GROWUPGROWOLD
HUMAN DEVELOPMENT, BIRTH TO DEATH

Moffitt & Caspi
Genes Environment Health Behavior

Concept Paper Form

Provisional Paper Title: Identifying DNA methylated-based predictors of miscarriage risk
Proposing Author: Trevor Doherty and Dr Therese Murphy
Author's Email: therese.murphy@tudublin.ie
P.I. Sponsor: Professor Avshalom Caspi
Today's Date: 8/28/2023

Please describe your proposal in 2-3 pages with sufficient detail for helpful review.

Objective of the study:

The primary aim of this study is to assess the predictive capability of second-generation epigenetic clocks for sporadic and recurrent miscarriage (RM) risk. We will compute a range of DNA methylation (DNAm)-based estimators of biological aging using the DNAm data from Dunedin. Strongly predictive DNAm-based biological aging estimators will be combined with additional predictive measures of miscarriage, such as Poly Cystic Ovary Syndrome (PCOS) diagnosis and BMI, to develop a holistic algorithm of miscarriage risk. Our study will, for the first time, examine DNA methylation age acceleration in the context of miscarriage risk prospectively and may lead to the identification of potential DNAm-based biomarkers for miscarriage.

The primary Objectives of this study are:

1) Examine the relationship between a range of DNAm-based predictors of age acceleration measures, at age 26, with future history of both sporadic and recurrent miscarriage between the age of 26 and 37.

2) Validate our findings by examining predictive DNAm-based age acceleration measures (identified in objective 1) at age 38 and subsequent risk of miscarriage between the age of 38-45.

3) Strongly predictive DNAm-based estimators will be combined with additional predictive measures of miscarriage, such as PCOS diagnosis and BMI, to develop a holistic algorithm of miscarriage risk.

Data analysis methods:

Compute a range of DNAm-based estimators of biological aging
We will compute a range of DNAm-based estimators of biological aging using the DNAm data from Dunedin (see below), some of which are readily available already in the Dunedin dataset.

- DunedinPACE (Belsky et al, 2022) (Age 26/38).
- GrimAge (Lu et al, 2019) (Age 26/38)

- GrimAge2 (Lu et al, 2022) (Age 26/38)
- PhenoAge (Levine et al, 2018) (Age 26/38)
- Epigenetic Age (Zhang et al, 2019) (Age 26/38)
- DNAmTL (Lu et al, 2019) – already calculated for age 26 and 38 by proposing authors for Doherty et al, 2023 paper.
- DNAmTL (Doherty et al, 2023) – already calculated for age 26 and 38 by proposing authors for Doherty et al, 2023 paper.

If not already available in Dunedin data set, The Horvath Lab webtool will be used to calculate the following measures:

- GrimAge (Lu et al, 2019) (Age 26/38)
- GrimAge2 (Lu et al, 2022) (Age 26/38)
- PhenoAge (Levine et al, 2018) (Age 26/38)
- Epigenetic Age (Zhang et al, 2019) (Age 26/38)

For each epigenetic clock, we will calculate epigenetic 'age acceleration' by regressing participants' clock-estimated ages on their chronological ages and computing the residuals. This measure is often used as an estimate of the aging rate.

Examining the relationship between DNAm-based biological age estimates and miscarriage risk

Step 1: Logistic regression will be used to examine the relationship between a range of DNAm-based estimators of biological aging (see above) measured at age 26 and sporadic miscarriage (1 or more miscarriage, n=67) and controls (no history of miscarriage, n=361) that occurred between the age of 26 and 37.

Step 2: Logistic regression will be used to examine the relationship between a range of DNAm-based estimators of biological aging (see above) measured at age 26 and recurrent miscarriage (2 or more miscarriage, n=21) and controls (no history of miscarriage, n=361) between the age of 26 and 37.

Step 2: We will validate our findings in Dunedin Age 38 DNA methylation cohort. Logistic regression will be used to examine the relationship between DNAm-based estimators of biological aging (identified in Step 1 above) measured at age 38 and i) sporadic miscarriage and ii) recurrent miscarriage that occurred between the age of 38 to 45.

