final analysis group, N = 2411 | p-value2 |
| --- | --- | --- | --- | --- |
| Maternal prenatal smoking | 1,692 | 289 (20%) | 56 (23%) | 0.2 |
| Maternal prenatal alcohol use | 1,688 | 172 (12%) | 21 (8.9%) | 0.2 |
| Unknown |  | 0 | 4 |  |
| Maternal prenatal any drug use | 1,686 | 73 (5.0%) | 11 (4.7%) | 0.8 |
| Unknown |  | 0 | 6 |  |
| Constructed - Poverty ratio - mother's household income/poverty threshold | 1,692 | 2.21 (2.42) | 2.39 (2.56) | 0.3 |
| Any postnatal maternal smoking when child age 1 or 5 | 1,604 |  |  | 0.14 |
| Maternal smoking at age 1 or age 5 |  | 574 (40%) | 70 (46%) |  |
| No maternal smoking at age 1 and 5 |  | 877 (60%) | 83 (54%) |  |
| Unknown |  | 0 | 88 |  |
| Maternal/primary care giver smoking in month prior to visit (pks/day) | 1,674 |  |  | 0.8 |
| Less than pack a day |  | 332 (23%) | 49 (22%) |  |
| No smoking |  | 1,018 (70%) | 156 (70%) |  |
| Pack or more a day |  | 101 (7.0%) | 18 (8.1%) |  |
| Unknown |  | 0 | 18 |  |
| Ancestry categorization from child principal components of genetic data | 1,607 |  |  | <0.001 |
| African ancestry |  | 907 (63%) | 80 (51%) |  |
| European ancestry |  | 236 (16%) | 22 (14%) |  |
| Hispanic ancestry |  | 308 (21%) | 54 (35%) |  |
| Unknown |  | 0 | 85 |  |
| Constructed - Focal baby's gender | 1,692 |  |  | <0.001 |
| 1 Boy |  | 703 (48%) | 145 (60%) |  |
| 2 Girl |  | 748 (52%) | 96 (40%) |  |
| Visit | 1,692 |  |  | 0.058 |
| C |  | 728 (50%) | 105 (44%) |  |
| T |  | 723 (50%) | 136 (56%) |  |
| Global methylation | 1,692 | 0.475 (0.008) | 0.474 (0.009) | 0.093 |
| Pediatric clock | 1,692 | 10.29 (1.92) | 10.61 (1.94) | 0.018 |
| GRIM clock | 1,692 | 27.5 (4.7) | 28.6 (4.9) | <0.001 |
| Epithelial cell proportion (saliva) | 1,692 | 0.06 (0.13) | 0.08 (0.16) | 0.10 |
| Immune cell proportion (saliva) | 1,692 | 1.00 (0.14) | 0.98 (0.18) | 0.059 |
| Polymethylation score: coefficients for sustained smoking from newborn cordblood meta-analysis, w/ cell-type control (mean-centered) | 1,692 | 0.00 (0.23) | -0.01 (0.24) | 0.5 |
| Polymethylation score: coeffcients for sustained smoking from older children peripheral blood meta-analysis (mean-centered) | 1,692 | 0.00 (0.26) | -0.03 (0.28) | 0.043 |
| AHHR: cg05575921 | 1,692 | 0.77 (0.06) | 0.75 (0.07) | <0.001 |
| 1n (%); Mean (SD) | | | | |
| 2Pearson's Chi-squared test; Welch Two Sample t-test | | | | |

