rplec: An R package of placental epigenetic clock to estimate aging by DNA-methylation-based gestational age

Herdiantri Sufriyana,[1] Emily Chia-Yu Su.[1,2,3,]*

[1] Institute of Biomedical Informatics, College of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan.
[2] Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan.
[3] Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei, Taiwan.

[*] Corresponding author:
Emily Chia-Yu Su. Institute of Biomedical Informatics, College of Medicine, National Yang Ming Chiao Tung University, 155, Section 2, Linong Street, Beitou District, Taipei 112304, Taiwan. Tel: +886-2-28267000 Ext. 67391. E-mail: emilysu@nycu.edu.tw.

Abstract

Background: Latest placental epigenetic clocks (PlECs) were claimed to be robust when applied to cases with either maternal or fetal adverse conditions. However, the accuracies in estimating gestational age (GA) were lower in earlier trimesters. We aimed to develop an R package of PlEC to estimate aging by DNA-methylation-based GA (DNAm-GA). **Methods:** We utilized 1742 samples of placental DNA methylation, provided by the 2024 Placental Clock DREAM Challenge. Our PlEC was trained using only used beta values at the common CpG sites in either the Infinium HumanMethylation-450 (*n*=930)/-850 BeadChip arrays (*n*=912) from Illumina, in which 100 samples were used for the validation set. External validation was independently evaluated by the challenge organizer using a publicly-unavailable test set (*n*=384). Elastic regression was applied to develop a three-stage prediction model to estimate: (1) DNAm-GA among normal samples; (2) first residual DNAm-GA among samples with known phenotypes that were leading to earlier termination; and (3) second residual DNAm-GA depending on the estimated GA from the previous stages. An R package was developed to simplify our scikit-learn models into a single function and to utilize DNAm-GA for placental aging study. **Results:** Our PlEC required beta values at 10,433 CpG sites and achieved the top performance in the validation set. Based on the test set, the root mean squared-error (RMSE) was 1.245 weeks. The RMSE for preterm samples were lower (0.558, 95% confidence interval [CI] 0.545, 0.570) compared to the two previous PlECs using the common dataset: (1) Lee et al (1.696, 95% CI 1.667, 1.724); and (2) Mayne et al (4.018, 95% CI 3.927, 4.108). We developed *rplec* R package with only two functions for preprocessing input and estimating DNAm-GA and two functions for conducting quality control and utilizing DNAm-GA for placental aging study. The simplified version of PlEC achieved similar performance with the original scikit-learn model with RMSE 0.102 (95% CI 0.101, 0.104), which was reasonably imperfect since Python and R handle floating/decimal numbers, differently. **Conclusions:** Our R package precisely estimated DNAm-GA and our analytical framework could utilize DNAm-GA for placental aging study. Our PlEC also allows individual assessment of placental aging in clinical settings via the residual DNAm-GA. Future studies are needed to refine the first residual GA estimation and reduce the number of predictors while maintaining the accuracy.

Keywords: placenta, epigenetic clock, biological age, DNA methylation, machine learning, R package.

Introduction

Epigenetic clock is emerging in the last decade, as a tool for identifying whether a medical condition involves an accelerated aging [1-3]. It is also considered as having a pivotal role in placental pathology [4], leading to great obstetrical syndromes that result in preterm delivery [5]. While placenta is inaccessible before termination of pregnancy, placental epigenetic clock (PlEC) for clinical application is more feasible via maternal blood because of the advancement of cell-free fetal DNA [6, 7]. However, currently available PlECs are less precise for placentas from pregnant women who underwent preterm delivery.

Placental pathology plays a putative role in great obstetrical syndromes [8, 9]. They consist of preeclampsia, fetal growth restriction, preterm labor, preterm premature rupture of the membranes, late spontaneous abortion, and placental abruption [10]. These conditions lead to earlier pregnancy termination up to 10. 6% (95% confidence interval [CI] 9.0, 12.0) because of either spontaneous preterm delivery or medically-indicated termination [11]. The conditions contribute both maternal mortality and neonatal morbidity [12, 13]. The latter likely result in admission to neonatal intensive care unit. Although the conditions have a devastating impact in maternal and child health, a hallmark of placental pathology, that causes those conditions, is still unclear [14]. Placental aging role may be central in their pathophysiological derangements. Understanding placental aging is helpful for discovering the effective therapeutic/preventive strategy [15]. Therefore, a widely accessible tool is needed for studying placental aging, including PlEC which is a tool that precisely estimate gestational age (GA) based on placental DNA methylation.

Latest PlEC were claimed to be robust when applied to cases with either maternal or fetal adverse conditions because of three reasons [16]. First, they included samples with these conditions to develop the placental clocks. Second, ones for control samples and uncomplicated term accurately estimated GA in complicated pregnancies. Third, correlations between DNA methylation levels and GA remained significant after adjusting for preeclampsia (PE). However, simply including pathological samples would not remove bias, particularly due to confounders and colliders [17-20]. The accuracy was also arguable since it was lower in earlier trimesters [16], including another placental clock from Mayne, Leemaqz [21]. Eventually, the confounding adjustment was insufficient since PE is not the only condition that might affect both placental methylation and GA [22-24]. Previous PlECs [16, 21] also did not correct selection bias, which might reduce the accuracy in population with different distributions of adverse pregnancy conditions.

Recently, the Dialogue for Reverse Engineering Assessment and Methods (DREAM) held the Placental Clock DREAM Challenge, a crowd-sourced competition to develop and validate a more precise PlEC [25, 26]. This challenge provided an independent test set, which avoid the participants to evaluate their model until the final model submission, by containerizing their model via Docker. We participated in this challenge and our prediction model was selected as the top performer according to the test set [27]. This model was developed by multistage predictive modeling which resulted in a large and complex PlEC. It may be less usable for non-computational scientists in medicine and molecular biology to identify placental aging in the condition of interest. To improve the usability of our PlEC, we aimed to develop an R package of PlEC to estimate aging by DNA-methylation-based GA (DNAm-GA).