Developing a holistic algorithm of miscarriage risk

Using logistic regression we will evaluate the predictive ability of strongly predictive DNAm-based estimators with additional predictive measures of miscarriage risk (e.g. PCOS diagnosis and BMI) as co-variates, to develop a holistic algorithm of miscarriage risk.

<u>Variables needed at which ages:</u>

All variables are needed at age 26 and 38 unless otherwise stated.

- Unnormalised and normalised DNA methylation values (Females only) – Unnormalised only required if the following estimates are not available
- DNAm based epigenetic clock measures *(including Age acceleration estimates (i.e. residuals after regressing on Age):*
    - DunedinPACE (Belsky et al, 2022) (Age 26, Age 38).
    - GrimAge (Lu et al, 2019) (Age 26/38)

- GrimAge2 (Lu et al, 2022) (Age 26/38)
- PhenoAge (Levine et al, 2018) (Age 26/38)
- Epigenetic Age (Zhang et al, 2019) (Age 26/38)
- DNAmTL (Lu et al, 2019) – already calculated for age 26 and 38 by proposing authors for Doherty et al, 2023 paper.
- PCA-EN TL (Doherty et al, 2023) – already calculated for age 26 and 38 by proposing authors for Doherty et al, 2023 paper.

- Measured cellular composition measures
- Sentrix ID/Chip ID
- Smoking
- BMI
- Poly cystic ovary syndrome (PCOS) diagnosis (if available)

Significance of the Study (for theory, research methods or clinical practice):

Miscarriage, as defined by the World Health Organization (WHO) as the spontaneous loss of an embryo or fetus up to 20–22 weeks of gestation (1), is a common, complex trait affecting ~15-20% of clinically confirmed pregnancies. Recurrent miscarriage (RM), which is considered to be a more severe phenotype, is currently defined as two or more miscarriages (1). Miscarriage is associated with excessive bleeding, infection, anxiety, depression, infertility, and an increased lifetime risk of cardiovascular disease (2). Risk factors include maternal age, aneuploidy, parental chromosomal abnormalities, maternal thrombophilias, obesity, and endocrine and immunological dysregulation (3) but the etiology and pathogenesis of ~50% of recurrent miscarriage cases remain unclear.

Miscarriage is partly driven by genetic variation potentially related to placental biology (4). Previous research has suggested a role for epigenetic dysregulation in recurrent miscarriage, and several studies have reported that imprinted gene defects lead to early pregnancy loss in animal studies (5). To date, there is limited research examining the role of DNA methylation in miscarriage and sample sizes are limited (6), and no studies have specifically examined DNA methylation age (DNAm Age) acceleration in women as a biomarker for miscarriage risk. Given the strong association with maternal age and miscarriage risk, we hypothesis that accelerated biological aging (as measured by DNA methylation) may be a biomarker for future miscarriage risk. Statistical and machine learning techniques have previously been employed to estimate variables such as biological age using DNAm data. "Epigenetic Clocks" use DNAm changes at specific sites, which are highly correlated with chronological age, to estimate biological age. Differences between DNAm age and chronological age reflect age acceleration or deceleration, providing a measure of biological aging. Advancements in epigenetic clocks have introduced second-generation models like phenoAge, GrimAge2, and DunedinPACE (a DNA methylation biomarker of the pace of aging), which estimate mortality, life-span, and health-span (7-9). This study aims to assess the predictive capability of DNAm-based estimators of biological age for sporadic and recurrent miscarriage. Despite the clinical burden of both sporadic and RM, specific diagnostic biomarkers and candidate regulatory targets have not yet been identified and would be of great utility to obstetricians for early detection and management of patients at risk for miscarriage.