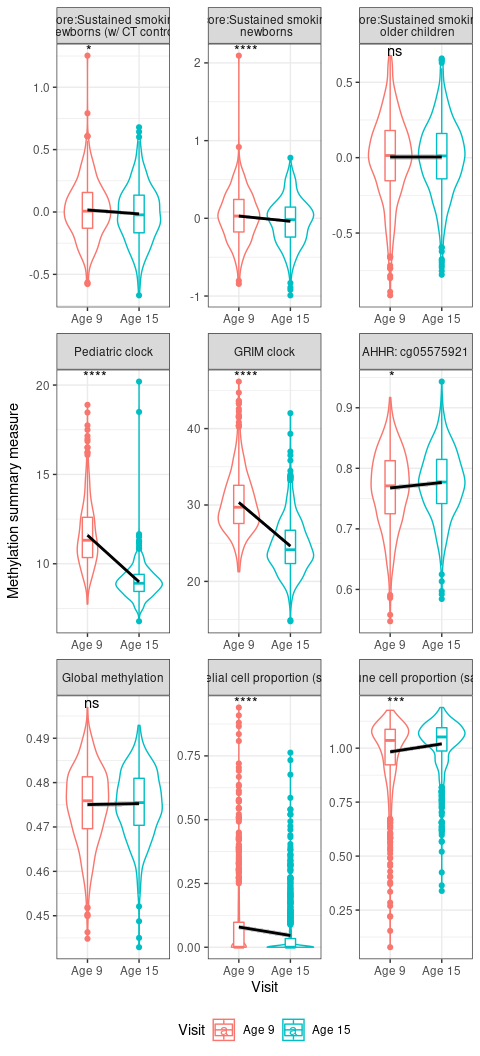
# Supplemental Figure 2 - Correlation of methylation summary measures



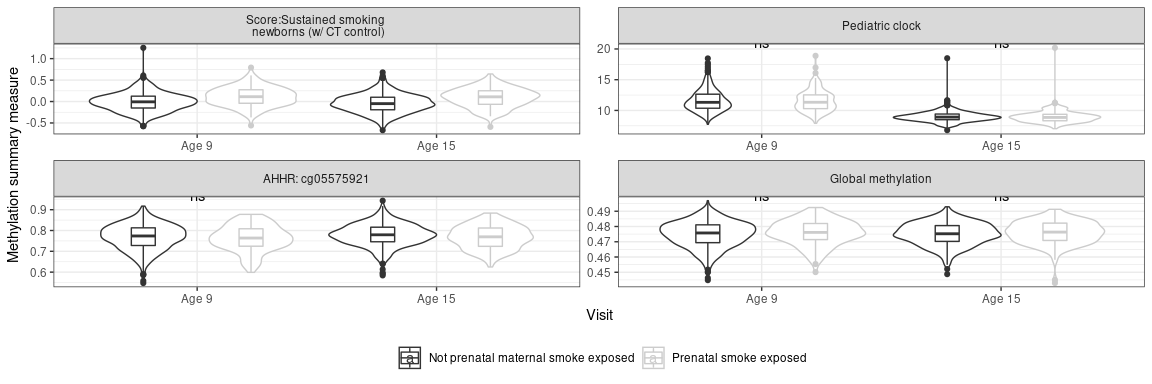
# Supplemental Figure 3 - Correlation between visits among DNA methylation summary measures



# Supplemental Figure XX: Distributions of methylation summary measures at age 9 and age 15

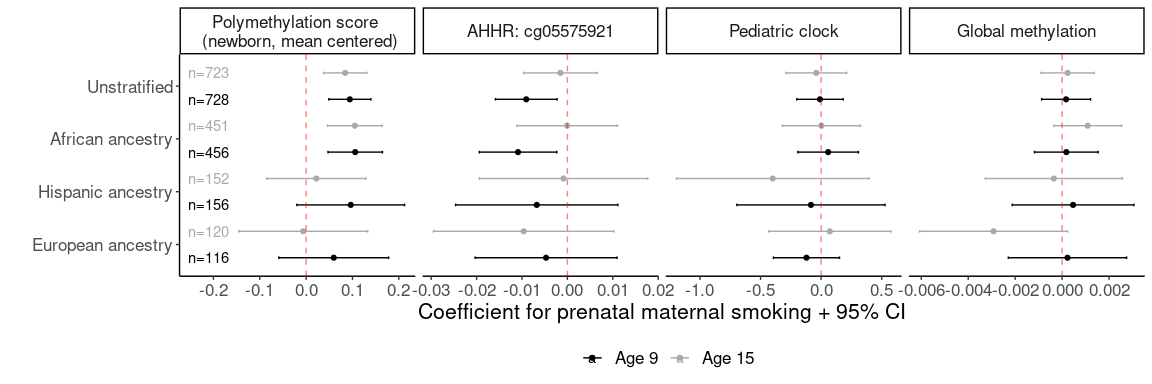


# Figure 2: Polymethylation scores are correlated and stable between age 9 and age 15 in a diverse sample of 784 US children



# Table 1: Bivariate associations between prenatal maternal smoking, DNA methylation summary measures and important covariates among a diverse sample of

# Figure 3 - Polymethylation scores are associated with prenatal maternal smoke exposure consistently across age and ancestry group in a sample of 784 US children



# Internal Table and Figure: Question: Remove postnatal maternal smoking variable from models and/or use base model as main paper figures?

Somewhat concerned about extrapolated inference: very few exposed children (post-ancestry stratification) who were not also exposed to postnatal smoking.

ancestry

PostnatalMaternalSmokingAny

smkPreg\_binary

Freq

African ancestry

No maternal smoking at age 1 and 5

No

534

European ancestry

No maternal smoking at age 1 and 5

No

122

Hispanic ancestry

No maternal smoking at age 1 and 5

No

207

African ancestry

Maternal smoking at age 1 or age 5

No

190

European ancestry

Maternal smoking at age 1 or age 5

No

45

Hispanic ancestry

Maternal smoking at age 1 or age 5

No

64

African ancestry

No maternal smoking at age 1 and 5

Yes

6

European ancestry

No maternal smoking at age 1 and 5

Yes

6

Hispanic ancestry

No maternal smoking at age 1 and 5

Yes

2

African ancestry

Maternal smoking at age 1 or age 5

Yes

177

European ancestry

Maternal smoking at age 1 or age 5

Yes

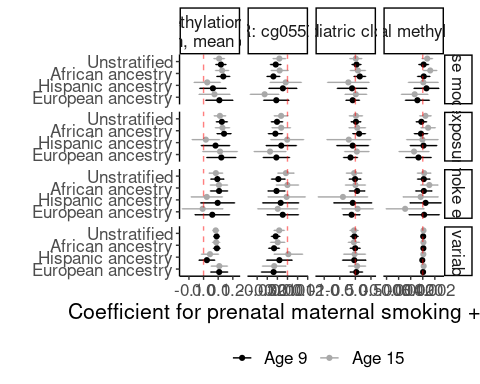
63

Hispanic ancestry

Maternal smoking at age 1 or age 5

Yes

35



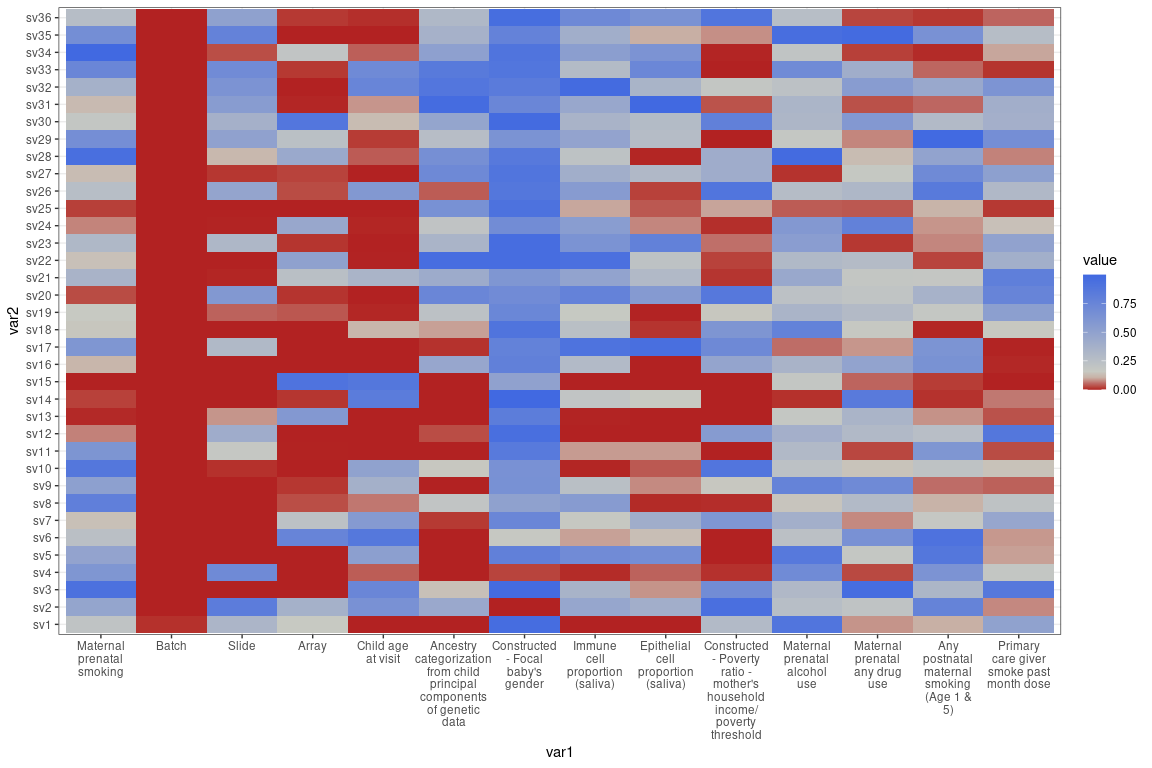
# Supplemental Table XX: Associations between prenatal maternal smoking and DNA methylation summary measures in multivariable models among 784 US children by age

|  | | Global methylation | | | | Pediatric clock | | | | Sustained smoking polymethylation score (newborns, cell-type controlled) | | | | Sustained smoking polymethylation score (older children) | | | | AHHR: cg05575921 | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Characteristic | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value |
| Age 9 models | Prenatal maternal smoke exposed | 728 | 0.00 | 0.00, 0.00 | 0.7 | 728 | -0.01 | -0.20, 0.18 | >0.9 | 728 | 0.09 | 0.05, 0.14 | <0.001 | 728 | 0.08 | 0.04, 0.11 | <0.001 | 728 | -0.01 | -0.02, 0.00 | 0.009 |
| Age 15 models | Prenatal maternal smoke exposed | 723 | 0.00 | 0.00, 0.00 | 0.7 | 723 | -0.04 | -0.29, 0.21 | 0.8 | 723 | 0.08 | 0.04, 0.13 | <0.001 | 723 | 0.07 | 0.04, 0.10 | <0.001 | 723 | 0.00 | -0.01, 0.01 | 0.7 |
| Longitudinal models | Prenatal maternal smoke exposed | 1,451 | 0.00 | 0.00, 0.00 | 0.6 | 1,451 | 0.00 | -0.18, 0.18 | >0.9 | 1,451 | 0.10 | 0.06, 0.14 | <0.001 | 1,451 | 0.08 | 0.05, 0.11 | <0.001 | 1,451 | -0.01 | -0.01, 0.00 | 0.10 |
| 1CI = Confidence Interval | | | | | | | | | | | | | | | | | | | | | |

# Supplemental Table XX: Associations between prenatal maternal smoking and DNA methylation summary measures in multivariable models among 784 US children by ancestry

|  | | Global methylation | | | | Pediatric clock | | | | Sustained smoking polymethylation score (newborns, cell-type controlled) | | | | Sustained smoking polymethylation score (older children) | | | | AHHR: cg05575921 | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Characteristic | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value |
| African ancestry | Prenatal maternal smoke exposed | 907 | 0.00 | 0.00, 0.00 | 0.3 | 907 | 0.08 | -0.15, 0.31 | 0.5 | 907 | 0.11 | 0.06, 0.17 | <0.001 | 907 | 0.08 | 0.04, 0.11 | <0.001 | 907 | -0.01 | -0.01, 0.00 | 0.2 |
| European ancestry | Prenatal maternal smoke exposed | 236 | 0.00 | 0.00, 0.00 | 0.7 | 236 | 0.10 | -0.20, 0.41 | 0.5 | 236 | 0.05 | -0.05, 0.15 | 0.3 | 236 | 0.07 | 0.01, 0.14 | 0.035 | 236 | -0.01 | -0.02, 0.01 | 0.2 |
| Hispanic ancestry | Prenatal maternal smoke exposed | 308 | 0.00 | 0.00, 0.00 | 0.8 | 308 | -0.28 | -0.76, 0.21 | 0.3 | 308 | 0.04 | -0.05, 0.13 | 0.4 | 308 | 0.07 | 0.00, 0.14 | 0.038 | 308 | 0.00 | -0.02, 0.01 | 0.5 |
| Unstratified | Prenatal maternal smoke exposed | 1,451 | 0.00 | 0.00, 0.00 | 0.6 | 1,451 | 0.00 | -0.18, 0.18 | >0.9 | 1,451 | 0.10 | 0.06, 0.14 | <0.001 | 1,451 | 0.08 | 0.05, 0.11 | <0.001 | 1,451 | -0.01 | -0.01, 0.00 | 0.10 |
| 1CI = Confidence Interval | | | | | | | | | | | | | | | | | | | | | |