Methods

Development of PlEC

Study design

Since this study developed a diagnostic prediction model, a cross-sectional paradigm was applied to the training set which were collected using diverse study designs. We acquired the training set from the Placental Clock DREAM Challenge. Their dataset comprised DNA methylation data from 1842 placental samples from 15 Gene Expression Omnibus (GEO) datasets and one dataset from ArrayExpress repository (Table S1). The data were obtained using either the Infinium HumanMethylation-450 (n=930)/-850 BeadChip arrays (n=912) from Illumina [27]. However, as a participant, we were only given 1742 samples, while the remaining 100 samples (i.e., the 850k array) were used for determining rank in leaderboard as the validation set. In this challenge, there were two tasks: (1) GA estimation using the common probes between the 450K and 850K arrays; (2) GA estimation using all the probes of the 850K arrays. We could only choose one model per task for final assessment, reasonably the best model using the validation set among maximum 10 submissions.

After the model selection in leaderboard was closed by the organizer, a prediction model was independently selected by the challenge organizer based on an independent test set (n=384) among the chosen models for any tasks from each participant. The test set was not publicly available and was collected proportionally to include all trimesters and various medical conditions at the Wayne State University School of Medicine and Detroit Medical Center Pregnancy and Perinatology Research Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

Phenotype data were not shared by the organizer due to inconsistently available across datasets. However, we obtained these data from the repositories based on the dataset identifiers in the challenge data. To deal with the availability problem, we imputed them using methylation data (see “Data preprocessing”).

Data preprocessing

We preprocessed the data by filtering probes and normalizing beta values. Probes were filtered out if they fell near SNPs, aligned to multiple locations, target locations on X and Y chromosomes, and were reliably detected in all samples (p-value < 0.01). Eventually, the filtered beta values were normalized using the beta mixture quantile dilation (BMIQ) method. For phenotype data, we manually curated the data using domain knowledge before applying data-driven imputation (see “Data availability”). We developed prediction models using beta values and random forest to impute the missing phenotypes.

Feature selection

We classified two types of predictors: (1) beta values at individual CpG sites; and (2) phenotypes, i.e., the medical conditions. Since phenotypes would not be provided at the time of prediction, we also used beta values to predict the phenotypes. Hence, the imputation models were also employed as parts of prediction models in this study.

A CpG site was selected for a prediction model if it was a differentially methylated position (DMP) with p-value of $\leq$0.05. The multiple-testing effects were corrected using the Benjamini-

Hochberg (BH) method. If there was no DMPs, we relaxed the criteria by omitting the BH correction but with *p*-value of ≤0.001.

Among those which were available and imputable in metadata across all datasets, a phenotype was selected for a prediction model if it was correlated to GA. This correlation must be directly leading to earlier termination based on domain knowledge. We applied the same method for correcting the multiple-testing effects. If a phenotype was selected, its prediction model was incorporated into the multistage prediction model.

Predictive modeling

In this study, we only described our model that was top-performing in the competition. It is because we do not have access to the models and their method writeups from other participants. Therefore, here we explained our model development to estimate GA in three stages.

First, we developed a prediction model using beta values to estimate GA among only normal samples. Normal status was defined as having none of the phenotypes that correlated to GA, except preterm phenotype. It is because we need to cover the whole range of GA using this model. It was subsequently used to estimate GA of all samples.

Second, for each of the selected phenotypes, the differences between true and predicted GA (i.e., GA residuals) were estimated by a prediction model using beta values among samples with that phenotype only. Using all samples, we also developed a prediction model using beta values to classify samples whether they had any of the selected phenotypes. The outputs of both types of prediction models were subsequently used to develop an ensemble prediction model to estimate GA residuals among all samples. The predictor in the aforementioned model was a multiply between the predicted GA residuals and the risk probability of each phenotype. We use the ensemble model to correct the normal GAs using the estimated GA residuals.

Third, we estimated the remaining GA residuals after the correction using one of three models, depending on whether the corrected GA prediction was ≤36, 37 to 40, or 41 weeks' gestation. The estimated GA was ceiled to follow clinicians' approach. This stage was intended to estimate clinician decision for terminating a pregnancy. The termination characteristics in the GA groups might be differrent in terms of medical and non-medical indications.

All the prediction models in any stages applied elastic net regression. For either estimating GA (i.e., regression task) or classifying a phenotype, we tuned the models using grid search. The best tuning was selected among the combinations of alpha values and L1 ratios.

Development of *rplec* R package

To facilitate user, we developed an R package to use our PlEC which is a multistage prediction model. It consists of $2n + 5$ prediction models (see "Predictive modeling") where n was the number of phenotypes that were correlated to GA. The input must be normalized among the specific probes for each sample and each model has the specific parameters for scaling the input. Furthermore, the ultimate goal of PlEC is to study placental aging in a condition of interest, which requires an understanding of how PlEC works and utilize its properties to identify accelerated or decelerated aging. Therefore, we need to achieve two objectives in developing this R package: (1) to simplify the usage of our model; and (2) to utilize DNAm-GA for placental aging studies.

Simplifying the usage of our model

For the first objective, we need two functions to normalize DNA methylation values and estimate DNAm-GA. We need these functions to run per sample/thus, a user could run only a batch of

samples each time. It is important to fit their computational resource (e.g., random-access memory [RAM] size).