References cited:

1. The ESHRE Guideline Group on RPL and others, ESHRE guideline: recurrent pregnancy loss, Human Reproduction Open, Volume 2018, Issue 2, 2018, hoy004, https://doi.org/10.1093/hropen/hoy004
2. Laisk, T., Soares, A.L.G., Ferreira, T. *et al.* The genetic architecture of sporadic and multiple

consecutive miscarriage. *Nat Commun* **11**, 5980 (2020). https://doi.org/10.1038/s41467-020-19742-5
3. Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. BMC Med. 2013 Jun 26;11:154. doi: 10.1186/1741-7015-11-154. PMID: 23803387; PMCID: PMC3699442.
4. Laisk, T., Soares, A.L.G., Ferreira, T. *et al.* The genetic architecture of sporadic and multiple consecutive miscarriage. *Nat Commun* **11**, 5980 (2020). https://doi.org/10.1038/s41467-020-19742-5
5. Anvar Z, Chakchouk I, Demond H, Sharif M, Kelsey G, Van den Veyver IB. DNA Methylation Dynamics in the Female Germline and Maternal-Effect Mutations That Disrupt Genomic Imprinting. Genes (Basel). 2021 Aug 6;12(8):1214. doi: 10.3390/genes12081214. PMID: 34440388; PMCID: PMC8394515.
6. Pi L, Zhang Z, Gu Y, Wang X, Wang J, Xu J, Liu J, Zhang X, Du J. DNA methylation profiling in recurrent miscarriage. PeerJ. 2020 Jan 7;8:e8196. doi: 10.7717/peerj.8196. PMID: 31938574; PMCID: PMC6953351.
7. Daniel W Belsky, Avshalom Caspi, David L Corcoran, Karen Sugden, Richie Poulton, Louise Arseneault, Andrea Baccarelli, Kartik Chamarti, Xu Gao, Eilis Hannon, Hona Lee Harrington, Renate Houts, Meeraj Kothari, Dayoon Kwon, Jonathan Mill, Joel Schwartz, Pantel Vokonas, Cuicui Wang, Benjamin S Williams, Terrie E Moffitt (2022) DunedinPACE, a DNA methylation biomarker of the pace of aging eLife 11:e73420
8. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, Aviv A, Lohman K, Liu Y, Ferrucci L, Horvath S. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018 Apr 18;10(4):573-591. doi: 10.18632/aging.101414. PMID: 29676998; PMCID: PMC5940111.
9. Lu AT, Binder AM, Zhang J, Yan Q, Reiner AP, Cox SR, Corley J, Harris SE, Kuo PL, Moore AZ, Bandinelli S, Stewart JD, Wang C, Hamlat EJ, Epel ES, Schwartz JD, Whitsel EA, Correa A, Ferrucci L, Marioni RE, Horvath S. DNA methylation GrimAge version 2. Aging (Albany NY). 2022 Dec 14;14(23):9484-9549. doi: 10.18632/aging.204434. Epub 2022 Dec 14. PMID: 36516495; PMCID: PMC9792204.

Data Security Agreement

Provisional Paper Title: Identifying DNA methylated-based predictors of miscarriage risk
Proposing Author: Trevor Doherty and Dr Therese Murphy
Today's Date: 8/28/2023

☒	I am current on Human Subjects Training (CITI (www.citiprogram.org) or equivalent)
☒	My project is covered by the Duke ethics committee OR I have /will obtain ethical approval from my home institution.
☒	I will treat all data as "restricted" and store in a secure fashion. My computer or laptop is: a) encrypted (recommended programs are FileVault2 for Macs, and Bitlocker for Windows machines) b) password-protected c) configured to lock-out after 15 minutes of inactivity AND d) has an antivirus client installed as well as being patched regularly.

☒	I will not "sync" the data to a mobile device.
☒	In the event that my laptop with data on it is lost, stolen or hacked, I will immediately contact Moffitt or Caspi.
☒	I will not share the data with anyone, including my students or other collaborators not specifically listed on this concept paper.
☒	I will not post data online or submit the data file to a journal for them to post. *Some journals are now requesting the data file as part of the manuscript submission process. Study participants have not given informed consent for unrestricted open access, so we have a managed-access process. Speak to Temi or Avshalom for strategies for achieving compliance with data-sharing policies of journals.*
☒	I will delete all data files from my computer after the project is complete. Collaborators and trainees may not take a data file away from the office. This data remains the property of the Study and cannot be used for further analyses without an approved concept paper for new analyses.
☒	I have read the Data Use Guidelines and agree to follow the instructions.

Signature: