An epigenetic clock to estimate postmenstrual and postnatal age in preterm infants

Stefan Graw[1], Marie Camerota[2], Brian S. Carter[3], Jennifer Helderman[4], Julie A. Hofheimer[5], Elisabeth C. McGowan[5], Charles R. Neal[6], Steven L. Pastyrnak[7], Lynne M. Smith[8], T. Michael O'Shea[5], Barry M. Lester[2], Carmen J. Marsit[1], Todd M. Everson[1]

1 Emory University, 2 Brown University, 3 Children's Mercy Kansas City, 4 Wake Forest School of Medicine, 5 University of North Carolina, 6 University of Hawai'i, 7 Helen DeVos Children's Hospital, 8 The Lundquist Institute

Background

- Rate of preterm birth is close to 10%
- Preterm infants have varying degrees of prematurity at birth and during the neonatal period affecting organ structure and function
- Preterm birth is associated with acute and long term morbidities
- Epigenetic clocks have been developed to estimate biological aging, which has been associated with age-related phenotypes in adults, such as frailty, chronic diseases and mortality
- Variety of epigenetic clocks for different tissue types and age ranges

Gap: Lack of epigenetic clock focused on age prediction of preterm infants during the neonatal period
Objective: Develop an epigenetic estimator for neonatal age metrics in preterm infants

Methods

Neonatal Neurobehavior and Outcomes in Very Preterm Infants (NOVI) study

- 543 very preterm infants born at less than 30 weeks gestational age
- Buccal cells were collected at NICU discharge
- DNA methylation of buccal cells was profiled with the Infinium MethylationEPIC BeadChip (EPIC)
- Data set was split into training set with 434 preterm infants (80%) and testing set with 109 preterm infants (20%)

Independent Saliva set (GSE72120)

- DNA methylation was profiled in saliva samples of 34 preterm infants using the Infinium HumanMethylation450 BeadChip (450K)

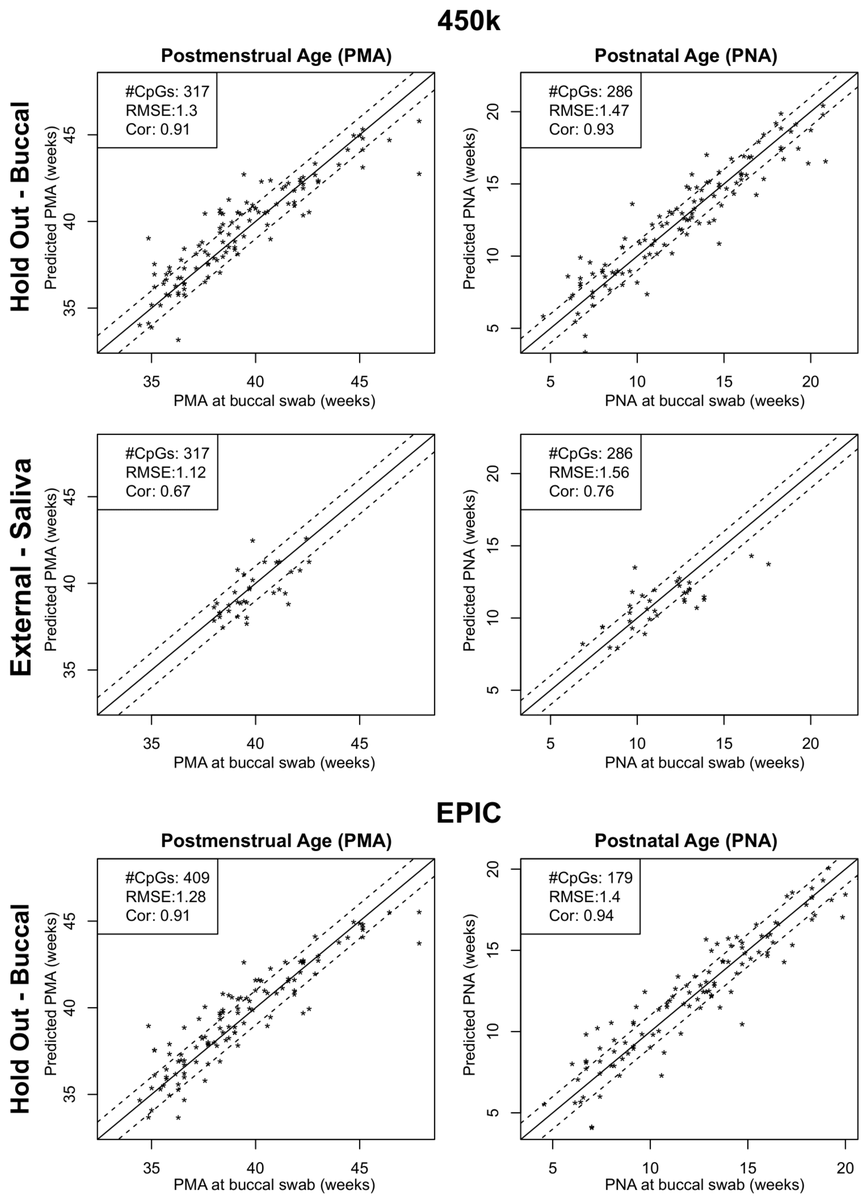
Elastic net regression

Elastic net regression was applied to our training set and identified two sets of CpGs predictive of postmenstrual age in weeks (PMA, time from conception to buccal swab) and postnatal age in weeks (PNA, time from birth to buccal swab), compatible the EPIC array and another two sets of CpGs compatible with the 450K array. We tested the predictors of PMA and PNA in our held out buccal samples (EPIC and 450K) and in an external independent salvia set (450K)

$$\hat{\beta} = \arg\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 + \frac{1-\alpha}{2}||\beta||_2^2 + \alpha||\beta||_1$$

where $\alpha = 0.5$

Results



- PMA and PNA sets compatible with the EPIC array: 409 and 179 CpGs
- PMA and PNA sets compatible with the 450k array: 317 and 286 CpGs
- Overlap between 450k and EPIC array:
  - PMA: 138 CpGs (out of 317 and 409 CpGs)
  - PNA: 67 CpGs (out of 286 and 179 CpGs)
- Prediction performance was evaluated by root-mean-square error (RMSE) and correlation

		PMA			PNA		
		#CpGs	RMSE	Cor	#CpGs	RMSE	Cor
450k	Hold out - Buccal	317	1.30	0.91	286	1.47	0.93
	External - Saliva	317	1.12	0.67	286	1.56	0.76
EPIC	Hold out - Buccal	409	1.28	0.91	179	0.76	0.94

Discussion

We identified four sets of CpGs predictive of PMA and PNA compatible with the 450k and EPIC array. Within the testing set, all four sets demonstrated very strong correlation between the predicted and measured PMA and PNA (0.91-0.94) and RMSE ranging from 1.28 to 1.47. The two sets of CpGs compatible with the 450k array could be evaluated in an independent external data set in which DNA methylation was profiled for saliva instead of buccal samples. To our knowledge, there currently exists no other data set that profiled DNA methylation in buccal cells of preterm infants during the neonatal period. Nevertheless, the two sets of CpGs compatible with the 450k array achieved similar prediction performance within the saliva data set with RMSE of 1.12 and 1.56 for PMA and PNA respectively. The diminished correlation can be attributed to the smaller range of PMA and PNA in the saliva data set.
These epigenetic estimators of PMA and PNA in preterm infants might help us to develop a tool for more accurate determination of neonatal aging. It may also help us gain insight into early life biological aging, including whether neonatal age acceleration is connected to positive or negative health outcomes in very preterm infants, who are at heightened risk of neonatal morbidities and long-term developmental impairments.

Acknowledgements

This work was supported by NIH Grants NICHD R01HD072267 (Lester and O'Shea), R01HD084515 (Lester and Everson), UH3OD023347 (Lester, Marsit, and O'Shea), and the HERCULES Center (P30 ES019776).