The BMIQ normalization is inherently performed for one sample across beta values at all the available CpG sites. In our model development, we applied BMIQ using *champ.norm* function from ChAMP R package. Its installation is quite challenging due to massive dependencies to other R packages. We created a function by modifying a part of *champ.norm*, that was specifically used for preparing input of our model. This function normalizes beta values per sample across beta values at the selected CpG sites specifically for our model.

The models in the multistage prediction model were originally developed using scikit-learn and other Python libraries. However, we need to simplify each model into a simple additive operation in R language to reduce the need for expensive computation and dependencies to the Python libraries. It is possible due to the nature of elastic net regression algorithm. The mathematical definition is shown for the multistage prediction model (Equation 1), where we developed the model α to estimate GA among normal samples, as denoted by $\bar{y}_\alpha$. Deviation of $\bar{y}_\alpha$ from the true GA, particularly among abnormal samples, is estimated by $\bar{y}_\rho + \bar{y}_\tau$.

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{y}}_{\boldsymbol{\alpha}} + \bar{\boldsymbol{y}}_{\boldsymbol{\rho}} + \bar{\boldsymbol{y}}_{\boldsymbol{\tau}} \quad \textbf{Equation 1}$$

We developed the model ρ to estimate the first residual $\bar{y}_\rho$ among samples across phenotypes J where $j \epsilon J$ and each was j related to the true GA (Equation 2). The models j_1 and j_2 were respectively used for: (1) computing the probability of j occurring in a sample; and (2) estimating the residual among samples with phenotype j. Finally, we employed the model ρ to estimate $\bar{y}_\rho$. If the true GA is deviated from $\bar{y}_\alpha$, particularly among abnormal samples, then the residual is estimated by $\bar{y}_\rho + \bar{y}_\tau$. Any deviation of $\bar{y}_\alpha + \bar{y}_\rho$ from the true GA may be contributed from other phenotypes beyond J.

$$\bar{\boldsymbol{y}}_{\boldsymbol{\rho}} = \boldsymbol{\beta}_{\mathbf{0}_{\boldsymbol{\rho}}} + \sum_{j=1}^{p_\rho} \boldsymbol{\beta}_{j_\rho} \cdot \left(\frac{1}{1+e^{-\left(\beta_{0_{j_1}} + \sum_{\mu=1}^{p_{j_1}} \beta_{\mu_{j_1}} x_{\mu_{j_1}}\right)}} \cdot \left(\boldsymbol{\beta}_{\mathbf{0}_{j_2}} + \sum_{\mu=1}^{p_{j_2}} \boldsymbol{\beta}_{\mu_{j_2}} \boldsymbol{x}_{\mu_{j_2}} \right) \right) \quad \textbf{Equation 2}$$

We developed the model τ to estimate the second residual $\bar{y}_\tau$ among samples across different GA periods of pregnancy termination (Equation 3). The placental samples were mostly, if not all, collected at the time of pregnancy termination which might be either spontaneous or medically-indicated. Based on domain knowledge, we considered that the pregnancy termination had different characteristics for preterm, term before the date, and term after the date. Respectively, the residual $\bar{y}_\tau$ optimally estimated by $\bar{y}_{\tau_1}$, $\bar{y}_{\tau_2}$, and $\bar{y}_{\tau_3}$. Earlier termination leading to preterm delivery is more likely due to phenotypes beyond J, that are correlated to the true GA.

$$\bar{\boldsymbol{y}}_{\boldsymbol{\tau}} = \begin{cases} \bar{\boldsymbol{y}}_{\boldsymbol{\tau_1}}, \ \boldsymbol{if} \ \lceil \bar{\boldsymbol{y}}_{\boldsymbol{\alpha}} + \bar{\boldsymbol{y}}_{\boldsymbol{\rho}} \rceil \leq \mathbf{36} \\ \bar{\boldsymbol{y}}_{\boldsymbol{\tau_2}}, \ \boldsymbol{if} \ \mathbf{37} \leq \lceil \bar{\boldsymbol{y}}_{\boldsymbol{\alpha}} + \bar{\boldsymbol{y}}_{\boldsymbol{\rho}} \rceil \leq \mathbf{40} \\ \bar{\boldsymbol{y}}_{\boldsymbol{\tau_3}}, \ \boldsymbol{if} \ \lceil \bar{\boldsymbol{y}}_{\boldsymbol{\alpha}} + \bar{\boldsymbol{y}}_{\boldsymbol{\rho}} \rceil \geq \mathbf{41} \end{cases} \quad \textbf{Equation 3}$$

Either model M and ρ from scikit-learn in Python can be re-implemented using a simple mathematical operation in R (Equations 2 and 4). We only need to obtain each parameter β for predictor i from the scikit-learn model, and the average and standard deviation to scale the normalized beta value x_i. Eventually, for technical validation, we compared $\bar{y}$ from the multistage prediction model in scikit-learn with $\bar{y}$ from the function in our R package.

$$\bar{\boldsymbol{y}}_{\boldsymbol{M}} = \boldsymbol{\beta}_{\boldsymbol{0}_{\boldsymbol{M}}} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{p_\alpha}} \boldsymbol{\beta}_{\boldsymbol{i}_{\boldsymbol{M}}} \boldsymbol{x}_{\boldsymbol{i}_{\boldsymbol{M}}}, \boldsymbol{where\ M\ \epsilon\ \{\alpha, \tau_1, \tau_2, \tau_3, j_1, j_2\}} \quad \textbf{Equation 4}$$

Utilizing DNAm-GA for placental aging studies

For the second objective, we need two functions to perform quality control and identify placental aging. We need the former function to ensure our PlEC is well-generalized using the user's own dataset. Meanwhile, the latter function is used to compare aging across GA between case and control groups in the same dataset.

Quality control is performed by several metrics. We provide a function to create calibration plot and compute RMSE, MAE, and the coefficient *r*. A user is suggested for proceeding to placental aging identification if these metrics are similar to those in the validation set.

For identifying placental aging, we defined it as the deviation of placental DNAm-GA from the true GA y beyond that from the measurement error e (Equation 5). If the error value is acceptable, then "abnormal" placental aging is the residual $\bar{y}_\rho + \bar{y}_\tau$, as denoted by $\bar{y}_A$. It is because we developed the model α to estimate placental aging $\bar{y}_\alpha$ among normal samples. However, normal placenta may have accelerated/decelerated placental aging to some extent but not leading to abnormal condition.

$$\boldsymbol{y} = \bar{\boldsymbol{y}}_{\boldsymbol{\alpha}} + \bar{\boldsymbol{y}}_{\boldsymbol{\rho}} + \bar{\boldsymbol{y}}_{\boldsymbol{\tau}} + \boldsymbol{e} \quad \textbf{Equation 5}$$

To determine whether placental aging is abnormal, the residual $\bar{y}_A$ is compared between case and control groups. Specifically, the null hypothesis is defined in Equation 6 where $\bar{y}_A^{Group}(t_k)$ represents placental aging for each group at timepoint t as a percentile ($k = [1, 2, \ldots, 100]$) between a and b which are respectively the minimum and maximum GA across all groups in the user's dataset. Accordingly, the alternative hypothesis is stated in Equation 7. If H_0 is rejected, it indicates that placental aging is significantly different between case and control groups, suggesting abnormal placental aging in case group if the control represents the normal one.

$$\boldsymbol{H_0}: \bar{\boldsymbol{y}}_{\boldsymbol{A}}^{\boldsymbol{Case}}(\boldsymbol{t_k}) = \bar{\boldsymbol{y}}_{\boldsymbol{A}}^{\boldsymbol{Control}}(\boldsymbol{t_k})\ \forall \boldsymbol{t_k} \in [\boldsymbol{a}, \boldsymbol{b}] \quad \textbf{Equation 6}$$

$$\boldsymbol{H_1}: \bar{\boldsymbol{y}}_{\boldsymbol{A}}^{\boldsymbol{Case}}(\boldsymbol{t_k}) \neq \bar{\boldsymbol{y}}_{\boldsymbol{A}}^{\boldsymbol{Control}}(\boldsymbol{t_k})\ \boldsymbol{at\ least\ one\ t_k} \in [\boldsymbol{a}, \boldsymbol{b}] \quad \textbf{Equation 7}$$

Notably, placental samples case and control groups are sampled at different GA. Meanwhile, the comparison is paired between case and control for every t_k. Hence, we need to interpolate and extrapolate the residual $\bar{y}_A$ as a function g_A over a common range of t from a to b (Equation 8).

$$\bar{\boldsymbol{y}}_{\boldsymbol{A}}^{\boldsymbol{Group}}(\boldsymbol{t_k}) = \boldsymbol{approx}(\{\boldsymbol{g}_{\boldsymbol{A}}, \bar{\boldsymbol{y}}_{\boldsymbol{A}} \mid \boldsymbol{Group}\}, \boldsymbol{t_k}) \quad \textbf{Equation 8}$$

We also provide an aging plot which is a line plot where x-axis is GA in weeks and y-axis is case-control aging difference in weeks. A ribbon plot is shown to depict the uncertainty interval via bootstrapping. Furthermore, we allow a user to determine the values of a and b for testing the hypothesis within a specific range of GA that demonstrated accelerated/decelerated placental aging in the plot.

For use case examples, we randomly selected a sample from each combination of GA periods and preeclampsia status in the training set. The GA period consisted of ≤12, 13 to 26, 27 to 36, 37 to 40, and >40 weeks' gestation. This selection was intended to cover all trimesters before and after the ideal date of delivery (40 weeks' gestation).

Statistical analysis

The precision of GA estimation was primarily evaluated using root mean squared-error (RMSE) of the actual and predicted GA. We also computed mean absolute difference (MAE) and Pearson's correlation coefficient (*r*). The challenge organizer also used the same metrics, but only RMSE was used to select the best model in the test set. Point and interval estimates were inferred using 30-time bootstrapping in training set. We compared our model with the previous models [16, 21], primarily using preterm samples in the common dataset for training set (GSE74738). In addition, we visualized the abovementioned metrics if they were less than or equal to the top performer using the validation set. The visualization was shown in the training set across different phenotypes that were eligible (see "Feature selection"), trimesters, and datasets (including leaderboard for our submitted models).

For *rplec* R package, a technical validation was conducted by comparing all the metrics of between original and simplified versions respectively using the scikit-learn model and *rplec* function. Meanwhile, in placental aging analysis, we provided either Mann-Whitney U or permutation test for testing the hypothesis. To provide the uncertainty interval in the plot, we resample $\{g_A, \bar{y}_A\}$ within each group for each bootstrap iteration to compute the standard deviation (SD). The ribbon ranges from 1 to -1 SD from $\bar{y}_A$. Therefore, a user can identify whether placental aging in case group is accelerated (>0) or decelerated (<0) compared to control group.

Results

Model characteristics

Our final model required BMIQ normalization across beta values at 346,407 CpG sites (Table S2) which were the results of filtering probes in task 1 (the common probes between the 450K and 850K arrays). Due to time constraint during challenge, we used the same model from task 1 for both tasks. This model stacked 31 models (Table S3) and used beta values at 10,433 CpG sites (Table S4). Some of them were used for classifying a phenotype and estimating the residual GA from 13 phenotypes which were eligible (see "Feature selection"). These phenotypes were: (1) fetal growth restriction; (2) preeclampsia; (3) early preeclampsia; (4) preterm labor; (5) anencephaly; (6) spina bifida; (7) gestational diabetes mellitus; (8) diandric triploid; (9) miscarriage; (10) large-for-gestational-age; (11) subfertility; (12) hemolysis, elevated liver enzyme, and platelet syndrome; and (13) chorioamnionitis. The thirteen phenotypes were selected from 22 phenotypes (Figure S1).

Predictive performance of PlEC

Using the validation set, our final model ranked second and first respectively for tasks 1 and 2 (Table 1) among 42 and 13 models (Tables S5 and S6). All the top-4 final models in the validation

set retained the top performance order in the test set, except the "opipek" model. Our final model achieved RMSE 1.245 weeks in the test set.

Table 1. Predictive performance of the final model from each participant.

#	Task 1						Task 2					
	Validation set				Test set		Validation set				Test set	
	Model	***r***	**MAE**	**RMSE**	**Val. #**	**RMSE**	**Model**	***r***	**MAE**	**RMSE**	**Val. #**	**RMSE**
1	opipek	0.970	0.743	0.997	**2**	**1.245**	**herdiants**	**0.966**	**0.888**	**1.077**	**1**	**1.245**
2	**herdiants**	**0.966**	**0.888**	**1.077**	3	1.603	opipek	0.968	0.875	1.110	3	1.438
3	tpatel	0.951	0.997	1.206	4	1.683	tpatel	0.946	0.985	1.279	4	2.718
4	ialsag01	0.961	1.129	1.319	1	2.512	Chih-Han	0.897	1.499	1.886	2	3.721
5	Chih-Han	0.884	1.581	2.035	5	2.754	ay6621	0.608	2.572	3.380	6	28.140
6	tpanja	0.892	1.554	7	7	2.801	danielsu	0.048	13.322	17.106		
7	danielsu	0.732	2.169	2.678	6	3.006						
8	ay6621	0.608	2.572	3.38								
9	dw1227test	0.608	2.572	3.38								

Only the final model from each participant is shown. See the complete list of models in Tables S5 and S6. The results for our model are bold. MAE, mean absolute error (weeks); *r*, Pearson's correlation coefficient; RMSE, root mean squared-error (weeks).

We evaluated whether our model improved the performance on preterm samples compared to the previous models [16, 21]. Since we only have an access to the training set, this evaluation must be conducted using the common dataset with those for training the previous models, i.e., GSE74738. Our model outperformed the two previous models in all of the evaluation metrics. The RMSE for preterm samples were lower (0.558, 95% confidence interval [CI] 0.545, 0.570) compared to the two previous PlECs using the common dataset: (1) Lee et al (1.696, 95% CI 1.667, 1.724); and (2) Mayne et al (4.018, 95% CI 3.927, 4.108). Furthermore, our modeling technique improved the predictive performance across different phenotypes gradually from stage 1 to stage 2, while the previous model was similar to our model in stage 1 (Figure S2).

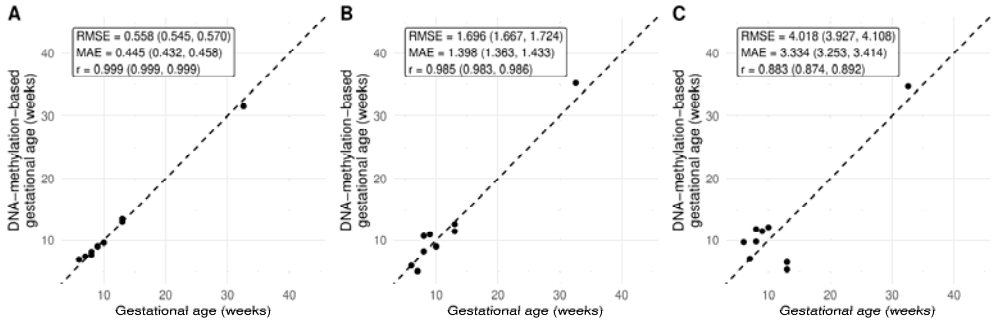

Figure 1. Predictive performance among preterm samples for: (1) our model; (2) Lee, Choufani [16](RPC); and (3) Mayne, Leemaqz [21]. See text regarding the dataset. MAE, mean absolute error (weeks); *r*, Pearson's correlation coefficient; RPC, robust placental clock; RMSE, root mean squared-error (weeks).

Predictive performance of the simplified PlEC in *rplec*

Using the training set, we conducted technical validation for the simplified version of PlEC in our *rplec* R package. This version demonstrated similar predictive performance in all of the evaluation metrics with the original one (Figures 2A and 2B). The simplified version of PlEC achieved similar performance with the original scikit-learn model with RMSE 0.102 (95% CI 0.101, 0.104). A

slightly different results between both version (Figure 2C) is reasonable since Python and R handle floating/decimal numbers, differently.

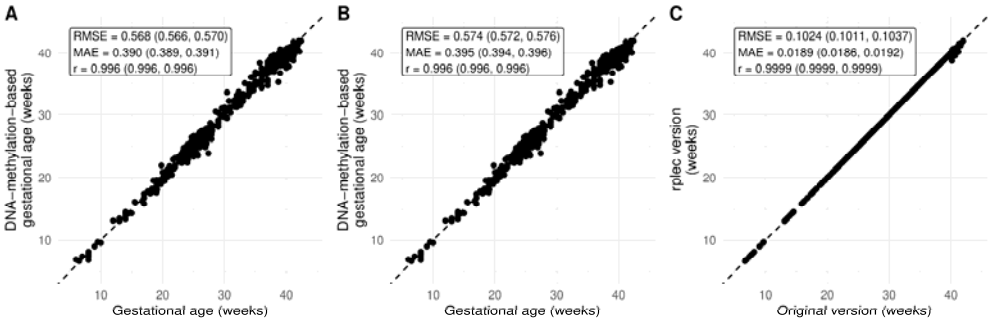

Figure 2. Technical validation of *rplec* function for GA estimation: (1) the original version using the scikit-learn model vs. the true GA; (2) the simplified version using *rplec* function vs. the true GA; and (3) the simplified version vs. the original version. MAE, mean absolute error (weeks); *r*, Pearson's correlation coefficient; RPC, robust placental clock; RMSE, root mean squared-error (weeks).

Use case example in utilizing DNAm-GA for placental aging studies

A user needs to install and load our R package either from the comprehensive R archive network (CRAN) or our GitHub repository (lines 1 to 3 Code snippet 1). We provided built-in data of beta values, GA, and phenotype (case/control) for 10 samples (lines 4 to 7 Code snippet 1), that were eligible as a use case (see "Utilizing DNAm-GA for placental aging studies"). After normalization and DNAm-GA estimation (lines 8 to 12 Code snippet 1), quality control is suggested to evaluate the predictive performance of our PlEC in the user's own dataset (line 13 Code snippet 1). If the predictive performance is accepted (Figure 3A), then placental aging is estimated using the residual DNAm-GA (lines 14 to 15 Code snippet 1). Eventually, a user need examine whether placental aging is significantly different between case and control groups (lines 17 Code snippet 1).

Code snippet 1. Placental aging analysis.

```
1:  install.packages("rplec")
2:  devtools::install_github("herdiantrisufriyana/rplec")
3:  library(rplec)
4:  data(beta_values_case)
5:  data(beta_values_control)
6:  data(ga)
7:  data(phenotype)
8:  norm_beta_values_case <- bmiq_norm_450k(beta_values_case)
9:  norm_beta_values_control <- bmiq_norm_450k(beta_values_control)
10: dnam_ga_case <- plec(norm_beta_values_case)
11: dnam_ga_control <- plec(norm_beta_values_control)
12: dnam_ga <- rbind(dnam_ga_case, dnam_ga_control)
13: qc(dnam_ga, ga, phenotype)
14: aging_case <- plec(norm_beta_values_case, type = "residual")
15: aging_control <- plec(norm_beta_values_control, type = "residual")
16: aging <- rbind(aging_case, aging_control)
17: ipla(aging, ga, phenotype)
18: ipla(aging, ga, phenotype, method = "Mann-Whitney U")
19: ipla(aging, ga, phenotype, method = "Mann-Whitney U", from = 5, to = 20)
```

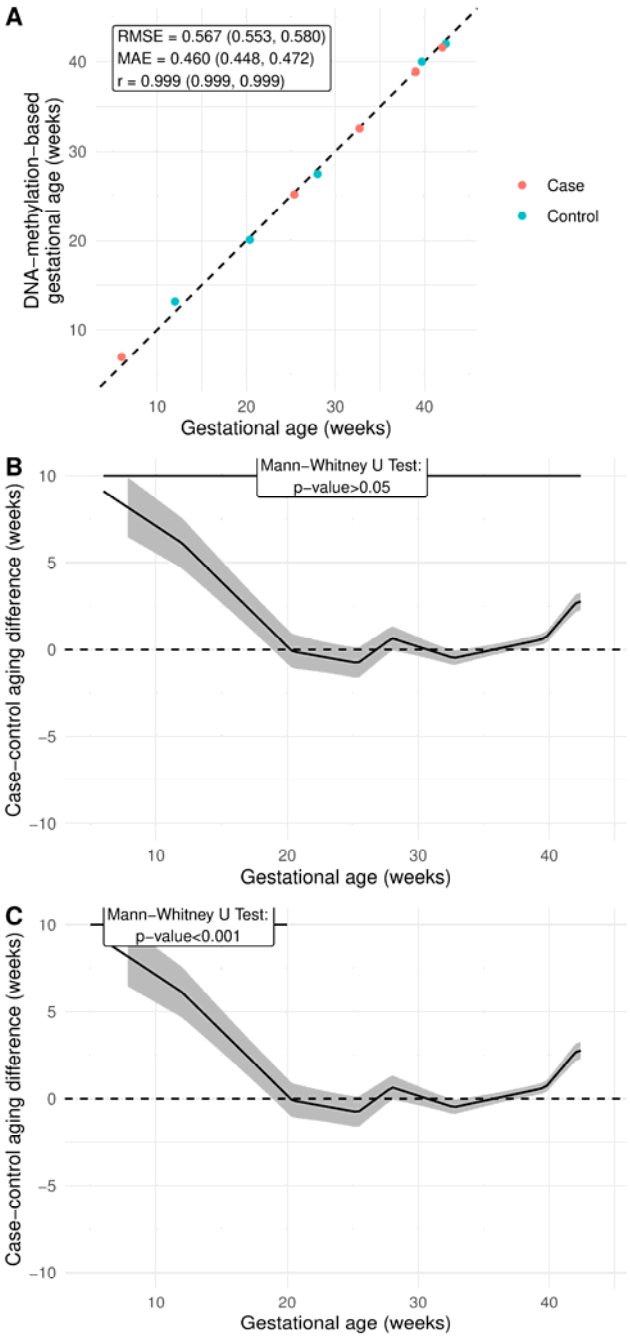

Figure 3. Placental aging analysis output in *rplec*: (A) quality control; (B) case-control aging difference with hypothesis testing across GA; and (C) case-control aging difference with hypothesis testing from 5 to 20 weeks' gestation. MAE, mean absolute error (weeks); *r*, Pearson's correlation coefficient; RMSE, root mean squared-error (weeks).

In this example (Table S9), we can observe difference in placental aging between case and control (Figures 3B and 3C). Using Mann-Whitney U test (line 18 Code snippet 1), placental aging in case group is not significantly different to that in control group across all trimesters (Figure 3B).

However, a larger difference is observed from ~5 to ~20 weeks' gestation. In this period, placental aging in case group is significantly different to that in control group (Figure 3C; line 19 Code snippet 1). The y-axis indicates that the case group demonstrated an accelerated aging from ~6 to ~9 weeks within ~8 to ~12 weeks' gestation. In other words, the placental age in this group was similar to ~17 to ~18 weeks' gestation in control group during first trimester, but the accelerated aging is resolved after 20 weeks' gestation.

Discussion

Finding summary

Our PlEC required beta values at 10,433 CpG sites which were available in either the Infinium HumanMethylation-450/-850 BeadChip arrays from Illumina. The precision in estimating DNAm-GA was 1.245 weeks based on an independent external validation. Compared to the previous PlECs [16, 21], our model maintained the precision for preterm samples.

We have simplified our PlEC from 31 scikit-learn models into a single function in an R package, i.e., *rplec*, with a built-in function for preprocessing the input. The simplified version maintained the precision of the original model. We also provide another function to utilize our PlEC for a placental aging study, appropriately based on the design of our PlEC.

Research and clinical implications

Generally, there are two main challenges in epigenetic clock research [3]. First, the challenge remains in identifying whether the acceleration is the precision error of a DNAm biological clock or a genuine biological aging effect. Second, biological age acceleration can be defined as: (1) absolute age acceleration (AAA), which is the difference between DNAm and chronological age; or (2) relative age acceleration (RAA), which is the residual of linear regression of DNAm age versus chronological age. In general, the second definition is preferred because it is arguably able to adjust sampling error. However, using RAA, the effect of age acceleration may not be detected on the outcome when it is used only in the case group, e.g., the effect on body mass index among individuals with clinical obesity [28].

Our modeling technique entangled the biological aging effect from the precision error (Equation 5). A user can re-evaluate the precision of our PlEC using their own dataset. The error may be technical or biological variation including that in the control group. Our statistical framework also corrects sampling error. It is because this framework only compares the residual DNAm-GA which refers to placental aging beyond those in normal aging process (Equations 6 and 7). Therefore, we can expect our PlEC and the statistical framework to refine current method for placental aging studies.

Several studies had used PlEC for placenta aging-related studies [29-31]. However, its clinical application is hampered regardless the advancement of cell-free fetal DNA, because RAA cannot be used to identify placental aging in a single sample. Our modeling technique allows the usage of the residual DNAm-GA for a single sample and comparison to those from the control samples at the same time. Hence, the residual DNAm-GA is expected to move forward the usage of PlEC for clinical application.

Strength and limitation

Our study demonstrated that PlEC precision could be improved solely by modeling techniques beyond adding more data and applying more sophisticated algorithms. The residual DNAm-GA

also offers individual assessment of placental aging based on AAA while also adjusts sampling error similar to RAA. Our PlEC can be re-evaluated in every placental aging study using quality control function in the R package.

However, we also considered several limitations in this study. First, a model in our PlEC may have variations of accuracy and precision respectively in classifying and estimating GA for a particular phenotype. This limitation may lead to poor precision when a user conducts quality control using their own dataset. Yet, the PlEC quality can be improved by partially retraining the model with more data for that phenotype. Second, our PlEC also did not include several phenotypes which might affect GA via early termination, e.g., placental abruption. Nevertheless, our modeling framework allows future studies to add a model and re-train some models in our PlEC. Third, our PlEC still required the whole array. While it can be a burden for clinical application, our PlEC is still useful for research purpose and reducing the predictors further with similar accuracy.

Conclusions and future studies

We have developed an R package of PlEC with the precision within 1.245 weeks based on an independent external validation. This package was designed to allow users to preprocess their data and conduct quality control using their own data. Eventually, we provide a function for a user utilize DNAm-GA for placental aging study. In addition, our PlEC allows individual assessment of placental aging in clinical settings.

Future studies are needed to extend our PlEC for the phenotypes beyond those in this study, particularly when quality control indicates a poor precision among samples with these phenotypes. The models for the phenotypes in our PlEC can also be improved. The number of predictors should be reduced to allow extensive application in clinical settings while maintaining the accuracy.

Acknowledgments

This study was funded by: (1) the Postdoctoral Accompanies Research Project from the National Science and Technology Council (NSTC) of Taiwan (grant nos.: NSTC111-2811-E-038-003-MY2 and NSTC113-2811-E-A49A-003) to HS; and (2) the National Science and Technology Council in Taiwan (grant no. NSTC113-2221-E-A49-193-MY3), the Ministry of Science and Technology (MOST) of Taiwan (grant nos.: MOST110-2628-E-038-001 and MOST111-2628-E-038-001-MY2), the University System of Taipei Joint Research Program (grant no.: USTP-NTOU-TMU-112-04), and the Higher Education Sprout Project from the Ministry of Education (MOE) of Taiwan (grant no.: DP2-111-21121-01-A-05 and DP2-TMU-112-A-13) to ECYS. These funding bodies had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

CRediT authorship

HS: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Visualization, Funding acquisition. **ECYS:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition. All authors have read and approved the manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of interests

The authors declare that they have no competing interests.