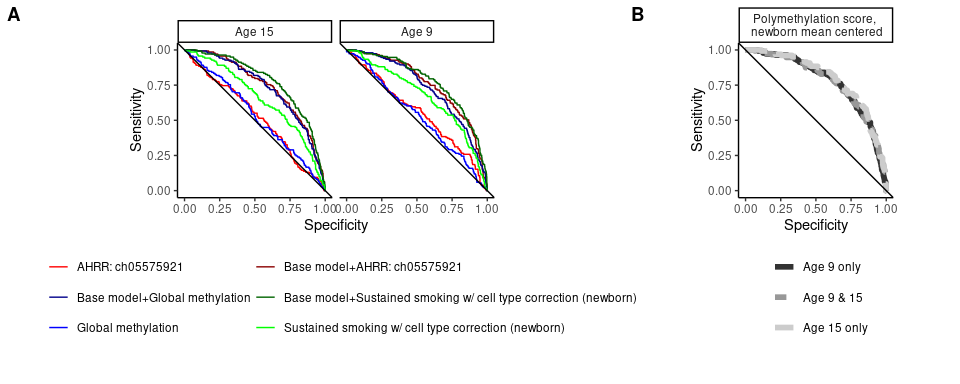
# Supplemental Figure XX: Surrogate variable associations with known covariates



# Supplemental Table XX: Surrogate variable models

|  | | Global methylation | | | | Pediatric clock | | | | Sustained smoking polymethylation score (newborns, cell-type controlled) | | | | Sustained smoking polymethylation score (older children) | | | | AHHR: cg05575921 | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Characteristic | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value | N | Beta | 95% CI1 | p-value |
| African ancestry | Prenatal maternal smoke exposed | 907 | 0.00 | 0.00, 0.00 | <0.001 | 907 | -0.12 | -0.31, 0.07 | 0.2 | 907 | 0.10 | 0.08, 0.12 | <0.001 | 907 | 0.06 | 0.05, 0.08 | <0.001 | 907 | -0.01 | -0.02, -0.01 | <0.001 |
| European ancestry | Prenatal maternal smoke exposed | 236 | 0.00 | 0.00, 0.00 | 0.004 | 236 | 0.15 | -0.08, 0.37 | 0.2 | 236 | 0.13 | 0.08, 0.18 | <0.001 | 236 | 0.09 | 0.06, 0.12 | <0.001 | 236 | -0.01 | -0.02, 0.00 | 0.004 |
| Hispanic ancestry | Prenatal maternal smoke exposed | 308 | 0.00 | 0.00, 0.00 | 0.7 | 308 | -0.10 | -0.41, 0.21 | 0.5 | 308 | 0.05 | 0.00, 0.10 | 0.031 | 308 | 0.04 | 0.01, 0.07 | 0.016 | 308 | 0.00 | -0.01, 0.01 | 0.8 |
| Unstratified | Prenatal maternal smoke exposed | 1,451 | 0.00 | 0.00, 0.00 | <0.001 | 1,451 | -0.09 | -0.23, 0.05 | 0.2 | 1,451 | 0.10 | 0.08, 0.11 | <0.001 | 1,451 | 0.06 | 0.05, 0.08 | <0.001 | 1,451 | -0.01 | -0.01, -0.01 | <0.001 |
| 1CI = Confidence Interval | | | | | | | | | | | | | | | | | | | | | |

# Figure 4 - Polymethylation scores more accurately delineate prenatal smoke exposure status than non-targeted DNA methylation summary measures and *a priori* CpGs



# Supplemental Table with AUC and p-value comparison to sustained smoking, newborn cell type controls polymethylation score

Age

Model

AUC

P value from comparison to sustained smoking score (newborns, cell type control, no other variables

Age 9

Global methylation

0.53

0.0000

Age 9

Pediatric clock

0.52

0.0001

Age 9

Any smoking (newborn)

0.62

0.0050

Age 9

Sustained smoking w/ cell type correction (newborn)

0.67

1.0000

Age 9

Sustain smoking (older children)

0.64

0.1126

Age 9

AHRR: ch05575921

0.55

0.0034

Age 9

CYP1A1: cg05549655

0.62

0.0408

Age 9

MYO1G: cg22132788

0.61

0.0373

Age 9

Sustained smoking (newborn)

0.62

0.0046

Age 15

Global methylation

0.54

0.0001

Age 15

Pediatric clock

0.50

0.0000

Age 15

Any smoking (newborn)

0.61

0.0187

Age 15

Sustained smoking w/ cell type correction (newborn)

0.66

1.0000

Age 15

Sustain smoking (older children)

0.60

0.0233

Age 15

AHRR: ch05575921

0.53

0.0009

Age 15

CYP1A1: cg05549655

0.63

0.3032

Age 15

MYO1G: cg22132788

0.58

0.0222

Age 15

Sustained smoking (newborn)

0.61

0.0200

# Supplemental Table with AUC and p-value comparison to base model with no methylation variables (i.e., using only child sex, income to poverty ratio, saliva/epithelial cell proportions and sample plate)

Age

Model

AUC

P value from comparison to base model (no methylation measure)

Age 9

Base model+No methylation

0.73

1.0000

Age 9

Base model+Global methylation

0.73

0.7716

Age 9

Base model+Pediatric clock

0.73

0.7815

Age 9

Base model+Any smoking (newborn)

0.76

0.0031

Age 9

Base model+Sustained smoking w/ cell type correction (newborn)

0.78

0.0007

Age 9

Base model+Sustain smoking (older children)

0.79

0.0002

Age 9

Base model+AHRR: ch05575921

0.76

0.0105

Age 9

Base model+CYP1A1: cg05549655

0.76

0.0204

Age 9

Base model+MYO1G: cg22132788

0.76

0.0140

Age 9

Base model+Sustained smoking (newborn)

0.77

0.0025

Age 15

Base model+No methylation

0.73

1.0000

Age 15

Base model+Global methylation

0.74

0.4490

Age 15

Base model+Pediatric clock

0.73

0.5445

Age 15

Base model+Any smoking (newborn)

0.77

0.0012

Age 15

Base model+Sustained smoking w/ cell type correction (newborn)

0.78

0.0007

Age 15

Base model+Sustain smoking (older children)

0.78

0.0004

Age 15

Base model+AHRR: ch05575921

0.74

0.1836

Age 15

Base model+CYP1A1: cg05549655

0.77

0.0069

Age 15

Base model+MYO1G: cg22132788

0.76

0.0164

Age 15

Base model+Sustained smoking (newborn)

0.77

0.0011

# Question for group: add EWAS results to this paper

* Arguments in favor: recent paper in Clinical Epigenetics included both
* Arguments against: enough here for a compelling story; deadline of December 1st; submit and if asked for we have it available.

## QQ plots

```{r. cache=T} par(mfrow=c(3, 1), mar=c(5.1, 4.1, 5.1, 2.1)) qqman::qq(child\_m1$P.Value, main=‘Age 9: Base model EWAS’) mtext(paste0(expression(lambda), ‘=’, round(calcLambda(child\_m1), 3)), side=3)

qqman::qq(teen\_m1$P.Value, main=‘Age 15: Base model EWAS’) mtext(paste0(expression(lambda), ‘=’, round(calcLambda(teen\_m1), 3)), side=3)

qqman::qq(lme\_smk$P.Value, main=‘Mixed effect models: Base model EWAS’) mtext(paste0(expression(lambda), ‘=’, round(calcLambda(lme\_smk), 3)), side=3)

## Beta vs beta plot: Age 9 vs age 15  
  
![](Results\_files/figure-docx/unnamed-chunk-20-1.png)<!-- -->  
  
## Upset plots  
  
![](Results\_files/figure-docx/unnamed-chunk-21-1.png)<!-- -->  
  
## Gap hunter   
  
  
  
  
## Additional covariate control and ancestry stratified models at top hits   
  
  
  
  
![](Results\_files/figure-docx/compare over time-1.png)<!-- -->![](Results\_files/figure-docx/compare over time-2.png)<!-- -->  
  
![](Results\_files/figure-docx/compare over ancestries-1.png)<!-- -->

## # A tibble: 1 × 3

## outcome est sign

## 

## 1 KIAA1875: cg07387044 <dbl [3]> -1

```

## Compare across models and studies