Data availability

The source codes were shared (https://github.com/herdiantrisufriyana/placental_clock_dream_task1 and https://github.com/herdiantrisufriyana/placental_clock_dream_task2) for the development of our PlECs. Meanwhile, the R package can be installed either from the CRAN or our GitHub repository (https://github.com/herdiantrisufriyana/rplec). Data are publicly available in GEO where the accession numbers are provided in the source codes. However, the curated version (https://www.synapse.org/Synapse:syn59520082/wiki/628063) can only be accessed by request to the challenge organizer.

References

1. Bocklandt, S., et al., *Epigenetic predictor of age.* PLoS One, 2011. **6**(6): p. e14821.
2. Horvath, S. and K. Raj, *DNA methylation-based biomarkers and the epigenetic clock theory of ageing.* Nat Rev Genet, 2018. **19**(6): p. 371-384.
3. Teschendorff, A.E. and S. Horvath, *Epigenetic ageing clocks: statistical methods and emerging computational challenges*. Nat Rev Genet, 2025.
4. Pan, M., et al., *The role of placental aging in adverse pregnancy outcomes: A mitochondrial perspective.* Life Sci, 2023. **329**: p. 121924.
5. Ciampa, E.J., et al., *Hypoxia-inducible factor 1 signaling drives placental aging and can provoke preterm labor.* Elife, 2023. **12**.
6. Goldwaser, T. and S. Klugman, *Cell-free DNA for the detection of fetal aneuploidy.* Fertil Steril, 2018. **109**(2): p. 195-200.
7. Lo, Y.M.D., et al., *Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies.* Science, 2021. **372**(6538).
8. Brosens, I., P. Puttemans, and G. Benagiano, *Placental bed research: I. The placental bed: from spiral arteries remodeling to the great obstetrical syndromes.* Am J Obstet Gynecol, 2019. **221**(5): p. 437-456.
9. Harris, L.K., et al., *Placental bed research: II. Functional and immunological investigations of the placental bed.* Am J Obstet Gynecol, 2019. **221**(5): p. 457-469.
10. Brosens, I., et al., *The "Great Obstetrical Syndromes" are associated with disorders of deep placentation.* Am J Obstet Gynecol, 2011. **204**(3): p. 193-201.
11. Chawanpaiboon, S., et al., *Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis.* Lancet Glob Health, 2019. **7**(1): p. e37-e46.
12. Say, L., et al., *Global causes of maternal death: a WHO systematic analysis.* Lancet Glob Health, 2014. **2**(6): p. e323-33.
13. Saleem, S., et al., *Neonatal deaths in infants born weighing ≥ 2500 g in low and middle-income countries.* Reprod Health, 2020. **17**(Suppl 2): p. 158.
14. Hoffman, M.K., *The great obstetrical syndromes and the placenta.* Bjog, 2023. **130 Suppl 3**(Suppl 3): p. 8-15.
15. Li, Y., et al., *Molecular mechanisms of aging and anti-aging strategies.* Cell Commun Signal, 2024. **22**(1): p. 285.
16. Lee, Y., et al., *Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels.* Aging (Albany NY), 2019. **11**(12): p. 4238-4253.
17. Greifer, N. and E.A. Stuart, *Matching Methods for Confounder Adjustment: An Addition to the Epidemiologist's Toolbox.* Epidemiol Rev, 2022. **43**(1): p. 118-129.

18. Liu, T., et al., *Can statistic adjustment of OR minimize the potential confounding bias for meta-analysis of case-control study? A secondary data analysis.* BMC Med Res Methodol, 2017. **17**(1): p. 179.
19. Holmberg, M.J. and L.W. Andersen, *Collider Bias.* Jama, 2022. **327**(13): p. 1282-1283.
20. Digitale, J.C., et al., *Key concepts in clinical epidemiology: collider-conditioning bias.* J Clin Epidemiol, 2023. **161**: p. 152-156.
21. Mayne, B.T., et al., *Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation.* Epigenomics, 2017. **9**(3): p. 279-289.
22. Jiang, M., et al., *A case control study of risk factors and neonatal outcomes of preterm birth.* Taiwan J Obstet Gynecol, 2018. **57**(6): p. 814-818.
23. Shi, D., et al., *Placental DNA methylation analysis of selective fetal growth restriction in monochorionic twins reveals aberrant methylated CYP11A1 gene for fetal growth restriction.* Faseb j, 2023. **37**(10): p. e23207.
24. Konwar, C., et al., *DNA methylation profiling of acute chorioamnionitis-associated placentas and fetal membranes: insights into epigenetic variation in spontaneous preterm births.* Epigenetics Chromatin, 2018. **11**(1): p. 63.
25. Stolovitzky, G., D. Monroe, and A. Califano, *Dialogue on reverse-engineering assessment and methods: the DREAM of high-throughput pathway inference.* Ann N Y Acad Sci, 2007. **1115**: p. 1-22.
26. Meyer, P. and J. Saez-Rodriguez, *Advances in systems biology modeling: 10 years of crowdsourcing DREAM challenges.* Cell Syst, 2021. **12**(6): p. 636-653.
27. *Placental Clock DREAM Challenge*. 2024; Available from: https://www.synapse.org/Synapse:syn59520082/wiki/628063.
28. Horvath, S., et al., *Obesity accelerates epigenetic aging of human liver*. Proc Natl Acad Sci U S A, 2014. **111**(43): p. 15538-43.
29. Polinski, K.J., et al., *Epigenetic gestational age and the relationship with developmental milestones in early childhood.* Hum Mol Genet, 2023. **32**(9): p. 1565-1574.
30. Saeed, H., et al., *Placental accelerated aging in antenatal depression.* Am J Obstet Gynecol MFM, 2024. **6**(1): p. 101237.
31. Tekola-Ayele, F., et al., *Sex differences in the associations of placental epigenetic aging with fetal growth.* Aging (Albany NY), 2019. **11**(15): p. 5412-5432.